WOMEN IN MICROBIOLOGY

EDITED BY: Rachel Ann Foster, Sonja-Verena Albers, Ludmila Chistoserdova, Virginia P. Edgcomb, M. Pilar Francino, Sara Hallin, Jana Seifert, Giovanna Suzzi, Sabine Dagmar Zimmermann, Colleen Hansel and Henrietta Venter

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

WOMEN IN MICROBIOLOGY

Topic Editors:
Rachel Ann Foster, Stockholm University, Sweden
Sonja-Verena Albers, University of Freiburg, Germany
Ludmila Chistoserdova, University of Washington, United States
Virginia P. Edgcomb, Woods Hole Oceanographic Institution, United States
M. Pilar Francino, Fundación para el Fomento de la Investigación Sanitaria y
Biomédica de la Comunitat Valenciana (FISABIO), Spain
Sara Hallin, Swedish University of Agricultural Sciences, Sweden
Jana Seifert, University of Hohenheim, Germany
Giovanna Suzzi, University of Teramo, Italy
Sabine Dagmar Zimmermann, Délégation Languedoc Roussillon (CNRS), France
Colleen Hansel, Woods Hole Oceanographic Institution, United States
Henrietta Venter, University of South Australia, Australia

Table of Contents

The Antimicrobial Activity of the AGXX® Surface Coating Requires a Small Particle Size to Efficiently Kill *Staphylococcus aureus*

Nico Linzner and Haike Antelmann*

Freie Universität Berlin, Institute for Biology-Microbiology, Berlin, Germany

Methicillin-resistant *Staphylococcus aureus* (MRSA) isolates are often resistant to multiple antibiotics and pose a major health burden due to limited treatment options. The novel AGXX® surface coating exerts strong antimicrobial activity and successfully kills multi-resistant pathogens, including MRSA. The mode of action of AGXX® particles involves the generation of reactive oxygen species (ROS), which induce an oxidative and metal stress response, increased protein thiol-oxidations, protein aggregations, and an oxidized bacillithiol (BSH) redox state in *S. aureus*. In this work, we report that the AGXX® particle size determines the effective dose and time-course of *S. aureus* USA300JE2 killing. We found that the two charges AGXX®373 and AGXX®383 differ strongly in their effective concentrations and times required for microbial killing. While 20–40 µg/ml AGXX®373 of the smaller particle size of 1.5–2.5 µm resulted in >99.9% killing after 2 h, much higher amounts of 60–80 µg/ml AGXX®383 of the larger particle size of >3.2 µm led to a >99% killing of *S. aureus* USA300JE2 within 3 h. Smaller AGXX® particles have a higher surface/volume ratio and therefore higher antimicrobial activity to kill at lower concentrations in a shorter time period compared to the larger particles. Thus, in future preparations of AGXX® particles, the size of the particles should be kept at a minimum for maximal antimicrobial activity.

Keywords: *Staphylococcus aureus*, AGXX®, metal particles, antimicrobial activity, contact killing

INTRODUCTION

The increasing prevalence of antibiotic resistant strains in hospitals and the community poses a major burden to human health and limits treatment options of life-threatening bacterial infections (Christaki et al., 2020). Of particular importance are ESKAPE pathogens, such as multi-resistant *Staphylococcus aureus* isolates, which can cause severe systemic infections and acquire quickly new resistance elements through horizontal gene transfer (Mulani et al., 2019; De Oliveira et al., 2020). Promising alternative antimicrobial compounds could be reactive oxygen species (ROS) producing antimicrobials, which target rather non-specifically multiple cellular targets, such as proteins, lipids, and nucleic acids and therefore can eliminate drug-resistant bacteria (Lewis, 2013; Vatansever et al., 2013).

Metals, like silver (Ag^+) and copper (Cu^{2+}), have been used since ancient times for medication of bacterial infections and were shown to exert their microbicidal activity *via* ROS generation and protein and membrane damage (Lemire et al., 2013; Mijnendonckx et al., 2013). ROS generated from Ag^+ has been shown to damage the bacterial cell envelope by disruption of the peptidoglycan cell wall, lipoteichoic acids, and the phosphatidylethanolamine membrane lipids (Gunawan et al., 2020). Furthermore, silver condensed the DNA and caused protein damage *via* its interaction with protein thiols, the release of Fe^{2+} from FeS clusters, or by mismetallation of Zn^{2+}-containing proteins (Barras et al., 2018). However, the widespread use of silver in the treatment of wounds and burns has selected for various silver resistance mechanisms (Gupta et al., 1999; Silver, 2003; Atiyeh et al., 2007; Mijnendonckx et al., 2013). To circumvent the problem of silver resistance, silver nanoparticles were produced, which act as antimicrobials *via* the release of silver ions (Xiu et al., 2012). The application of silver ions and nanoparticles has raised concerns due to their toxic effects in eukaryotic cells (Mijnendonckx et al., 2013). Silver was shown to precipitate in various tissues and organs in the mice, including the kidney, liver, jejunum, colon, and the brain (Boudreau et al., 2016; Recordati et al., 2021). Neurotoxic effects of silver have been reported, including argyria, which are rare cases of irreversible gray pigmentations of the skin and the eyes due to silver sulfide precipitates (Lansdown, 2010; Mijnendonckx et al., 2013; Recordati et al., 2021).

Recently, metal-based nanoenzymes, such as the oxidase-like silver-palladium bimetallic alloy nanocage $AgPd_{0.38}$ were shown to produce ROS at the surface and selectively kill drug-resistant bacteria, but did not show toxicity in mammalian cells (Gao et al., 2021). In addition, very promising metal-based antimicrobial surface coatings were developed, including AGXX® and the combination of functionalized graphene oxide (GOX) with AGXX®, termed as GOX-AGXX® (Landau, 2013; Guridi et al., 2015; Vaishampayan et al., 2021). AGXX® is composed of Ag^+ and ruthenium (Ru^+), which form an electric field *via* two redox cycles, leading to electron transfer to molecular oxygen and ROS generation (**Figure 1**; Guridi et al., 2015; Clauss-Lendzian et al., 2018). In the first cycle, elementary Ag^0 is oxidized to Ag^+, which is regenerated by bacterial reducing pathways, such as the thioredoxin (Trx)/ thioredoxin reductase (TrxR) system. In the second cycle, higher valent Ru^{x+1} is reduced to Ru^x, and re-oxidized to Ru^{x+1} with subsequent ROS formation, such as superoxide anion, H_2O_2, and the highly toxic OH•. These redox cycles lead to a constant regeneration and ensure longevity and sustainability of the antimicrobial metal coating (Landau, 2013; Guridi et al., 2015). Moreover, AGXX® is predicted to cause low toxicity in human cells, since it releases only small amounts of 0.1–0.2 mg/l of silver ions after 12 weeks of exposure in distilled water (Guridi et al., 2015), although detailed toxicity studies are lacking.

The novel combined GOX-AGXX® coating acts *via* a "catch and kill" mechanism to facilitate bacterial killing (Vaishampayan et al., 2021). GOXs are oxidized graphite sheets, which are grafted with polycationic polymer chains and bind the negatively-charged cell envelope of bacteria (Vaishampayan et al., 2021). Thus, GOX captures bacteria *via* electrostatic attractions, leading to inhibition of bacterial proliferation as bacteriostatic effect (Vaishampayan et al., 2021). The second material AGXX® kills the GOX-captured bacteria *via* the described ROS formation.

AGXX® can be electroplated on various materials, including plastics, steel, glass, ceramics, fleeze, and cellulose fibers. AGXX® coatings are used for sterilization of medical implants, catheters, and plasters as well as in industrial water pipelines to successfully eradicate bacteria, which are in close contact with the AGXX® surface by a process termed as "contact killing" (Maillard and Hartemann, 2013; Guridi et al., 2015; Villapún et al., 2016; Clauss-Lendzian et al., 2018; Vaishampayan et al., 2021).

AGXX® acts as promising broad-spectrum antimicrobial and revealed strong bactericidal effects against various multidrug resistant pathogens, such as *S. aureus*, *Staphylococcus epidermidis*, *Escherichia coli*, *Enterococcus faecalis*, *Enterococcus faecium*, and *Legionella erythra* (Landau, 2013; Guridi et al., 2015; Heiss et al., 2017; Clauss-Lendzian et al., 2018; Loi et al., 2018; Vaishampayan et al., 2018). In addition, AGXX® and GOX-AGXX® both inhibit biofilm formation, which was supported by the downregulation of virulence factor regulons (e.g., SaeRS and Agr) and genes for biofilm formation in *S. aureus* (Vaishampayan et al., 2018, 2021). RNAseq analyses further revealed that AGXX®373 causes an oxidative and metal stress response as well as proteotoxic effects in *S. aureus* (Loi et al., 2018). AGXX® treatment leads to protein-*S*-bacillithiolation of GapDH, increased protein aggregations and an oxidative shift in the bacillithiol (BSH) redox potential, supporting ROS generation as its main mode of action (**Figure 1**; Loi et al., 2018). We have further shown that the low molecular weight thiol BSH and the HypR-regulated flavin disulfide reductase MerA function in the defense against AGXX® in *S. aureus* (Loi et al., 2018; Linzner et al., 2021). Furthermore, the inductions of heat-shock specific proteases and chaperones and antioxidant responses were observed after AGXX® and GOX-AGXX® treatment in *E. faecium* and *S. aureus* (Clauss-Lendzian et al., 2018; Vaishampayan et al., 2021).

While global transcriptomic analyses revealed insights into the mode of action of these metal coatings in bacterial pathogens, the question arises how the efficiency can be improved for complete killing. Previous studies revealed that the antimicrobial activity of metal nanoparticles depends on the particle size (Raghupathi et al., 2011; Kadiyala et al., 2018). In our studies, we found that the two charges, AGXX®373 and AGXX®383, differ strongly in their antimicrobial activities, which correlates with their particle sizes and affects the efficient concentrations and times for *S. aureus* killing. While 20–40 µg/ml AGXX®373 with particle size of 1.5–2.5 µm kills *S. aureus* completely within 2 h, 2–3-fold higher amounts of 60–80 µg/ml AGXX®383 with particle size of >3.2 µm were required for the same extent of *S. aureus* killing. Thus, small AGXX® particles with a higher surface area are more efficient antimicrobials and kill bacteria at lower concentrations compared to larger size particles.
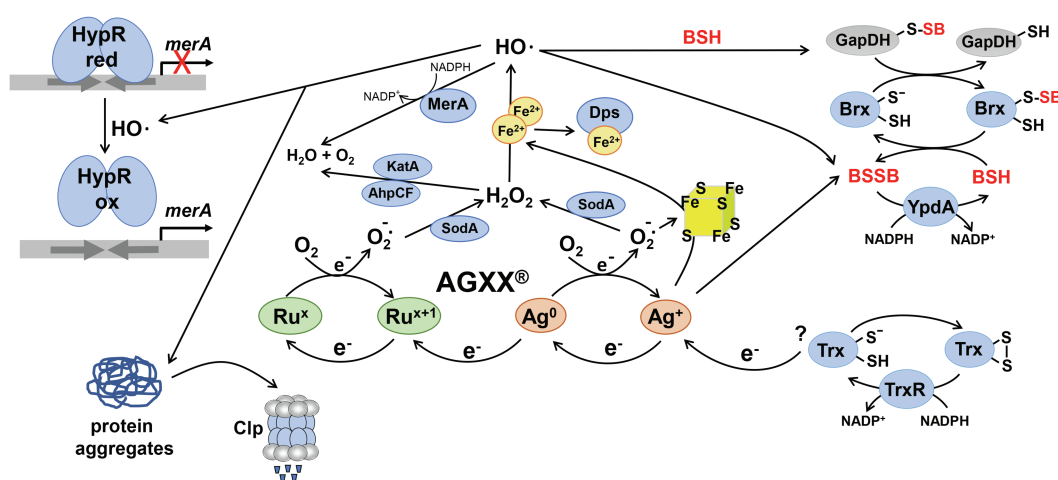
FIGURE 1 | The proposed mode of action of AGXX® and the responses in *Staphylococcus aureus*. AGXX® is composed of silver (Ag^+) and ruthenium (Ru^x), which are connected by two redox cycles and form an electric field, leading to electron transfer to molecular oxygen (O_2), with subsequent reactive oxygen species (ROS) generation, such as superoxide anion $O_2 \bullet^-$, hydrogen peroxide (H_2O_2), and hydroxyl radicals (OH•). First, Ag^0 is oxidized to Ag^+, which is back-reduced to Ag^0 possibly by electrons from cellular donors, such as the thioredoxin (Trx)/thioredoxin reductase (TrxR) system. Second, Ru^{x+1} is reduced by electrons from Ag^0, leading to formation of Ru^x, which is re-oxidized to Ru^{x+1}. ROS are generated in the AGXX® redox cycles upon oxidation of Ag^0 and Ru^x. Furthermore, Ag^+ ions and $O_2 \bullet^-$ have been described to damage FeS clusters and release Fe^{2+}, which potentiates OH• generation *via* the Fenton chemistry. In *S. aureus*, AGXX® induces antioxidant enzymes, such as catalases, peroxiredoxins, and superoxide dismutases for detoxification of ROS (Loi et al., 2018). AGXX® leads to thiol-oxidation of the HypR repressor, resulting in upregulation of the HypR-controlled flavin disulfide reductase MerA, which probably detoxifies OH•. In addition, AGXX® exposure resulted in increased protein *S*-bacillithiolation of GapDH and an oxidative shift of the BSH redox potential, supporting an impaired thiol-redox homeostasis. GapDH-SSB can be de-bacillithiolated in the Brx/BSH/YpdA redox pathway to regenerate GapDH. Consequently, BSH and MerA were shown to protect *S. aureus* against AGXX® toxicity (Loi et al., 2018). In addition, AGXX® causes oxidative protein unfolding and protein aggregates, which are degraded by the Clp protease complex. The figure is modified from (Guridi et al., 2015) including previous results from (Loi et al., 2018).

MATERIALS AND METHODS

Preparation of AGXX® Microparticles

The AGXX®373 and AGXX®383 particles were produced by Largentec GmbH (Berlin, Germany) using different silver powders with the particle sizes of 1.5–2.5 µm (MaTecK, Germany) and >3.2 µm (Toyo, Japan) as described previously (Loi et al., 2018). Briefly, both silver powders were chemically coated with ruthenium in alkaline medium, where the Ru(III) ions were first oxidized by NaOCl to RuO_4 (Guridi et al., 2015; Heiss et al., 2017). Reduction of RuO_4 to Ru or RuO_x was performed by addition of $NaNO_2$ as described (Chen et al., 2010). The black AGXX® powder was further conditioned with 50 mM ascorbic acid for 2 h, followed by filtration, washing with deionized water and drying.

Bacterial Cultivation and Survival Assays

The multi-resistant *S. aureus* USA300JE2 strain (Fey et al., 2013) was used for the AGXX® survival experiments and cultivated in RPMI medium as described (Fritsch et al., 2020). For survival assays under AGXX® stress, *S. aureus* USA300JE2 was grown in RPMI medium to an optical density at 500 nm (OD_{500}) of 0.5 and exposed to 10–40 µg/ml AGXX®373 and 40–100 µg/ml AGXX®383 microparticles as indicated in the figure legends. Ten microliter of serial dilutions of the cultures was spotted onto Luria Bertani (LB) agar plates and colony

forming units (CFUs) monitored after overnight incubation at 37°C. For quantification, 100 µl of serial dilutions was plated onto LB agar plates and CFUs counted after overnight incubation. The survival rates were calculated of the treated cultures in comparison to the untreated control culture at an OD_{500} of 0.5. The survival of the untreated control was set to 100%.

RESULTS

The Smaller AGXX®373 Particles Show Higher Antimicrobial Activities Toward *S. aureus* USA300JE2 in Comparison to Larger Sized AGXX®383 Particles

Previously, we analyzed the oxidative mode of action of AGXX®373 in *S. aureus* using 5 µg/ml AGXX® as sub-lethal concentration, which inhibits the growth (Loi et al., 2018). In this study, we were interested in the applied aspects of the AGXX® coating. We investigated the time-course required for *S. aureus* killing upon challenge with the two different charges AGXX®373 and AGXX®383, which varied in their particle size. These AGXX®373 and AGXX®383 charges were generated from different silver powders obtained from the companies MaTeck (Germany) and Toyo (Japan), respectively. While AGXX®373 particles had a small particle size of 1.5–2.5 µm, the size of AGXX®383 particles was in the range of >3.2 µm.

To investigate the impact of the particle size and the exposure time, we performed various survival assays of *S. aureus* cultures exposed to 10–40 μg/ml of AGXX®373 and 40–100 μg/ml AGXX®383, respectively, in a time-dependent manner (**Figures 2, 3**). The spotted survival assays revealed that 20–40 μg/ml AGXX®373 were highly toxic, resulting in microbial killing after 2 h, while 50–100 μg/ml AGXX®383 were required to kill *S. aureus* to a similar extent within 3 h (**Figures 2A,B**). For comparison, a similar CFU reduction was observed after treatment of *S. aureus* with 30 μg/ml AGXX®373 for 2 h as with 60 μg/ml AGXX®383 for 3 h, pointing to a >2-fold increased efficiency of the smaller AGXX®373 particles. In addition, the *S. aureus* viability rate decreased continuously over time with both AGXX® charges, leading to almost complete killing after 4 h with 20–40 μg/ml AGXX®373 and 60–100 μg/ml AGXX®383 (**Figures 2A,B**). These data indicate that the AGXX® particle size and the exposure time affect significantly its antimicrobial activity.

Next, the survival rates of *S. aureus* cells were quantified after different times of exposure to AGXX®373 or AGXX®383 in more detail using CFU counting. In agreement with the above results, treatment of *S. aureus* with 20–40 μg/ml AGXX®373 resulted in a time-depending killing of 99.9% cells within 2 h and almost 99.999% of cells lost their viability after 4 h (**Figures 3A,B**; **Supplementary Table S1**). However, much higher concentrations of 60–100 μg/ml AGXX®383 were required to achieve a >99% killing efficiency after 3 h and >99.99%

after 4 h (**Figures 3C,D**). While only 0.02% of *S. aureus* cells survived after 2 h exposure to 30 μg/ml AGXX®373, the survival was only slightly reduced to ~60% after 2 h treatment with 60 μg/ml AGXX®383 (**Figures 3A,C**). Furthermore, the decrease in bacterial viability was significant with 10–40 μg/ml AGXX®373 over the entire 1–4 h time course, whereas 50–60 μg/ml AGXX®383 did not show significant effects on *S. aureus* survival within 1–2 h. However, long-term exposure of *S. aureus* to 60–100 μg/ml AGXX®383 for 8–24 h led to a nearly 100% killing of cells (**Figures 3C,D**). Starting with a cell count of ~10^8 of the log phase bacteria, the CFU dropped to ~10^3 after 6 h, was further decreased to 23–63 colonies after 12 h and to 0–140 colonies after 24 h of exposure to 60–100 μg/ml AGXX®383 (**Figure 3D; Supplementary Table S2**). Importantly, almost complete killing of *S. aureus* was observed after exposure to 50–60 and 100 μg/ml AGXX®383, while the CFUs were slightly higher after treatment with 80 μg/ml AGXX®383 for unknown reasons.

In conclusion, our results confirmed that both AGXX®373 and AGXX®383 charges act strongly bactericidal and kill *S. aureus* USA300JE2 efficiently and nearly complete in a time- and dose-dependent manner. However, the concentrations and exposure times required for complete killing of *S. aureus* cultures were 2-3-fold higher for the larger size AGXX®383 particles compared to AGXX®373. Thus, the increased particle size of AGXX®383 led to a decreased antimicrobial activity compared to the smaller AGXX®373 particles. Taken together, our results
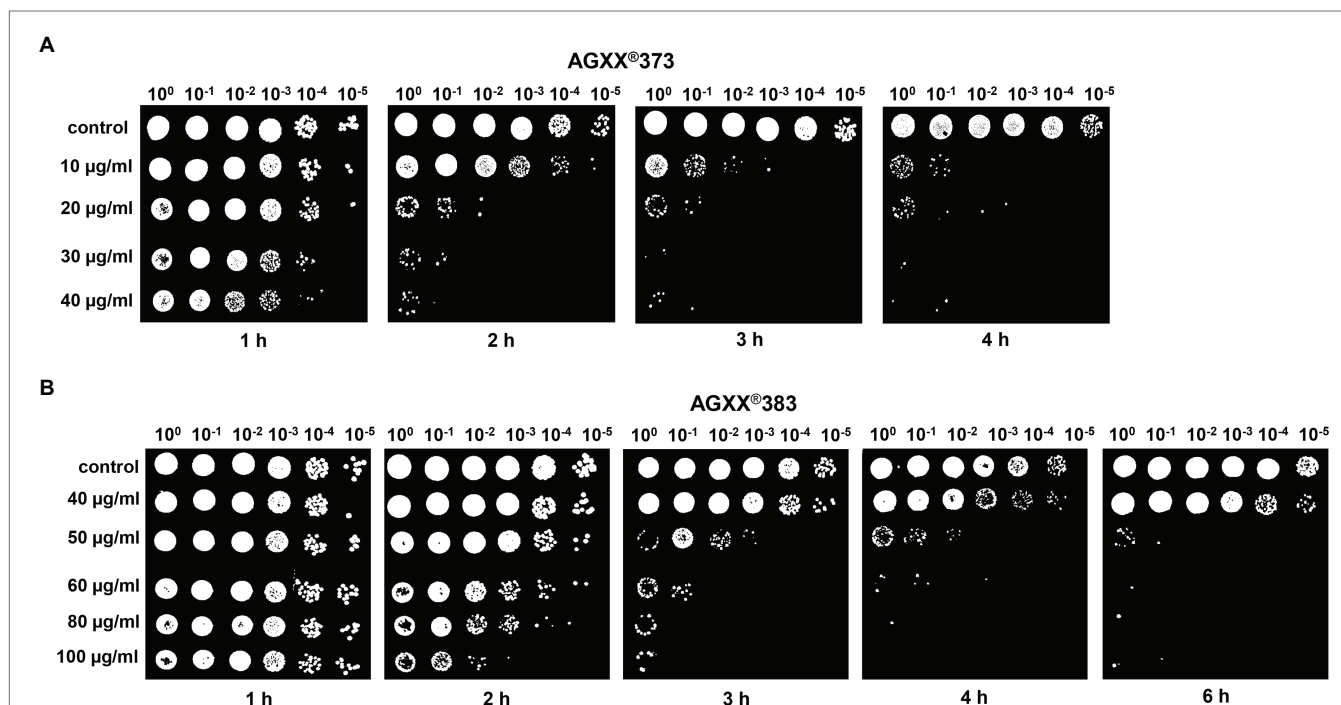
FIGURE 2 | The charges AGXX®373 **(A)** and AGXX®383 **(B)** cause different antimicrobial effects in *S. aureus* USA300JE2. Survival assays of *S. aureus* USA300JE2 cells were performed at an OD_{500} of 0.5 after exposure to 10–40 μg/ml AGXX®373 **(A)** or 40–100 μg/ml AGXX® 383 **(B)**. Serial dilutions of the cultures were spotted after 1–6 h of AGXX® exposure onto Luria Bertani (LB) agar plates to observe colony forming units (CFUs) after 24 h incubation. The presented results are representatives of three biological replicates. The results indicate that AGXX®373 particles have a stronger killing effect compared to AGXX®383 particles in *S. aureus*.
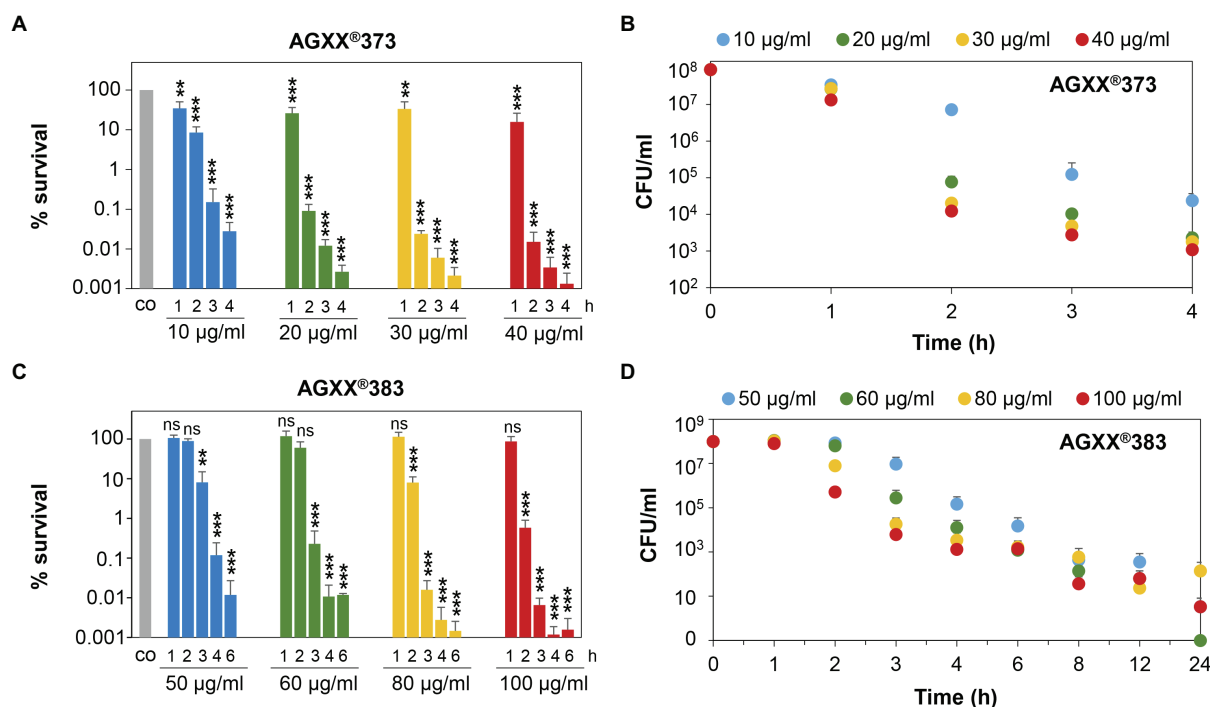
FIGURE 3 | Survival assays indicate 2-3-fold increased killing efficiencies of smaller AGXX®373 versus larger AGXX®383 particles in *S. aureus*. *S. aureus* USA300JE2 was exposed to 10–40 µg/ml AGXX®373 particles **(A,B)** and 50–100 µg/ml AGXX®383 particles **(C,D)** at an OD_{500} of 0.5. After 1–24 h of AGXX® treatment, CFUs were determined from serial dilutions plated overnight on LB agar plates. The survival ratios of the CFUs after AGXX® stress were calculated relative to the untreated control (co), which was set to 100%. Mean values of 3–5 biological replicates are presented, and error bars represent the SD. For statistical analyses of the killing effect, the 100% survival of the untreated control (co) was compared to 10, 20, 30, and 40 µg/ml AGXX®373 **(A)** or 50, 60, 80, and 100 µg/ml AGXX®383 **(C)**, respectively, and the calculations were performed using a Student's unpaired two-tailed *t*-test by the graph prism software. Symbols are *ns* > 0.05, **$p \leq 0.01$, and ***$p \leq 0.001$.

indicate that a low particle size and long exposure time are crucial for maximal killing efficiency of the AGXX® powder.

DISCUSSION

In the present study, we have quantified the killing effect of different particles sizes of the AGXX® antimicrobial surface coating in *S. aureus*. We used the smaller AGXX®373 and larger AGXX®383 particle charges to analyze time-dependent killing effects in survival assays. Our results showed that both AGXX® particles had strong antimicrobial activities, which were dependent on the particle size, exposure time, and the concentration. Overall, the antimicrobial efficiency of AGXX® particles was most efficient with high concentrations, long exposure times, and smaller particles sizes. In comparison, we would suggest to use charge AGXX®373 for future applications, due to its small particle size of 1.5–2.5 µm and its strong killing effect with low concentrations of 20–40 µg/ml. Much higher concentrations of 60–100 µg/ml were required to obtain the same killing efficiency with AGXX®383, which had a particle size of ~3.4 µm. Thus, the particle size has a strong impact on the antimicrobial activity and efficiency of the AGXX®

coating. Similar connections between the particle size and the antimicrobial activity were previously obtained for other metal nanoparticles, such as 4,6-diamino-2-pyrimidine thiol (DAPT)-capped gold nanoparticles, zinc oxide, and silver nanoparticles (Raghupathi et al., 2011; Lu et al., 2013; Zheng et al., 2021). The antimicrobial activities of the larger gold nanoparticles with the size of >5.5 nm toward *E. coli*, *Pseudomonas aeruginosa*, and *Klebsiella pneumoniae* were significantly lower compared to nanoparticles of 1.8–5.5 nm (Zheng et al., 2021). Furthermore, the antibacterial activities of smaller zinc oxide and silver nanoparticles were higher against *S. aureus* and *Streptococcus mutans*, respectively, compared to larger size nanoparticles (Raghupathi et al., 2011; Lu et al., 2013). The impact of the particle size on the antimicrobial activity can be explained by the higher surface/volume ratio of smaller particles, which provide an increased surface for "contact killing" of microbes by the metal coating. The term "contact killing" indicates that bacteria and yeast cells are rapidly killed on metallic surfaces, such as copper and silver alloys by their close contact with the metal surface, causing damage of the cellular envelope and macromolecules (Grass et al., 2011; Quaranta et al., 2011). The probability that bacteria come in close contact with the metal particles increases with a larger surface/volume

ratio, leading to an increased antimicrobial activity (Kadiyala et al., 2018). Thus, our results support previous data on other metal nanoparticles and their antimicrobial activity, which strongly depend on the particle size. Given the wide range of applications of the AGXX® surface coating in medicine, agriculture, and industries, it is very important to keep the size of metal particles as low as possible and reproducible to achieve efficient antimicrobial activities.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

NL designed and performed the experiments, analyzed the data, and wrote the manuscript draft. NL and HA revised and approved the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.731564/full#supplementary-material

REFERENCES

Atiyeh, B. S., Costagliola, M., Hayek, S. N., and Dibo, S. A. (2007). Effect of silver on burn wound infection control and healing: review of the literature. *Burns* 33, 139–148. doi: 10.1016/j.burns.2006.06.010

Barras, F., Aussel, L., and Ezraty, B. (2018). Silver and antibiotic, new facts to an old story. *Antibiotics* 7:79. doi: 10.3390/antibiotics7030079

Boudreau, M. D., Imam, M. S., Paredes, A. M., Bryant, M. S., Cunningham, C. K., Felton, R. P., et al. (2016). Differential effects of silver nanoparticles and silver ions on tissue accumulation, distribution, and toxicity in the Sprague Dawley rat following daily oral gavage administration for 13 weeks. *Toxicol. Sci.* 150, 131–160. doi: 10.1093/toxsci/kfv318

Chen, J.-Y., Wang, L.-Y., and Wu, P.-W. (2010). Preparation and characterization of ruthenium films via an electroless deposition route. *Thin Solid Films* 518, 7245–7248. doi: 10.1016/j.tsf.2010.04.086

Christaki, E., Marcou, M., and Tofarides, A. (2020). Antimicrobial resistance in bacteria: mechanisms, evolution, and persistence. *J. Mol. Evol.* 88, 26–40. doi: 10.1007/s00239-019-09914-3

Clauss-Lendzian, E., Vaishampayan, A., De Jong, A., Landau, U., Meyer, C., Kok, J., et al. (2018). Stress response of a clinical *Enterococcus faecalis* isolate subjected to a novel antimicrobial surface coating. *Microbiol. Res.* 207, 53–64. doi: 10.1016/j.micres.2017.11.006

De Oliveira, D. M. P., Forde, B. M., Kidd, T. J., Harris, P. N. A., Schembri, M. A., Beatson, S. A., et al. (2020). Antimicrobial resistance in ESKAPE pathogens. *Clin. Microbiol. Rev.* 33, e00181–e00119. doi: 10.1128/CMR.00181-19

Fey, P. D., Endres, J. L., Yajjala, V. K., Widhelm, T. J., Boissy, R. J., Bose, J. L., et al. (2013). A genetic resource for rapid and comprehensive phenotype screening of nonessential *Staphylococcus aureus* genes. *MBio* 4, e00537–e00512. doi: 10.1128/mBio.00537-12

Fritsch, V. N., Loi, V. V., Busche, T., Tung, Q. N., Lill, R., Horvatek, P., et al. (2020). The alarmone (p)ppGpp confers tolerance to oxidative stress during the stationary phase by maintenance of redox and iron homeostasis in *Staphylococcus aureus*. *Free Radic. Biol. Med.* 161, 351–364. doi: 10.1016/j.freeradbiomed.2020.10.322

Gao, F., Shao, T., Yu, Y., Xiong, Y., and Yang, L. (2021). Surface-bound reactive oxygen species generating nanozymes for selective antibacterial action. *Nat. Commun.* 12:745. doi: 10.1038/s41467-021-20965-3

Grass, G., Rensing, C., and Solioz, M. (2011). Metallic copper as an antimicrobial surface. *Appl. Environ. Microbiol.* 77, 1541–1547. doi: 10.1128/AEM.02766-10

Gunawan, C., Faiz, M. B., Mann, R., Ting, S. R. S., Sotiriou, G. A., Marquis, C. P., et al. (2020). Nanosilver targets the bacterial cell envelope: the link with generation of reactive oxygen radicals. *ACS Appl. Mater. Interfaces* 12, 5557–5568. doi: 10.1021/acsami.9b20193

Gupta, A., Matsui, K., Lo, J. F., and Silver, S. (1999). Molecular basis for resistance to silver cations in salmonella. *Nat. Med.* 5, 183–188. doi: 10.1038/5545

Guridi, A., Diederich, A. K., Aguila-Arcos, S., Garcia-Moreno, M., Blasi, R., Broszat, M., et al. (2015). New antimicrobial contact catalyst killing antibiotic resistant clinical and waterborne pathogens. *Mater. Sci. Eng. C Mater. Biol. Appl.* 50, 1–11. doi: 10.1016/j.msec.2015.01.080

Heiss, A., Freisinger, B., and Held-Fohn, E. (2017). Enhanced antibacterial activity of silver-ruthenium coated hollow microparticles. *Biointerphases* 12:05G608. doi: 10.1116/1.5003803

Kadiyala, U., Kotov, N. A., and Vanepps, J. S. (2018). Antibacterial metal oxide nanoparticles: challenges in interpreting the literature. *Curr. Pharm. Des.* 24, 896–903. doi: 10.2174/1381612824666180219130659

Landau, U. (2013). AGXX - Eine nachhaltige Lösung für die Entkeimung wässriger Lösungen. *Galvanotechnik* 11, 2169–2184.

Lansdown, A. B. (2010). A pharmacological and toxicological profile of silver as an antimicrobial agent in medical devices. *Adv. Pharmacol. Sci.* 2010:910686. doi: 10.1155/2010/910686

Lemire, J. A., Harrison, J. J., and Turner, R. J. (2013). Antimicrobial activity of metals: mechanisms, molecular targets and applications. *Nat. Rev. Microbiol.* 11, 371–384. doi: 10.1038/nrmicro3028

Lewis, K. (2013). Platforms for antibiotic discovery. *Nat. Rev. Drug Discov.* 12, 371–387. doi: 10.1038/nrd3975

Linzner, N., Loi, V. V., Fritsch, V. N., and Antelmann, H. (2021). Thiol-based redox switches in the major pathogen *Staphylococcus aureus*. *Biol. Chem.* 402, 333–361. doi: 10.1515/hsz-2020-0272

Loi, V. V., Busche, T., Preuss, T., Kalinowski, J., Bernhardt, J., and Antelmann, H. (2018). The AGXX® antimicrobial coating causes a thiol-specific oxidative stress response and protein *S*-bacillithiolation in *Staphylococcus aureus*. *Front. Microbiol.* 9:3037. doi: 10.3389/fmicb.2018.03037

Lu, Z., Rong, K., Li, J., Yang, H., and Chen, R. (2013). Size-dependent antibacterial activities of silver nanoparticles against oral anaerobic pathogenic bacteria. *J. Mater. Sci. Mater. Med.* 24, 1465–1471. doi: 10.1007/s10856-013-4894-5

Maillard, J. Y., and Hartemann, P. (2013). Silver as an antimicrobial: facts and gaps in knowledge. *Crit. Rev. Microbiol.* 39, 373–383. doi: 10.3109/1040841X.2012.713323

Mijnendonckx, K., Leys, N., Mahillon, J., Silver, S., and Van Houdt, R. (2013). Antimicrobial silver: uses, toxicity and potential for resistance. *BioMetals* 26, 609–621. doi: 10.1007/s10534-013-9645-z

Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S., and Pardesi, K. R. (2019). Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: a review. *Front. Microbiol.* 10:539. doi: 10.3389/fmicb.2019.00539

Quaranta, D., Krans, T., Espirito Santo, C., Elowsky, C. G., Domaille, D. W., Chang, C. J., et al. (2011). Mechanisms of contact-mediated killing of yeast cells on dry metallic copper surfaces. *Appl. Environ. Microbiol.* 77, 416–426. doi: 10.1128/AEM.01704-10

Raghupathi, K. R., Koodali, R. T., and Manna, A. C. (2011). Size-dependent bacterial growth inhibition and mechanism of antibacterial activity of zinc oxide nanoparticles. *Langmuir* 27, 4020–4028. doi: 10.1021/la104825u

Recordati, C., De Maglie, M., Cella, C., Argentiere, S., Paltrinieri, S., Bianchessi, S., et al. (2021). Repeated oral administration of low doses of silver in mice: tissue distribution and effects on central nervous system. *Part. Fibre Toxicol.* 18:23. doi: 10.1186/s12989-021-00418-x

Silver, S. (2003). Bacterial silver resistance: molecular biology and uses and misuses of silver compounds. *FEMS Microbiol. Rev.* 27, 341–353. doi: 10.1016/S0168-6445(03)00047-0

Vaishampayan, A., Ahmed, R., Wagner, O., De Jong, A., Haag, R., Kok, J., et al. (2021). Transcriptomic analysis of stress response to novel antimicrobial coatings in a clinical MRSA strain. *Mater. Sci. Eng. C Mater. Biol. Appl.* 119:111578. doi: 10.1016/j.msec.2020.111578

Vaishampayan, A., De Jong, A., Wight, D. J., Kok, J., and Grohmann, E. (2018). A novel antimicrobial coating represses biofilm and virulence-related genes in methicillin-resistant *Staphylococcus aureus*. *Front. Microbiol.* 9:221. doi: 10.3389/fmicb.2018.00221

Vatansever, F., De Melo, W. C., Avci, P., Vecchio, D., Sadasivam, M., Gupta, A., et al. (2013). Antimicrobial strategies centered around reactive oxygen species-bactericidal antibiotics, photodynamic therapy, and beyond. *FEMS Microbiol. Rev.* 37, 955–989. doi: 10.1111/1574-6976.12026

Villapún, V. M., Dover, L. G., Cross, A., and Gonzalez, S. (2016). Antibacterial metallic touch surfaces. *Materials* 9:736. doi: 10.3390/ma9090736

Xiu, Z. M., Zhang, Q. B., Puppala, H. L., Colvin, V. L., and Alvarez, P. J. (2012). Negligible particle-specific antibacterial activity of silver nanoparticles. *Nano Lett.* 12, 4271–4275. doi: 10.1021/nl301934w

Zheng, W., Jia, Y., Zhao, Y., Zhang, J., Xie, Y., Wang, L., et al. (2021). Reversing bacterial resistance to gold nanoparticles by size modulation. *Nano Lett.* 21, 1992–2000. doi: 10.1021/acs.nanolett.0c04451

Check for updates

Physiology of the Nitrite-Oxidizing Bacterium *Candidatus* Nitrotoga sp. CP45 Enriched From a Colorado River

*Munira A. Lantz, Andrew M. Boddicker, Michael P. Kain, Owen M. C. Berg, Courtney D. Wham and Annika C. Mosier**

Department of Integrative Biology, University of Colorado Denver, Denver, CO, United States

Nitrogen cycling microbes, including nitrite-oxidizing bacteria (NOB), perform critical ecosystem functions that help mitigate anthropogenic stresses and maintain ecosystem health. Activity of these beneficial nitrogen cycling microbes is dictated in part by the microorganisms' response to physicochemical conditions, such as temperature, pH, and nutrient availability. NOB from the newly described *Candidatus* Nitrotoga genus have been detected in a wide range of habitats across the globe, yet only a few organisms within the genus have been physiologically characterized. For freshwater systems where NOB are critical for supporting aquatic life, *Ca.* Nitrotoga have been previously detected but little is known about the physiological potential of these organisms or their response to changing environmental conditions. Here, we determined functional response to environmental change for a representative freshwater species of *Ca.* Nitrotoga (*Ca.* Nitrotoga sp. CP45, enriched from a Colorado river). The physiological findings demonstrated that CP45 maintained nitrite oxidation at pH levels of 5–8, at temperatures from 4 to 28°C, and when incubated in the dark. Light exposure and elevated temperature (30°C) completely halted nitrite oxidation. *Ca.* Nitrotoga sp. CP45 maintained nitrite oxidation upon exposure to four different antibiotics, and potential rates of nitrite oxidation by river sediment communities were also resilient to antibiotic stress. We explored the *Ca.* Nitrotoga sp. CP45 genome to make predictions about adaptations to enable survival under specific conditions. Overall, these results contribute to our understanding of the versatility of a representative freshwater *Ca.* Nitrotoga sp. Identifying the specific environmental conditions that maximize NOB metabolic rates may ultimately direct future management decisions aimed at restoring impacted systems.

Keywords: nitrite-oxidizing bacteria, freshwater, nitrification, water quality, nitrotoga, antibiotics, physiology, temperature

INTRODUCTION

Nitrite-oxidizing bacteria (NOB) play fundamental roles in maintaining the health and resilience of freshwater habitats by regulating nitrogen transformations. NOB are responsible for the second step of nitrification (oxidizing nitrite to nitrate) and have three key impacts on aquatic habitats. First, they are responsible for the formation of nitrate, which provides a critical nitrogen source for

microbial and plant assimilation. Second, they provide a substrate for denitrification, which results in the formation of gaseous dinitrogen for evolution out of the aquatic habitat. Third, they reduce nitrite toxicity to fish and other aquatic organisms (and indirectly reduce ammonia toxicity by consuming the end-product of microbial ammonia oxidation). If nitrite oxidation rates are reduced or halted due to environmental perturbation, nitrite concentrations will increase and negative effects, such as toxicity, hypoxia, and loss of biodiversity, may propagate through the ecosystem.

Nitrobacter and *Nitrospira* are often identified as the primary NOB present in freshwater systems (Hovanec et al., 1998; Féray et al., 1999; Altmann et al., 2003; Cébron and Garnier, 2005). Sequences from the recently discovered *Candidatus* Nitrotoga (Alawi et al., 2007) have now also been reported in a wide range of freshwater habitats, including a subglacial lake (Christner et al., 2014), wastewater treatment plants (WWTPs)/activated sludge (Alawi et al., 2009; Lücker et al., 2014; Kitzinger et al., 2018; Wegen et al., 2019), a groundwater cave system (Chen et al., 2009), biologically active filter for drinking water (White et al., 2012), the tidal reach area of a freshwater river (Fan et al., 2016), a freshwater aquaculture plant (Hüpeden et al., 2016), and rivers (Li et al., 2011; Boddicker and Mosier, 2018). *Ca.* Nitrotoga-like sequences range in relative abundance from ∼0.01 to 10% of the total bacterial community (which is similar to estimates for other NOB) across globally distributed freshwater habitats (Boddicker and Mosier, 2018). The relationship between *Ca.* Nitrotoga relative abundance, nitrite oxidation rates, and *in situ* nitrite concentrations remains understudied.

Despite the detection of *Ca.* Nitrotoga gene sequences in a variety of freshwater habitats, little is known about the physiological potential of these organisms or their response to changing environmental conditions. Across the entire genus, physiology data have only been reported for a small number of *Ca.* Nitrotoga cultures, which were enriched from permafrost, activated sludge, an aquaculture system, and coastal sand (Alawi et al., 2007, 2009; Hüpeden et al., 2016; Ishii et al., 2017, 2020; Kitzinger et al., 2018; Wegen et al., 2019). These physiological studies have indicated that optimal temperatures for *Ca.* Nitrotoga physiological activity range from 10° to 28°C and that *Ca.* Nitrotoga has been cultured at pH values ranging from 6.8 to 8.3. There are no prior reports on the physiology of freshwater *Ca.* Nitrotoga so it is unknown whether or not these NOB have similar activity to *Ca.* Nitrotoga enriched from other habitats.

Here, we determine the physiology of a newly identified *Ca.* Nitrotoga species from a natural freshwater system, *Ca.* Nitrotoga sp. CP45 (Boddicker and Mosier, 2018). In this study, we measure nitrite oxidation under varying conditions of light, pH, temperature, and antibiotic exposure to better characterize and predict environmental parameters that may lead to physiological activity/inactivity within freshwater systems. We explore the CP45 genome to make predictions about which genes may be involved in the physiological responses that were observed. Physiology and genomic potential of *Ca.* Nitrotoga sp. CP45 is directly relevant to Colorado ecosystems, but can also be viewed as more broadly applicable to other freshwater systems since

CP45 is closely related to other freshwater *Ca.* Nitrotoga sp. found across the globe (based on prior gene sequence analyses; Boddicker and Mosier, 2018). Determining the physiological limits of the newly described *Ca.* Nitrotoga sp. can improve our understanding of how these bacteria respond under natural and stressed conditions in the environment.

MATERIALS AND METHODS

Cultivation and Genomic Sequencing of *Ca.* Nitrotoga sp. CP45

Ca. Nitrotoga sp. CP45 was previously enriched from water column samples collected from the Cache la Poudre River near Greeley, CO (site CP45; latitude 40.41774, longitude −104.64017). Briefly, cultures were grown in Freshwater Nitrite Oxidizer Medium (FNOM) (Boddicker and Mosier, 2018) with 0.3 mM nitrite and incubated at room temperature (∼23°C) in the dark. The cultures were enriched with NOB from the *Ca.* Nitrotoga genus (based on PCR and amplicon sequencing; *Ca.* Nitrotoga was 16–24% enriched at the time of genomic sequencing, and approximately 65% enriched upon further purification) (Boddicker and Mosier, 2018). The *Ca.* Nitrotoga sp. CP45 genome was previously sequenced and annotated (Boddicker and Mosier, 2018). Genome sequence and phylogenetic analyses showed that the NOB in the culture represented a new species within the *Ca.* Nitrotoga genus (Boddicker and Mosier, 2018). Here, further annotation of genes specific to processes described in this study was conducted using the eggNOG database and mapper (Huerta-Cepas et al., 2017, 2019), BacMet2 (Pal et al., 2014), virulence factor database (VFDB; Liu et al., 2019), Interpro (Blum et al., 2021), and BLAST (Altschul et al., 1990).

Measurement of Nitrite Concentrations

NOB activity in culture is commonly measured by following rates of substrate (nitrite) utilization or product (nitrate) formation (Prosser, 1989). Here, cultures were regularly monitored for nitrite utilization using a Griess nitrite color reagent composed of 10 g sulfanilamide (SULF), 1.0 g N-(1-napthyl) ethylenediamine dihydrochloride (NEDD), 100 ml 85% phosphoric acid, and MilliQ water to a final volume of 1 L (Griess-Romijn van Eck, 1966; American Public Health Association, American Water Works Association, and Water Environment Federation, 2017). The solution was stored in the dark at 4°C for up to 1 month in an amber bottle wrapped in aluminum foil. Griess nitrite color reagent was mixed with culture samples at a 1:1 ratio. Nitrite concentrations were determined by quantitative spectrophotometric measurements at 543 nm according to *Standard Methods*, method 4,500-NO_2^- B Colorimetric Method (American Public Health Association, American Water Works Association, and Water Environment Federation, 2017) using a BioTek Synergy HT plate reader (BioTek, Winooski, VT). Pseudoreplicates ($\times$ 2) were measured for each individual nitrite sample. Nitrite concentrations present in sampled media were calculated using resultant end point absorbance readings and associated nitrite standards (0–0.3 mM NO_2^-) prepared in sterile

FNOM. In all, more than 10,000 individual nitrite measurements were made for the experiments described below.

Growth Curve

Four large batches of culture (700 ml FNOM with 7 ml *Ca.* Nitrotoga sp. CP45 inoculum) were incubated at 23°C and grown over the course of 5 days. Nitrite was measured from subsamples collected approximately every 6 h. At select time points, 10 ml of each culture was filtered through 0.2-µm Supor 200 filters (Pall, New York, NY) and immediately frozen at −80°C. DNA was extracted as previously described (Boddicker and Mosier, 2018).

Nitrite concentrations were measured as described above. Nitrate + nitrite (NOx) concentrations were determined using Vanadium (III) Chloride in hydrochloric acid combined with the Griess reagents (SULF and NEDD) (Miranda et al., 2001). Following color development, absorbance was read as described above with pseudoreplicates ($\times$ 2) for each sample. NOx concentrations were calculated using resultant end point absorbance readings and associated standards (0–0.3 mM NO_3^-) prepared in sterile FNOM. Nitrate was determined by subtracting nitrite values from the NOx values.

Extracted DNA was amplified in triplicate quantitative PCR (qPCR) reactions using *Ca.* Nitrotoga 16S rRNA gene-specific primers NTG200F and NTG840R (Alawi et al., 2007). Reactions were carried out using a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA). Each 20-µl reaction included 10 µl of FailSafe Premix E (Lucigen, Madison, WI) [with 20X SYBR Green Nucleic Acid Dye (Invitrogen, Carlsbad, CA) and 0.2% Tween-20], 5.95 µl of nuclease-free water, 0.4 µl of 25 µM ROX Passive Reference Dye (Invitrogen, Carlsbad, CA), 1.6 µl of 25 mM $MgCl_2$, 0.4 µl of 10 µM forward and reverse primers, 0.05 µl of 50 mg/ml BSA, and 0.2 µl of AmpliTaq polymerase 5 U/µl (Invitrogen, Carlsbad, CA). Standard curves (10^3–10^8 copies/µl) were generated using purified PCR product from the *Ca.* Nitrotoga 16S rRNA primers on the CP45 enrichment culture. The qPCR was run with three technical replicates from each time point. Cycling conditions were 96°C for 2 min followed by 40 cycles of 96°C for 50 s, 58°C for 50 s, 72°C for 50 s, and 84°C for 10 s to acquire an image. R^2 values for the standard curves were > 0.995 for all runs. qPCR efficiencies ranged from 68.3 to 69.2%. Melt curve analysis was performed after each run with plate reads at a temperature increment of 0.3°C.

Physiology Experiments

All physiology treatments utilized FNOM with a final concentration of 0.3 mM nitrite distributed in sterile 20-ml borosilicate glass scintillation vials (Thermo Fisher Scientific, Waltham, MA). Prior to use, all scintillation vials were soaked in 1.9% HCl for a minimum of 24 h, followed by triplicate rinsing with DI water, triplicate rinsing with Milli-Q water, drying, and autoclaving at 15 psi and 250°F for 60 min followed immediately by a 30-min drying exhaustive cycle. The inoculum culture was regularly monitored for nitrite oxidation activity. Immediately after depletion of nitrite, *Ca.* Nitrotoga sp. CP45 was inoculated into FNOM at a ratio of 100 µl culture to 10 ml of FNOM in scintillation vials. Triplicate cultures were inoculated

for each experimental condition and grown under aerobic conditions. Each experiment (temperature, pH, light/dark, and antibiotics) was conducted at separate times, so the starting inoculum may vary slightly from experiment to experiment (e.g., physiological state and cell numbers). Therefore, data should be compared within an experiment, but not necessarily between experiments (e.g., comparisons among pH 5 and pH 7, but not pH 5 versus light).

Effects of light on Nitrotoga nitrite oxidation rates were tested with triplicate vials being exposed to continuous direct light with an illumination of 486 µEinsteins m^{-2} s^{-1}, while a separate triplicate vial set was wrapped in aluminum foil. Temperature range was tested in triplicate vials housed in incubators at the following temperatures: 4, 10, 15, 20, 23, 25, 28, and 30°C. pH treatments were performed using FNOM adjusted to pH 5.0, pH 5.7, pH 6.0, pH 7.0, and pH 8.0 using either 10% (3.26 M) hydrochloric acid (HCl, 22° Bé) or 1 M sodium hydroxide (NaOH). For antibiotic treatments, solutions of erythromycin, penicillin (penicillin G), sulfamethoxazole, and trimethoprim were made at concentrations of 5, 50, and 500 nM in a base of FNOM. All vials (with the exception of the temperature experiment) were incubated at 23°C. All vials were incubated in the dark (with the exception of the light treatment).

Nitrite concentrations were determined in 100-µl subsamples from each vial collected at approximately 6-h intervals (as described above). Nitrite oxidation rates were calculated across three time points within the period of logarithmic nitrite consumption for each vial sampled. Individual rates were calculated for each replicate and then averaged.

River Mesocosm Experiments

Three sampling sites were selected along the South Platte River Basin to evaluate nitrite oxidation capabilities of the endogenous microbial communities upon exposure to antibiotics. Each site was chosen due to contrasting landscape use and water chemistry (Storteboom et al., 2010a,b): (1) site SPCC along the South Platte River upstream of Clear Creek (−104.949010, 39.827140) representative of WWTP effluent-dominated water chemistry; (2) site SPKER along South Platte River at Colorado Highway 37 near Kersey (−104.563270, 40.412250) representative of a mixed WWTP effluent and animal feeding operation influenced water chemistry; and (3) CP45 at Cache la Poudre at County Road 45 in Greeley (−104.640170, 40.417740) dominated by animal feeding operations.

Surface sediment samples were obtained from each site and immediately placed on ice for processing at the lab within a few hours. Three sediment samples from each location were physically homogenized and 1.5 g (wet weight) was placed in a sterile scintillation vial with 15 ml of 0.3 mM nitrite FNOM plus 500 nM of an antibiotic (penicillin, erythromycin, sulfamethoxazole, or trimethoprim). Antibiotics were chosen based on their presence in river waterways as well as their antibacterial properties and pharmacokinetics. Biological controls contained river sediment and FNOM with no antibiotics. Blank controls contained FNOM and sediment samples (1.5 g) that were sterilized by autoclaving on a dry cycle at 15 psi at a

constant temperature of 250°C for 30 min, followed by cooling and vortexing for 30 s prior to a second dry autoclaving run.

Water column mesocosm experiments were set up utilizing the same three predetermined sites. For each site, 250 ml of surface water from three separate grab-bottle captures were placed into a sterile, acid-washed Nalgene and mixed for homogenization. Homogenized site water was amended with nitrite (0.3 mM final concentration) and one of the antibiotics at 500 nM final concentration (penicillin, erythromycin, sulfamethoxazole, or trimethoprim). Biological controls contained river water amended with nitrite, but with no antibiotics. Blank control water samples (45 ml) were sterilized by first filtering the sample with a 0.45-μm 25-mm PES Nalgene syringe filter, then double filtering using 0.2-μm 25-mm PES Nalgene syringe filters.

For sediment and water column mesocosm experiments, triplicates of each treatment vial were incubated at 23°C in the dark. Nitrite concentrations and nitrite oxidation potential rates were determined as described above. DNA was extracted from sediment and water column inoculum for each site in duplicate, and NOB were quantified by qPCR as described above using primers targeting *Ca.* Nitrotoga (NTG200F 5′-CTCGCGTTTTCGGAGCGG and NTG840R 5′-CTAAGGAAGTCTCCTCCC) (Alawi et al., 2007) and *Nitrospira* (Nspra675F 5′-GCGGTGAAATGCGTAGAKATCG and Nspra746R 5′-TCAGCGTCAGRWAYGTTCCAGAG) (Graham et al., 2007).

Environmental Parameters

Environmental parameters were measured at the river site where the *Ca.* Nitrotoga sp. CP45 culture inoculum was collected (see above). Briefly, surface water parameters were measured approximately every 4–8 weeks from May 2015 to July 2016. Temperature and pH were measured with a YSI Professional Plus handheld multiparameter meter (YSI Incorporated, Yellow Springs, Ohio).

Water quality data were obtained from the Northern Water Conservancy District (Northern Water Conservancy District, 2016) at a site 2 miles downstream from the CP45 sampling site (latitude 40.4244, longitude −104.6000). Data included water temperature, pH, turbidity, and dissolved inorganic nitrogen concentrations. Samples were obtained monthly from April of 2015 through November of 2016 (with the exception of January and December), analyzed at a USGS certified laboratory, and approved in a QA/QC process. At a sampling site 9 miles upstream from CP45 (latitude 40.5013, longitude −104.9673), water temperature was obtained at 15-min intervals around-the-clock from April of 2015 through May of 2016 using a HOBO ProV2 temperature logger (Onset Computer Corporation, Bourne, MA).

RESULTS AND DISCUSSION

Study Site Description

We previously enriched *Ca.* Nitrotoga sp. CP45 from the Cache la Poudre River water column in the South Platte River Basin in Colorado (Boddicker and Mosier, 2018). Genomic sequencing revealed the metabolic potential to couple the oxidation of nitrite with aerobic respiration for energy conservation *via* a novel nitrite oxidoreductase enzyme (Boddicker and Mosier, 2018). *Ca.* Nitrotoga sp. CP45 16S rRNA gene sequences were highly conserved with other *Ca.* Nitrotoga-like sequences found in freshwater habitats across the globe (Boddicker and Mosier, 2018).

As a representative freshwater *Ca.* Nitrotoga sp., here we extend this prior genomic sequencing work to evaluate the ability of *Ca.* Nitrotoga sp. CP45 to oxidize nitrite under a range of chemical and physical conditions to better understand how these NOB might respond in varying habitats across the river system. The South Platte River flows across the Colorado Front Range where multiple land-use types result in fluctuating contaminant inputs, including urban WWTP discharge, agricultural land-use areas with concentrated fertilizer and pesticide application, and land areas with a high density of animal feed operations (Storteboom et al., 2010a,b). Pristine headwaters in the Rocky Mountains near central Colorado migrate through a gradient of human activities, from urban sites in the Denver metropolitan area, to mixed setting sites dominated by urban and agricultural land use, to the most northern sites dominated by predominantly agricultural land use (Dennehy et al., 1998).

Ca. Nitrotoga sp. CP45 Growth and Nitrite Oxidation

Batch cultures of *Ca.* Nitrotoga sp. CP45 grown in media with 0.3 mM nitrite (NO_2^-) revealed that nitrite consumption corresponded with increased nitrate concentrations (at near-stoichiometric levels) and increased 16S rRNA gene copy numbers (utilized as a proxy for cell growth) (**Figure 1**). Maximum 16S rRNA gene copies at 69 h post-inoculation coincided with a logarithmic decline in nitrite concentrations, followed by a tapering off once nitrite was depleted (**Figure 1**). Accounting for genomic estimates showing that the *Ca.* Nitrotoga sp. CP45 genome had two copies of the 16S rRNA gene per cell (Boddicker and Mosier, 2018), qPCR estimates suggest that CP45 reached a maximum of 3.5×10^6 cells per ml culture. Mean nitrite oxidation rate (calculated across four measurements during logarithmic nitrite oxidation) averaged 167 ± 7 μM NO_2^- day^{-1} (**Figure 1**). *Ca.* Nitrotoga sp. CP45 oxidized NO_2^- at concentrations ranging from 0.15 to 4.5 mM NO_2^- (rates not determined).

Though some NOB are inhibited by ammonia (e.g., *Nitrospira* inhibited at 0.04–0.08 mg NH_3-N L^{-1} produced by the addition of ammonium in the media; Blackburne et al., 2007), recent studies reveal that other NOB benefit from the addition of ammonium for both growth and stable nitrite oxidation (Sorokin et al., 2012; Ishii et al., 2017, 2020; Spieck et al., 2019; Wegen et al., 2019). Specifically, some species of *Ca.* Nitrotoga have demonstrated stimulated nitrite oxidation with the addition of 0.1–30 mM ammonium (Ishii et al., 2017, 2020; Wegen et al., 2019). Here, *Ca.* Nitrotoga sp. CP45 oxidized 0.3 mM NO_2^- in the presence of 0.15–4.5 mM ammonium (rates not determined). These concentrations were higher than *in situ*
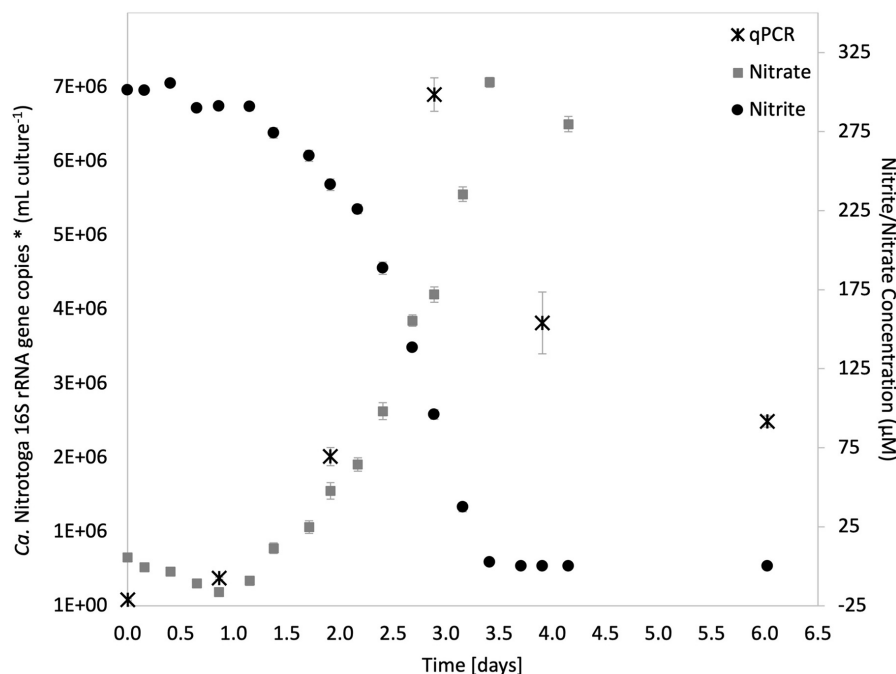
FIGURE 1 | Nitrite oxidation in the *Ca.* Nitrotoga sp. CP45 enrichment culture coupled with increased concentrations of nitrate and copies of *Ca.* Nitrotoga 16S rRNA genes. Nitrite (black circles) and nitrate (gray squares) values are averages of quadruplicate vials (each with duplicate sample measurements) at each sampling time point with error bars representing ± the standard error of the mean (SEM). Gene copy numbers measured by qPCR (black stars) are averages ± SEM error bars across three technical replicates. If not visible, error bars are smaller than data points.

river water ammonium concentrations at the CP45 sampling site (**Supplementary Figure 1**), suggesting that water column ammonium does not interfere with the potential for riverine nitrite oxidation. Prior genomic analyses indicated that CP45 encodes an ammonium transporter (*amtB*) potentially facilitating ammonium assimilation (Boddicker and Mosier, 2018). Further research should evaluate how inorganic nitrogen concentrations impact the ecology of *Ca.* Nitrotoga sp. CP45, including resource partitioning with other co-occurring NOB depending on enzyme kinetics and the range of substrate concentrations.

Impact of Light Exposure on *Ca.* Nitrotoga sp. CP45 Nitrite Oxidation

Microorganisms in aquatic environments are exposed to light at varying intensities given fluctuating depths and water transparency; however, no prior studies have evaluated the effect of light exposure on *Ca.* Nitrotoga. To test photosensitivity, *Ca.* Nitrotoga sp. CP45 was incubated under continuous illumination at 486 μEinsteins m^{-2} s^{-1} while separate biological replicates were maintained in darkness. *Ca.* Nitrotoga sp. CP45 exhibited no nitrite oxidation when grown in continuous illumination, compared to cultures incubated in the dark that oxidized nitrite at a rate of 194 μM day^{-1} (**Figure 2** and **Supplementary Table 1**). Photoeffect on *Ca.* Nitrotoga sp. CP45 could be the result of photooxidation of *c*-type cytochromes, previously reported in *Ca.* Nitrotoga genomes (Boddicker and Mosier, 2018; Kitzinger et al., 2018; Ishii et al., 2020). Cytochrome c has been shown

to absorb light in the visible spectrum causing photodynamic destruction of amino acid residues (Spikes and Livingston, 1969), resulting in light-induced cell death (Bock, 1970; Bock and Wagner, 2013). In some habitats, photochemical instability of nitrite (Zafiriou and True, 1979) could result in photoinhibition by decreasing substrate availability, though that is unlikely in the experiments conducted here since nitrite concentrations were constant over time in the CP45 incubations and in the controls that contained no biomass.

Photoinhibitory effects have been found to occur in several other nitrifying microbes (*Nitrosomonas*, *Nitrosococcus*, *Nitrospira*, *Nitrococcus*, and *Nitrobacter* species) (Hooper and Terry, 1974; Yoshioka and Saijo, 1984; Vanzella et al., 1989; Guerrero and Jones, 1996; Kaplan et al., 2000; Merbt et al., 2012). Photoinhibition documented in *Ca.* Nitrotoga sp. CP45 may also impact other *Ca.* Nitrotoga sp. with similar cell structures and metabolisms. This trait likely impacts the environmental range of these organisms, with nitrite oxidation activity occurring primarily in habitats devoid of light. *Ca.* Nitrotoga sp. CP45 was enriched from a river water column that exhibits daily and seasonal fluctuations in light exposure in part dependent on the levels of suspended particulate and organic matter. Turbidity (as a measure of water clarity) near the CP45 river site ranged from 4 to 118 NTU over time (**Supplementary Figure 2**), which spans the range of values typically seen in pristine streams and heavily sedimented rivers. High water column turbidity levels may alleviate some of the photoinhibition for nitrite oxidation by *Ca.* Nitrotoga sp. CP45 within the water column. However, turbidity
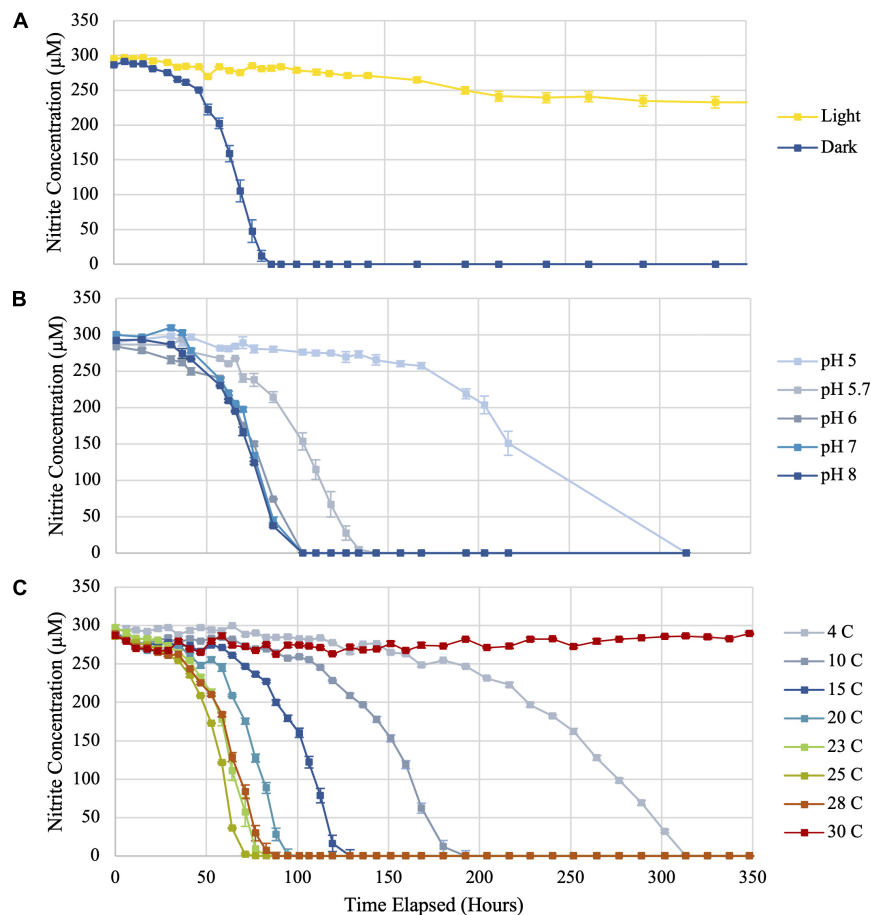
FIGURE 2 | Nitrite utilization by *Ca*. Nitrotoga sp. CP45 when **(A)** exposed to continuous light compared to incubation in darkness, **(B)** grown in media at varying pH values, and **(C)** incubated at eight different temperatures (4, 10, 15, 20, 23, 25, 28, and 30°C). Values plotted are averages of triplicate vials at each sampling time point with error bars representing ± the standard error of the mean (SEM). If not visible, error bars are smaller than data points.

could have secondary effects on NOB related to increased water temperature as the suspended particles absorb more heat or reduce dissolved oxygen concentrations from decreased photosynthesis (Austin et al., 2017). *Ca*. Nitrotoga sp. likely also benefit from photoprotection within riverbed sediments.

Impact of Varying pH on *Ca*. Nitrotoga sp. CP45 Nitrite Oxidation

Nitrite was oxidized in *Ca*. Nitrotoga sp. CP45 cultures at pH 5.0–8.0 (**Figures 2**, **3** and **Supplementary Table 1**). The highest rates of nitrite oxidation occurred at pH 7 and pH 8 (162 μM day^{-1}), with a slightly lower rate at pH 6. At the most acidic conditions tested (pH 5.0), mean nitrite oxidation rates were 42 μM day^{-1}. These results were consistent with prior *Ca*. Nitrotoga physiology studies demonstrating near-neutral pH culture conditions and optima: *Ca*. Nitrotoga arctica cultivated at pH 7.4–7.6 (optima not determined), *Ca*. Nitrotoga sp. AM1 cultivated at pH of 8.0–8.3 (optima at pH 8.3), *Ca*. Nitrotoga sp. HW29 cultivated at pH 6.8–7.4 (optima at pH 6.8), *Ca*. Nitrotoga sp. HAM-1 cultivated at pH 7.4–7.6 (optima not determined), and *Ca*. Nitrotoga sp. fabula

cultivated at pH 6.6–8.0 (optimum 7.1–7.6) (Alawi et al., 2007, 2009; Hüpeden et al., 2016; Ishii et al., 2017; Kitzinger et al., 2018). At the CP45 river enrichment site, the water pH ranged from 7.8 to 8.1 over a 10-month time period (**Supplementary Figure 2**), so these physiology tests suggest that *Ca*. Nitrotoga sp. CP45 is likely able to perform nitrite oxidation at all pH levels measured within the river year-round. It cannot be ruled out that pH changed over the course of the incubation; however, the media has some buffering capacity and NOB cultures typically become slightly more acidic over time due to proton production during nitrite oxidation (Prosser, 1989). Other community members in the enrichment culture may have impacted these results (see below); however, no other NOB were identified in the 16S rRNA gene or genome sequence dataset and no other electron donors or acceptors were added to the media.

Decreased nitrite oxidation in *Ca*. Nitrotoga sp. CP45 cultures at decreased pH values is likely the result of a combination of factors associated with abiotic chemistry and bacterial physiology. pH value has a direct impact on the equilibrium between nitrite (NO_2^-) and nitrous acid (HNO_2) and therefore directly affects the availability of nitrite in a system. According

FIGURE 3 | Temperature and pH optima for the nitrite-oxidizing activity of *Ca*. Nitrotoga sp. CP45. **(A)** Mean nitrite oxidation rates during incubation at varying temperatures. **(B)** Mean nitrite oxidation rates during incubation at varying pH conditions. Values are means of triplicate biological replicates with ± standard deviations (error bars). If not visible, error bars are smaller than data points. Values that are significantly different from the optima (highest rate; 25°C for temperature and pH 7 for pH) by *t*-test are indicated by asterisks as follows: *$p < 0.05$; **$p < 0.01$.

to the nitrous acid equilibrium, the concentration of nitrous acid increases as pH decreases, resulting in a corresponding decrease in the nitrite pool available for NOB energy generation (Philips et al., 2002). Nitrite oxidation has been shown to be inhibited at low pH due to the toxicity of nitrous acid (Prosser, 1989 and references therein). Additionally, acidic pH conditions can be toxic to NOB cells by causing a disruption in the proton concentration that is intricately involved in cellular bioenergetics through the proton motive force (Krulwich et al., 2011). Decreased pH has also been found to alter the catalytic activities of enzymes involved in nitrogen transformations by compromising their structural integrity (Schreiber et al., 2012; Blum et al., 2018). In non-neutral pH conditions, additional energy must be expended to combat pH stress, such as the expression of membrane pumps to actively uptake protons or the efflux of protons to maintain internal pH. Collectively, low nitrite pools and physiological stress at low pH can potentially lead to insufficient energy for growth (Prosser, 1989).

Nitrite oxidation at pH 5.0 suggests that *Ca*. Nitrotoga sp. CP45 may be acid tolerant, which has been shown for other NOB such as *Nitrospira* sp. (Takahashi et al., 2020) and *Nitrobacter* sp. IOacid (Hankinson and Schmidt, 1988) enriched from acid soils. We probed the *Ca*. Nitrotoga sp. CP45 genome for genes predicted to be involved in acid homeostasis. We identified a putative ActS/PrrB/RegB family redox-sensitive histidine kinase and an acid tolerance regulatory protein (ActR), which may be involved in the ActS/ActR two-component system that plays an important role in acid tolerance and oxidant resistance (Tiwari et al., 1996; Tang et al., 2017). Another potential mechanism of acid tolerance is to regulate proton and cation transport in order to reduce the overall influx of protons into the cytoplasm under low pH conditions (Booth, 1985; Baker-Austin and Dopson, 2007). The CP45 genome encoded a large number of inorganic ion transport proteins, including several cation:proton antiporters, cation-transporting ATPases

(P-type), and cation-related signal transduction proteins. The *Ca*. Nitrotoga sp. CP45 genome contained many genes for cell wall/cell membrane biosynthesis and DNA repair, which have been proposed as mechanisms of pH tolerance (Baker-Austin and Dopson, 2007 and references within), though further research is necessary to determine if these genes are used in structural modifications for pH homeostasis. Similar to the acidophilic ammonia-oxidizing archaeon, *Ca*. Nitrosotalea devanaterra (Lehtovirta-Morley et al., 2016), the *Ca*. Nitrotoga sp. CP45 genome encoded two carbonic anhydrases that may have a dual function for carbon cycling and cytoplasmic buffering to prevent acidification (Sachs et al., 2005). Some of these CP45 proteins are also found in other *Ca*. Nitrotoga genomes, which suggests that the proteins may not necessarily be involved in acid tolerance or that the other *Ca*. Nitrotoga strains may be capable of nitrite oxidation at lower pH conditions.

Since *Ca*. Nitrotoga sp. CP45 appeared to oxidize nitrite across all tested pH conditions (5.0–8.0), it would be useful to further probe survival and metabolism at pH conditions below pH 5.0 and above pH 8.0, which would be a considerable advantage in engineered systems such as WWTPs that rely on NOB function. Prior studies of nitrite oxidation in WWTP have shown that NOB (e.g., *Nitrobacter* and *Nitrospira*) operate optimally within a near-neutral range and exhibit reduced activity or cessation of nitrite oxidation at acidic conditions (Grunditz and Dalhammar, 2001; Park et al., 2007; Jiménez et al., 2011). Introduction of *Ca*. Nitrotoga sp. in dynamic WWTPs or other engineered systems may be beneficial due to their potential ability to maintain nitrite oxidation over a broader pH range.

Impact of Temperature on *Ca*. Nitrotoga sp. CP45 Nitrite Oxidation

Ca. Nitrotoga sp. CP45 oxidized nitrite when incubated at temperatures ranging from 4 to 28°C (**Figures 2**, **3** and

Supplementary Table 1), suggesting physiological temperature adaptability enabling maintenance of nitrite oxidation over the typical *in situ* temperature variability within the river year-round (**Supplementary Figure 2**). Nitrite oxidation rates were highest at 25°C (256 μM NO$_2^-$ day^{-1}). When incubated at 23 and 28°C, *Ca*. Nitrotoga sp. CP45 oxidized nitrite at similar rates to the optimum. Activity completely halted at 30°C.

CP45 nitrite oxidation at ambient temperatures is consistent with physiological characteristics of other *Ca*. Nitrotoga sp. The reported optimum temperature for described *Ca*. Nitrotoga species ranged from 10 to 28°C across a wide range of inoculum sources: 10°C for *Ca*. Nitrotoga arctica cultivated from Siberian Arctic permafrost soils (Alawi et al., 2007; Nowka et al., 2015); 16°C for *Ca*. Nitrotoga AM1 and 23°C for AM1P cultivated from marine sediments (Ishii et al., 2017, 2020); 17–22°C for *Ca*. Nitrotoga BS enriched from activated sludge (Wegen et al., 2019); 22°C for *Ca*. Nitrotoga HW29 cultivated from a recirculating aquaculture system (Hüpeden et al., 2016); and 24°C –28°C for *Ca*. Nitrotoga fabula strain KNB cultivated from activated sludge (Kitzinger et al., 2018). Our previous analyses described *Ca*. Nitrotoga-like sequences in samples with reported temperatures ranging from 0 to 33°C across many sample types (based on searches of the NCBI Sequence Read Archive with associated temperature metadata); however, *Ca*. Nitrotoga-like sequences were only identified in freshwater samples with reported temperatures < 15°C (Boddicker and Mosier, 2018). Our findings suggest that nitrite oxidation by *Ca*. Nitrotoga sp. in freshwater systems may occur over a broader range of temperatures.

Ca. Nitrotoga sp. CP45 oxidized nitrite at the lowest temperature tested (4°C) but activity was much slower than at optimum temperatures (**Figure 3**). The ability to continue nitrite oxidation at cold temperatures is useful for maintenance of ecological health in local rivers given the seasonal variation and consequential temperature fluctuation that occur across the year. At the river enrichment site, water temperature averaged 12–15°C over the year but dropped below 5°C from late November to the middle of March (**Supplementary Figure 2**). We analyzed the CP45 genome and identified several proteins that may be involved in counteracting the harmful effects of cold temperatures including cold-shock proteins such as CspA, oxidative stress response proteins such as catalase and superoxide dismutase, and proteins for the production of extracellular polysaccharides (EPS) such as alginate (Smirnova et al., 2001; Margesin and Miteva, 2011; Keto-Timonen et al., 2016). Though nitrite oxidation rates are likely quite low in the winter months, *Ca*. Nitrotoga sp. CP45 may be able to outcompete more temperature-sensitive NOB in the system. For instance, *Ca*. Nitrotoga were previously shown to outcompete *Nitrospira* and *Nitrobacter* spp. at lower temperatures (Alawi et al., 2009).

Impact of Antibiotic Exposure on *Ca*. Nitrotoga sp. CP45 Nitrite Oxidation

Antibiotics have emerged as a contaminant of concern in riverine systems influenced by human activity (Kolpin et al., 2002); however, their ecological effects have been poorly investigated despite the fact that many antibiotics possess broad-spectrum mechanisms of action that may negatively impact non-target organisms naturally present in the environment (Grenni et al., 2018). Here, we exposed *Ca*. Nitrotoga sp. CP45 cultures to four antibiotics commonly found in contaminated rivers (penicillin, erythromycin, sulfamethoxazole, and trimethoprim) at three concentrations each (5, 50, and 500 nM). Results indicated that in the presence of 5–50 nM of each antibiotic, *Ca*. Nitrotoga sp. CP45 oxidized nitrite at a slightly increased rate (though not significantly different) compared to the control culture with no antibiotics present (**Figure 4**). The 500 nM treatment had a more variable effect across the different antibiotics. *Ca*. Nitrotoga sp. CP45 had slightly increased nitrite oxidation rates with 500 nM penicillin and trimethoprim exposure (no significant difference). Rates decreased with 500 nM erythromycin ($P = 0.019$) and sulfamethoxazole ($P = 0.004$) exposure (**Figure 4**).

Overall, these findings show that *Ca*. Nitrotoga sp. CP45 is capable of maintaining nitrite oxidation in the presence of antibiotics. Increased nitrite oxidation rates in the presence of antibiotics (though not necessarily significantly different from the control) could be the result of decreased competition with other bacteria in the enrichment culture, increased resource availability from dying antibiotic-susceptible cells (e.g., release of nutrient rich macromolecules), direct nutrient supply from the antibiotic chemicals (e.g., carbon, nitrogen, and sulfur in the case of penicillin and sulfamethoxazole), or altered metabolism (e.g., heterotrophy versus autotrophy). Decreased nitrite oxidation rates at higher antibiotic concentrations (500 nM erythromycin and sulfamethoxazole) are likely indicative of physiological stress. The lowest antibiotic concentrations tested here were about twice as high as the *in situ* antibiotic concentrations previously measured in the river water column (for sulfamethoxazole and trimethoprim; Bai et al., 2018). This suggests that *Ca*. Nitrotoga sp. CP45 could likely maintain nitrite oxidation at much higher levels of antibiotic pollution in the river.

We extended these culture studies to evaluate the impact of antibiotics on nitrite oxidation by naturally occurring microbes in sediments at three sites with varying land-use patterns in the South Platte River Basin (**Figure 4**): SPCC, influenced primarily by WWTP effluent; CP45, influenced primarily by runoff from animal feeding operations (AFO); and SPKER with a mixed land-use of urban and agriculture. Each of these sources of runoff can be associated with antibiotic pollution that varies by land use (e.g., antibiotics found in human versus animal waste) (Barnes et al., 2002; Scribner et al., 2003; Koike et al., 2010; Roose-Amsaleg and Laverman, 2015). Overall, we observed rapid potential rates of nitrite oxidation in river sediments at each site. Fastest overall potential rates were observed at WWTP- and AFO-dominated sites (SPCC and CP45), which may be attributed to differences in microbial community composition, contaminant history, sediment structure, or abiotic factors (e.g., temperature). Potential rates of nitrite oxidation were similar to controls when incubated in the presence of elevated antibiotics. Quantitative PCR estimates of NOB abundance from the sediment inoculum at each site revealed that *Ca*. Nitrotoga 16S rRNA gene copies ($\sim$1.4 $\times$ 10^5 copies per gram of sediment) were more abundant than *Nitrospira* 16S rRNA gene copies at each site ($\sim$7.8 $\times$ 10^4
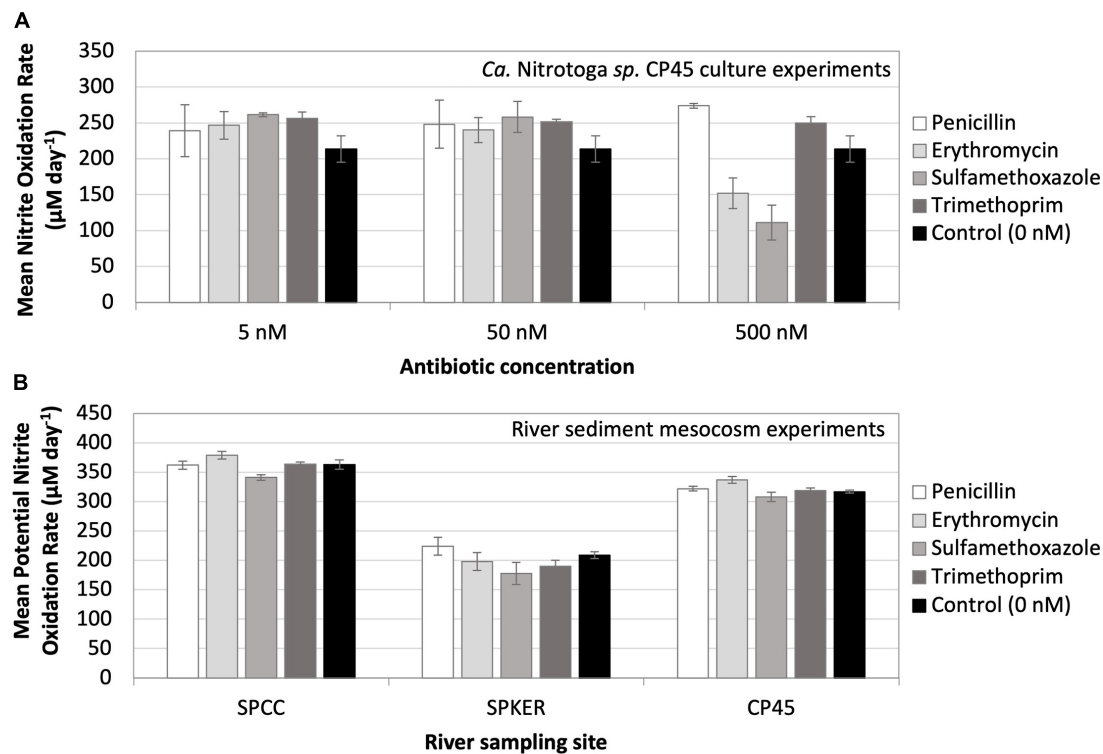
FIGURE 4 | (A) Mean nitrite oxidation rates (μM day^{-1}) by *Ca*. Nitrotoga sp. CP45 in the presence of 5, 50, and 500 nM concentrations of four different antibiotics (penicillin, erythromycin, sulfamethoxazole, and trimethoprim). Controls contained no antibiotics. **(B)** Mean potential nitrite oxidation rates (μM day^{-1}) from river mesocosm experiments where sediments from three river sites (SPCC, SPKER, and CP45) were exposed to four antibiotics (penicillin, erythromycin, sulfamethoxazole, and trimethoprim) at a concentration of 500 nM. Antibiotic-free sediments were utilized to represent inherent, endogenous microbial nitrite oxidation activity at each site. Each symbol represents the average of triplicate cultures or sediment incubations grown under each condition with standard deviation error bars.

copies per gram of sediment) (*Nitrobacter* genes were not amplifiable at any site). Mesocosm experiments using the river water column as inoculum (without added sediment) showed essentially no NOB amplification by qPCR and no nitrite oxidation with or without antibiotic amendments. The complex nature of the sediment microbial community and metabolite profile may influence the nitrogen and carbon pools in these river mesocosm experiments, so rates should be interpreted as the "potential" for nitrite oxidation under the given conditions (with the possibility of other biotic and abiotic influences). Further research is needed to determine the activity of specific community members upon antibiotic exposure.

Limited studies have evaluated the impact of antibiotics on nitrogen cycling. Studies focusing on the effect of antibiotics on nitrification in activated sludge revealed that benzylpenicillin exposure resulted in no effect on the rate of nitrate production, while erythromycin demonstrated an inhibition in ammonification, nitritation, and nitratation (Gomez et al., 1996; Alighardashi et al., 2009; Katipoglu-Yazan et al., 2013). Additionally, nitrite oxidation in a long-term laboratory-scale treatment plant was inhibited when exposed to a mixture of ciprofloxacin, gentamicin, sulfamethoxazole/trimethoprim, and vancomycin resulting in no nitrate formation (Schmidt et al., 2012). Variability in the overall effect of antibiotics in

different systems is likely a result of differences in tolerance of the underlying microbial community. Also, the complexity of each individual antibiotic class (e.g., absorption, pharmacokinetics, metabolite formation, degradation, and sediment adsorption) affects the bioavailability of the antibiotic once it enters the waterway (Luo et al., 2011).

Here, the uniform rates of potential nitrite oxidation by endogenous sediment communities across the three sampling sites and in culture would suggest mechanisms for tolerating the effects of antibiotics. NOB are Gram-negative bacteria that characteristically possess an additional outer membrane that can serve as protection against large antibacterial compounds that cannot penetrate the cell wall and enter the cytoplasm. Furthermore, many NOB can produce EPS (including *Ca*. Nitrotoga sp. CP45 based on genomic predictions; see above) that facilitate biofilm formation (Hüpeden et al., 2016), which may protect interior NOB from antibiotics at the surface of the biofilm. Acquired antibiotic resistance can occur by adaptive gene mutation in the presence of a stressor (such as antibiotics) or through the process of genetic exchange mechanisms (e.g., transformation, transduction, or conjugation) (Rosenblatt-Farrell, 2009). Previous research identified antibiotic pollution (Yang and Carlson, 2003; Pei et al., 2006; Pruden et al., 2006; Kim and Carlson, 2007; Bai et al., 2018) and

antibiotic resistance genes (Storteboom et al., 2010a,b) in the South Platte River Basin. The NOB studied here likely also have developed some mechanism of antibiotic resistance in order to maintain nitrite oxidation during antibiotic exposure. The *Ca.* Nitrotoga sp. CP45 genome was predicted to encode a broad-spectrum antibiotic efflux pump that may confer resistance to all four antibiotics tested here (Boddicker and Mosier, 2018; based on searches against the CARD database, Alcock et al., 2020). Expanded annotation analyses here indicated that the *Ca.* Nitrotoga sp. CP45 genome was predicted to also encode a MdtABC multidrug exporter system (Nagakubo et al., 2002), as well as several additional putative multidrug efflux proteins identified by manual annotation. Riverine NOB capable of nitrite oxidation in the mesocosm experiments likely exhibited similar resistance mechanisms. These antibiotic-resistant NOB may spread their resistance genes to other microbes in the river *via* mechanisms of genetic exchange. Future efforts should evaluate temporal trends in antibiotic exposure, gene expression of NOB under antibiotic stress, and the effects of contaminant mixtures common within river systems.

Influence of Community Members on Nitrite Oxidation by *Ca.* Nitrotoga sp. CP45

Ca. Nitrotoga sp. CP45 was grown in enrichment cultures, so other community members also likely influence their physiology. At the time of genomic sequencing, other organisms (besides *Ca.* Nitrotoga sp. CP45) in the CP45 enrichment culture were identified as p__Bacteroidetes, o__Burkholderiales, o__Rhizobiales, and o__Pseudomonadales by checkM (Parks et al., 2015); *Ca.* Nitrotoga sp. CP45 was enriched over time to > 65% abundance (Boddicker and Mosier, 2018). No other organisms capable of nitrite oxidation or ammonia oxidation were present based on evaluation of 16S rRNA gene sequences and protein coding genes sequenced in the culture metagenome. *Ca.* Nitrotoga sp. CP45 may utilize metabolites and/or cofactors produced by heterotrophs within the culture to compensate for incomplete biosynthetic pathways within their own genome and/or to reduce biosynthetic costs (e.g., Daims et al., 2016; Ngugi et al., 2016). Likewise, *Ca.* Nitrotoga sp. CP45 metabolism may influence other community members (e.g., production of organic carbon, pH impacts). As biomass builds up toward the end of nitrite oxidation, organic carbon likely results in a dynamic exchange of metabolites and cross-feeding among the community members, as well as other chemical changes (e.g., oxygen and pH). When biomass is available, some community members may be capable of aerobic denitrification using organic carbon as an electron donor (inorganic electron donors such as hydrogen or sulfur can be used by some organisms under certain circumstances) and nitrite/nitrate as an electron acceptor for conversion to nitrogen gases. This is unlikely to occur in the early phases of *Ca.* Nitrotoga sp. CP45 growth since the culture conditions contained nitrite as the only known electron donor and oxygen as the only electron acceptor. Adjustments to culture conditions (e.g., low or high pH) may alter the activity of the other non-NOB community members depending on their own

physiology (range and optima for each parameter tested). Abiotic sources and sinks may also influence the community dynamics. Nonetheless, our data suggest that *Ca.* Nitrotoga sp. CP45 oxidizes nitrite to nitrate (with corresponding increases in cell counts) during the exponential growth phase, so the nitrite data presented likely reflect metabolic processes specific to these cells. In some ways, these complex community relationships within the culture hint at the reality of *Ca.* Nitrotoga sp. CP45 living in a dynamic ecosystem with fluctuating biological, chemical, and physical factors.

CONCLUSION

Ca. Nitrotoga has emerged as a genus of versatile organisms with a diverse metabolic potential. Physiological tests showed that nitrite oxidation can occur in darkness, across all pH values tested (pH 5.0–8.0), at temperatures ranging from 4 to 28°C, and in the presence of antibiotics. The ability of *Ca.* Nitrotoga sp. CP45 to oxidize nitrite across a variety of conditions indicates that the environmental range of *Ca.* Nitrotoga may be greater than initially estimated. The physiological versatility of *Ca.* Nitrotoga may be ecologically valuable in natural and engineered environments that rely on the function of NOB to mitigate nitrite toxicity. By recognizing potential constraining variables that may limit nitrite oxidation, future management decisions may be guided to best manage elevated levels of nitrogen in contaminated systems that rely on the activity of NOB communities for mitigation. Understanding the influence of key microorganisms, such as NOB, strengthens the predictive power required to recognize and manage nitrogen levels that, if allowed to intensify, may otherwise result in a cascade of environmental and human health complications.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

ML and AM conceived and designed the experiments. ML, CW, and AM contributed to sampling, site information, and chemistry. ML, AB, MK, OB, and AM performed the experiments. ML, AB, and AM analyzed the data. ML, AB, MK, OB, CW, and AM contributed to the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb. 2021.709371/full#supplementary-material

Supplementary Figure 1 | The concentration of **(A)** ammonia as nitrogen (μM NH_3-N) and **(B)** nitrate and nitrite as nitrogen (μM NO_3^- + NO_2^--N) in the experimental dataset obtained by the Northern Water Conservation District at the nearby site approximately 2 miles away on the same river.

Supplementary Figure 2 | Water column data measured at the CP45 site (gray symbols) or obtained by the Northern Water Conservation District at the nearby site approximately 2 miles away (black symbols) for **(A)** pH, **(B)** turbidity, and **(C)** temperature. **(D)** Water temperature at 15-min intervals around the clock for the 12-month time period from a Northern Water Conservation District sampling site approximately 9 miles away from the experimental sampling site.

Supplementary Table 1 | Mean nitrite oxidation rates (μM day^{-1}) at varying treatment conditions, with accompanying standard deviation values of replicate cultures. Data should be compared within an experiment, but not necessarily between experiments (see section "Materials and Methods").

REFERENCES

Alawi, M., Lipski, A., Sanders, T., Pfeiffer, E. M., and Spieck, E. (2007). Cultivation of a novel cold-adapted nitrite oxidizing betaproteobacterium from the Siberian Arctic. *ISME J.* 1, 256–264. doi: 10.1038/ismej.2007.34

Alawi, M., Off, S., Kaya, M., and Spieck, E. (2009). Temperature influences the population structure of nitrite-oxidizing bacteria in activated sludge. *Environ. Microbiol. Rep.* 1, 184–190. doi: 10.1111/j.1758-2229.2009.00029.x

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525.

Alighardashi, A., Pandolfi, D., Potier, O., and Pons, M. N. (2009). Acute sensitivity of activated sludge bacteria to erythromycin. *J. Hazard. Mater.* 172, 685–692. doi: 10.1016/j.jhazmat.2009.07.051

Altmann, D., Stief, P., Amann, R., Beer, D. D., and Schramm, A. (2003). *In situ* distribution and activity of nitrifying bacteria in freshwater sediment. *Environ. Microbiol.* 5, 798–803. doi: 10.1046/j.1469-2920.2003.00469.x

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.

American Public Health Association, American Water Works Association, and Water Environment Federation (2017). *Standard Methods for the Examination of Water and Wastewater, 23rd Edition 4500-NO_2^- B. Colorimetric Method.* Washington, D.C: American Public Health Association, American Water Works Association, Water Environment Federation.

Austin, ÅN., Hansen, J. P., Donadi, S., and Eklöf, J. S. (2017). Relationships between aquatic vegetation and water turbidity: a field survey across seasons and spatial scales. *PLoS One* 12:e0181419. doi: 10.1371/journal.pone.0181419

Bai, X., Lutz, A., Carroll, R., Keteles, K., Dahlin, K., Murphy, M., et al. (2018). Occurrence, distribution, and seasonality of emerging contaminants in urban watersheds. *Chemosphere* 200, 133–142. doi: 10.1016/j.chemosphere.2018.02. 106

Baker-Austin, C., and Dopson, M. (2007). Life in acid: pH homeostasis in acidophiles. *Trends Microbiol.* 15, 165–171. doi: 10.1016/j.tim.2007.02.005

Barnes, K. K., Kolpin, D. W., Meyer, M. T., Thurman, E. M., Furlong, E. T., Zaugg, S. D., et al. (2002). *Water-Quality Data for Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999-2000 (2002-94).* Reston, VA: U.S. Geological Survey.

Blackburne, R., Vadivelu, V. M., Yuan, Z., and Keller, J. (2007). Kinetic characterisation of an enriched *Nitrospira* culture with comparison to *Nitrobacter*. *Water Res.* 41, 3033–3042. doi: 10.1016/j.watres.2007. 01.043

Blum, J.-M., Su, Q., Ma, Y., Valverde-Pérez, B., Domingo-Félez, C., Jensen, M. M., et al. (2018). The pH dependency of N-converting enzymatic processes, pathways and microbes: effect on net N2O production. *Environ. Microbiol.* 20, 1623–1640. doi: 10.1111/1462-2920.14063

Blum, M., Chang, H., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., et al. (2021). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* 29, D344–D354. doi: 10.1093/nar/gkaa977

Bock, E. (1970). Untersuchungen über die wechselwirkung zwischen licht und chemosynthese am beispiel von nitrobacter winogradskyi. *Arch. Mikrobiol.* 70, 217–239. doi: 10.1007/bf00407712

Bock, E., and Wagner, M. (2013). "Oxidation of inorganic nitrogen compounds as an energy source," in *The Prokaryotes*, eds M. Dworkin, S. Falkow, E. Rosenberg, K. H. Schleifer, and E. Stackebrandt (New York, NY: Springer), 83–118. doi: 10.1007/978-3-642-30141-4_64

Boddicker, A. M., and Mosier, A. C. (2018). Genomic profiling of four cultivated *Candidatus* Nitrotoga spp. predicts broad metabolic potential and environmental distribution. *ISME J.* 12, 2864–2882. doi: 10.1038/s41396-018- 0240-8

Booth, I. (1985). Regulation of cytoplasmic pH in Bacteria. *Microbiol. Rev.* 49, 359–378. doi: 10.1128/mmbr.49.4.359-378.1985

Cébron, A., and Garnier, J. (2005). *Nitrobacter* and *Nitrospira* genera as representatives of nitrite-oxidizing bacteria: detection, quantification and growth along the lower Seine River (France). *Water Res.* 39, 4979–4992. doi: 10.1016/j.watres.2005.10.006

Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., et al. (2009). Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J.* 3, 1093–1104. doi: 10.1038/ismej.2009.57

Christner, B. C., Priscu, J. C., Achberger, A. M., Barbante, C., Carter, S. P., Christianson, K., et al. (2014). A microbial ecosystem beneath the West Antarctic ice sheet. *Nature* 512, 310–313. doi: 10.1038/nature13667

Daims, H., Lucker, S., and Wagner, M. (2016). A new perspective on microbes formerly known as nitrite-oxidizing bacteria. *Trends Microbiol.* 24, 699–712. doi: 10.1016/j.tim.2016.05.004

Dennehy, K. F., Litke, D. W., Tate, C. M., Qi, S. L., McMahon, P. B., Bruce, B. W., et al. (1998). *Water Quality in the South Platte River Basin, Colorado, Nebraska, and Wyoming. 1992–1995.*

Fan, L., Song, C., Meng, S., Qiu, L., Zheng, Y., Wu, W., et al. (2016). Spatial distribution of planktonic bacterial and archaeal communities in the upper section of the tidal reach in Yangtze River. *Sci. Rep.* 6:39147. doi: 10.1038/srep39147

Féray, C., Volat, B., Degrange, V., Clays-Josserand, A., and Montuelle, B. (1999). Assessment of three methods for detection and quantification of nitrite-oxidizing bacteria and *Nitrobacter* in freshwater sediments (MPN-PCR, MPN-Griess, Immunofluorescence). *Microb. Ecol.* 37, 208–217. doi: 10.1007/s002489900144

Gomez, J., Mendez, R., and Lema, J. M. (1996). The effect of antibiotics on nitrification processes. *Appl. Biochem. Biotechnol.* 57, 869–876. doi: 10.1007/bf02941767

Graham, D. W., Knapp, C. W., Van Vleck, E. S., Bloor, K., Lane, T. B., and Graham, C. E. (2007). Experimental demonstration of chaotic instability in biological nitrification. *The ISME Journal* 1, 385–393. doi: 10.1038/ismej.2007.45

Grenni, P., Ancona, V., and Barra Caracciolo, A. (2018). Ecological effects of antibiotics on natural ecosystems: A review. *Microchem. J.* 136, 25–39. doi: 10.1016/j.microc.2017.02.006

Griess-Romijn van Eck, E. (1966). *Physiological and Chemical Tests for Drinking Water.* Rijswijk: Nederlands Normalisatie Instituut 504, 1056.

Grunditz, C., and Dalhammar, G. (2001). Development of nitrification inhibition assays using pure cultures of *Nitrosomonas* and *Nitrobacter*. *Water Res.* 35, 433–440. doi: 10.1016/s0043-1354(00)00312-2

Guerrero, M., and Jones, R. (1996). Photoinhibition of marine nitrifying bacteria. I. Wavelength-dependent response. *Mar. Ecol. Progr. Ser.* 141, 183–192. doi: 10.3354/meps141183

Hankinson, T. R., and Schmidt, E. L. (1988). An acidophilic and a neutrophilic *Nitrobacter* strain isolated from the numerically predominant nitrite-oxidizing population of an acid forest soil. *Appl. Environ. Microbiol.* 54, 1536–1540. doi: 10.1128/aem.54.6.1536-1540.1988

Hooper, A. B., and Terry, K. R. (1974). Photoinactivation of ammonia oxidation in *Nitrosomonas*. *J. Bacteriol.* 119, 899–906. doi: 10.1128/jb.119.3.899-906.1974

Hovanec, T. A., Taylor, L. T., Blakis, A., and Delong, E. F. (1998). *Nitrospira*-like bacteria associated with nitrite oxidation in freshwater aquaria. *Appl. Environ. Microbiol.* 64, 258–264. doi: 10.1128/aem.64.1.258-264.1998

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., et al. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.* 34, 2115–2122. doi: 10.1093/molbev/msx148

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernandez-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/nar/gky1085

Hüpeden, J., Wegen, S., Off, S., Lücker, S., Bedarf, Y., Daims, H., et al. (2016). Relative abundance of *Nitrotoga* spp. in a biofilter of a cold-freshwater aquaculture plant appears to be stimulated by slightly acidic pH. *Appl. Environ. Microbiol.* 82, 1838–1845. doi: 10.1128/aem.03163-15

Ishii, K., Fujitani, H., Sekiguchi, Y., and Tsuneda, S. (2020). Physiological and genomic characterization of a new '*Candidatus* Nitrotoga' isolate. *Environ. Microbiol.* 22, 2365–2382. doi: 10.1111/1462-2920.15015

Ishii, K., Fujitani, H., Soh, K., Nakagawa, T., Takahashi, R., and Tsuneda, S. (2017). Enrichment and physiological characterization of a cold-adapted nitrite-oxidizing *Nitrotoga* sp. from an eelgrass sediment. *Appl. Environ. Microbiol.* 83:e00549-17. doi: 10.1128/aem.00549-17

Jiménez, E., Giménez, J., Ruano, M., Ferrer, J., and Serralta, J. (2011). Effect of pH and nitrite concentration on nitrite oxidation rate. *Bioresour. Technol.* 102, 8741–8747. doi: 10.1016/j.biortech.2011.07.092

Kaplan, D., Wilhelm, R., and Abeliovich, A. (2000). Interdependent environmental factors controlling nitrification in waters. *Water Sci. Technol.* 42, 167–172. doi: 10.2166/wst.2000.0309

Katipoglu-Yazan, T., Pala-Ozkok, I., Ubay-Cokgor, E., and Orhon, D. (2013). Acute impact of erythromycin and tetracycline on the kinetics of nitrification and organic carbon removal in mixed microbial culture. *Bioresour. Technol.* 144, 410–419. doi: 10.1016/j.biortech.2013.06.121

Keto-Timonen, R., Hietala, N., Palonen, E., Hakakorpi, A., Lindstrom, M., and Korkeala, H. (2016). Cold shock proteins: a minireview with special emphasis on Csp-family of enteropathogenic *Yersinia*. *Front. Microbiol.* 7:1151. doi: 10.3389/fmicb.2016.01151

Kim, S. C., and Carlson, K. (2007). Temporal and spatial trends in the occurrence of human and veterinary antibiotics in aqueous and river sediment matrices. *Environ. Sci. Technol.* 41, 50–57. doi: 10.1021/es060737

Kitzinger, K., Koch, H., Lücker, S., Sedlacek, C. J., Herbold, C., Schwarz, J., et al. (2018). Characterization of the first "*Candidatus* Nitrotoga" isolate reveals metabolic versatility and separate evolution of widespread nitrite-oxidizing bacteria. *mBio* 9:e01186-18. doi: 10.1128/mBio.01186-18

Koike, S., Aminov, R. I., Yannarell, A. C., Gans, H. D., Krapac, I. G., Chee-Sanford, J. C., et al. (2010). Molecular ecology of macrolide–lincosamide–streptogramin B methylases in waste lagoons and subsurface waters associated with swine production. *Microb. Ecol.* 59, 487–498. doi: 10.1007/s00248-009-9610-0

Kolpin, D. W., Furlong, E. T., Meyer, M. T., Thurman, E. M., Zaugg, S. D., Barber, L. B., et al. (2002). Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999-2000: a national reconnaissance. *Environ. Sci. Technol.* 36, 1202–1211. doi: 10.1021/es011055j

Krulwich, T. A., Sachs, G., and Padan, E. (2011). Molecular aspects of bacterial pH sensing and homeostasis. *Nat. Rev. Microbiol.* 9, 330–343. doi: 10.1038/nrmicro2549

Lantz, M. (2019). *Impacts of Environmental Change on Freshwater Nitrite Oxidation.* University of Colorado Denver.

Lehtovirta-Morley, L. E., Sayavedra-Soto, L. A., Gallois, N., Schouten, S., Stein, L. Y., Prosser, J. I., et al. (2016). Identifying potential mechanisms enabling acidophily in the ammonia-oxidizing archaeon "*Candidatus* Nitrosotalea devanaterra". *Appl. Environ. Microbiol.* 82, 2608–2619. doi: 10.1128/AEM.04031-15

Li, D., Qi, R., Yang, M., Zhang, Y., and Yu, T. (2011). Bacterial community characteristics under long-term antibiotic selection pressures. *Water Res.* 45, 6063–6073. doi: 10.1016/j.watres.2011.09.002

Liu, B., Zheng, D. D., Jin, Q., Chen, L. H., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* 47, D687–D692.

Lücker, S., Schwarz, J., Gruber-Dorninger, C., Spieck, E., Wagner, M., and Daims, H. (2014). *Nitrotoga*-like bacteria are previously unrecognized key nitrite oxidizers in full-scale wastewater treatment plants. *ISME J.* 9, 708–720. doi: 10.1038/ismej.2014.158

Luo, Y., Xu, L., Rysz, M., Wang, Y., Zhang, H., and Alvarez, P. J. J. (2011). Occurrence and transport of tetracycline, sulfonamide, quinolone, and macrolide antibiotics in the haihe river basin, China. *Environ. Sci. Technol.* 45, 1827–1833. doi: 10.1021/es104009s

Margesin, R., and Miteva, V. (2011). Diversity and ecology of psychrophilic microorganisms. *Res. Microbiol.* 162, 346–361. doi: 10.1016/j.resmic.2010.12.004

Merbt, S. N., Stahl, D. A., Casamayor, E. O., Martí, E., Nicol, G. W., and Prosser, J. I. (2012). Differential photoinhibition of bacterial and archaeal ammonia oxidation. *FEMS Microbiol. Lett.* 327, 41–46. doi: 10.1111/j.1574-6968.2011.02457.x

Miranda, K. M., Espey, M. G., and Wink, D. A. (2001). A rapid, simple spectrophotometric method for simultaneous detection of nitrate and nitrite. *Nitric Oxide Biol. Chem.* 5, 62–71. doi: 10.1006/niox.2000.0319

Nagakubo, S., Nishino, K., Hirata, T., and Yamaguchi, A. (2002). The putative response regulator BaeR stimulates multidrug resistance of *Escherichia coli* via a novel multidrug exporter system, MdtABC. *J. Bacteriol.* 184, 4161–4167. doi: 10.1128/jb.184.15.4161-4167.2002

Ngugi, D. K., Blom, J., Stepanauskas, R., and Stingl, U. (2016). Diversification and niche adaptations of *Nitrospina*-like bacteria in the polyextreme interfaces of Red Sea brines. *ISME J.* 10, 1383–1399. doi: 10.1038/ismej.2015.214

Northern Water Conservancy District (2016). East Slope Temperature Data. Available online at: https://www.northernwater.org/WaterQuality/EastSlopeWaterTemperatureData.aspx; https://www.northernwater.org/DynData/WQDataMain.aspx

Nowka, B., Daims, H., and Spieck, E. (2015). Comparison of oxidation kinetics of nitrite-oxidizing bacteria: nitrite availability as a key factor in niche differentiation. *Appl. Environ. Microbiol.* 81, 745–753. doi: 10.1128/AEM.02734-14

Pal, C., Bengtsson-Palme, J., Rensing, C., Kristiansson, E., and Larsson, D. G. (2014). BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res.* 42, D737–D743. doi: 10.1093/nar/gkt1252

Park, S., Bae, W., Chung, J., and Baek, S.-C. (2007). Empirical model of the pH dependence of the maximum specific nitrification rate. *Process Biochem.* 42, 1671–1676. doi: 10.1016/j.procbio.2007.09.010

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Pei, R., Kim, S. C., Carlson, K. H., and Pruden, A. (2006). Effect of river landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Res.* 40, 2427–2435. doi: 10.1016/j.watres.2006.04.017

Philips, S., Laanbroek, H. J., and Verstraete, W. (2002). Origin, causes and effects of increased nitrite concentrations in aquatic environments. *Rev. Environ. Sci. Biotechnol.* 1, 115–141. doi: 10.1023/a:1020892826575

Prosser, J. (1989). Autotrophic nitrification in bacteria. *Adv. Microb. Physiol.* 30, 125–181. doi: 10.1016/s0065-2911(08)60112-5

Pruden, A., Pei, R., Storteboom, H., and Carlson, K. H. (2006). Antibiotic resistance genes as emerging contaminants: studies in Northern Colorado. *Environ. Sci. Technol.* 40, 7445–7450. doi: 10.1021/es060413l

Roose-Amsaleg, C., and Laverman, A. M. (2015). Do antibiotics have environmental side-effects? Impact of snythetic antibiotics on biogeochemical processes. *Environ. Sci. Pollut. Res.* 23, 4000–4012. doi: 10.1007/s11356-015-4943-3

Rosenblatt-Farrell, N. (2009). The landscape of antibiotic resistance. *Environ. Health Perspect.s* 117, A244–A250.

Sachs, G., Weeks, D., Wen, Y., Marcus, E., and Scott, D. (2005). Acid accumulation by *Helicobacter pylori*. *Physiology* 20, 429–438. doi: 10.1152/physiol.00032.2005

Schmidt, S., Winter, J., and Gallert, C. (2012). Long-term effects of antibiotics on the elimination of chemical oxygen demand, nitrification, and viable bacteria in laboratory-scale wastewater treatment plants. *Arch. Environ. Contam. Toxicol.* 63, 354–364. doi: 10.1007/s00244-012-9773-4

Schreiber, F., Wunderlin, P., Udert, K. M., and Wells, G. F. (2012). Nitric oxide and nitrous oxide turnover in natural and engineered microbial communities: biological pathways, chemical reactions, and novel technologies. *Front. Microbiol.* 3:372. doi: 10.3389/fmicb.2012.00372

Scribner, E. A., Battaglin, W. A., Dietze, J. E., and Thurman, E. M. (2003). *Reconnaissance Data for Glyphosate, Other Selected Herbicides, their Degradation Products, and Antibiotics in 51 Streams in Nine Midwestern States, 2002 (2003-217).* Reston, VA: U.S. Geological Survey.

Smirnova, G., Zakirova, O., and Oktyabrskii, O. (2001). The role of antioxidant systems in the cold stress response of *Escherichia coli*. *Microbiology* 70, 45–50.

Sorokin, D. Y., Lücker, S., Vejmelkova, D., Kostrikina, N. A., Kleerebezem, R., Rijpstra, W. I. C., et al. (2012). Nitrification expanded: discovery, physiology and genomics of a nitrite-oxidizing bacterium from the phylum Chloroflexi. *ISME J.* 6, 2245–2256. doi: 10.1038/ismej.2012.70

Spieck, E., Spohn, M., Wendt, K., Bock, E., Shively, J., Frank, J., et al. (2019). Extremophilic nitriteoxidizing Chloroflexi from Yellowstone hot springs. *ISME J.* 14. doi: 10.1038/s41396-019-0530-9 [Epub ahead of print].

Spikes, J. D., and Livingston, R. (1969). The molecular biology of photodynamic action: sensitized photoautoxidations in biological systems. *Adv. Radiat. Biol.* 3, 29–121. doi: 10.1016/b978-1-4832-3122-8.50008-1

Storteboom, H., Arabi, M., Davis, J. G., Crimi, B., and Pruden, A. (2010a). Identification of antibiotic-resistance-gene molecular signatures suitable as

tracers of pristine river, urban, and agricultural sources. *Environ. Sci. Technol.* 44, 1947–1953. doi: 10.1021/es902893f

Storteboom, H., Arabi, M., Davis, J. G., Crimi, B., and Pruden, A. (2010b). Tracking antibiotic resistance genes in the South Platte River Basin using molecular signatures of urban, agricultural, and pristine sources. *Environ. Sci. Technol.* 44, 7397–7404. doi: 10.1021/es101657s

Takahashi, Y., Fujitani, H., Hirono, Y., Tago, K., Wang, Y., Hayatsu, M., et al. (2020). Enrichment of comammox and nitrite-oxidizing *Nitrospira* from acidic soils. *Front. Microbiol.* 11:1737. doi: 10.3389/fmicb.2020.01737

Tang, G., Wang, S., Lu, D., Huang, L., Li, N., and Luo, L. (2017). Two-component regulatory system ActS/ActR is required for Sinorhizobium meliloti adaptation to oxidative stress. *Microbiol. Res.* 198, 1–7. doi: 10.1016/j.micres.2017.01.005

Tiwari, R., Reeve, W., Dilworth, M., and Glenn, A. (1996). Acid tolerance in *Rhizobium meliloti* strain WSM419 involves a two-component sensor-regulator system. *Microbiology* 142, 1693–1704. doi: 10.1099/13500872-142-7-1693

Vanzella, A., Guerrero, M., and Jones, R. (1989). Effect of CO and light on ammonium and nitrite oxidation by chemolithotrophic bacteria. *Mar. Ecol. Progr. Ser.* 57, 69–76. doi: 10.3354/meps057069

Wegen, S., Nowka, B., and Spieck, E. (2019). Low temperature and neutral pH define "*Candidatus* Nitrotoga sp." as a competitive nitrite oxidizer in coculture with *Nitrospira defluvii*. *Appl. Environ. Microbiol.* 85:e02569-18. doi: 10.1128/aem.02569-18

White, C. P., Debry, R. W., and Lytle, D. A. (2012). Microbial survey of a full-scale, biologically active filter for treatment of drinking water. *Appl. Environ. Microbiol.* 78, 6390–6394. doi: 10.1128/aem.00308-12

Yang, S., and Carlson, K. (2003). Evolution of antibiotic occurrence in a river through pristine, urban and agricultural landscapes. *Water Res.* 37, 4645–4656. doi: 10.1016/S0043-1354(03)00399-3

Yoshioka, T., and Saijo, Y. (1984). Photoinhibition and recovery of NH_4-oxidizing bacteria and NO_2-oxidizing bacteria. *J. Gen. Appl. Microbiol.* 30, 151–166. doi: 10.2323/jgam.30.151

Zafiriou, O. C., and True, M. B. (1979). Nitrite photolysis in seawater by sunlight. *Mar. Chem.* 8, 9–32. doi: 10.1016/0304-4203(79)90029-x

frontiers
in Microbiology

Check for updates

Corrigendum: Physiology of the Nitrite-Oxidizing Bacterium *Candidatus* Nitrotoga sp. CP45 Enriched From a Colorado River

Munira A. Lantz, Andrew M. Boddicker, Michael P. Kain, Owen M. C. Berg, Courtney D. Wham and Annika C. Mosier*

Department of Integrative Biology, University of Colorado Denver, Denver, CO, United States

A Corrigendum on

Physiology of the Nitrite-Oxidizing Bacterium *Candidatus* Nitrotoga sp. CP45 Enriched From a Colorado River

by Lantz, M. A., Boddicker, A. M., Kain, M. P., Berg, M. C. O., Wham, C. D., and Mosier, A. C. (2021). Front. Microbiology. 12:709371. doi: 10.3389/fmicb.2021.709371

In the original article, there was an error in the Acknowledgment section where Gracie RedShirt Tyon's important contributions were unintentionally missed.

A correction has been made to the **Acknowledgment**:

We thank Hannah Clark, Bhargavi Ramanathan, Nicklaus Deevers, Colin Beacom, Ashley Luntsford, Ashley Gonzales, Renee Kershaw, Anna Scopp, Sladjana Subotic, Brian Legvold, and Ben Wise for their assistance in field sampling, cultivation, and/or preliminary physiology experiments. We thank Alan Vajda, Kristen Keteles, and Karl Hermann for discussions about river contaminant chemistry. Portions of this manuscript were previously published as a part of University of Colorado Denver Ph.D. Dissertation submission (Lantz, 2019). "We honor and acknowledge that the samples used for the research presented here were collected from rivers that are from the traditional territories and ancestral homelands of the Cheyenne, Arapaho, and Ute Nations, as well as over 45 Indigenous Nations who called this area home. The confluence of the Platte and Cherry Creek Rivers was the epicenter for trade, information sharing, planning for the future, community, family and ally building, and conducting healing ceremonies. We recognize Indigenous peoples as the original inhabitants, stewards, and relatives of this land. Let us acknowledge the painful history of genocide and forced removal from this territory and pay our respect to the diverse Indigenous peoples still connected to this land. We thank all Tribal Nations and the ancestors of this place." We thank Gracie RedShirt Tyon (Lakota, Director of American Indian Student Services, University of Colorado at Denver) for assistance in understanding the history of this region and for writing this land acknowledgment statement.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

REFERENCES

Lantz, M. (2019). *Impacts of Environmental Change on Freshwater Nitrite Oxidation.* University of Colorado Denver.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Check for
updates

Expanding Characterized Diversity and the Pool of Complete Genome Sequences of *Methylococcus* Species, the Bacteria of High Environmental and Biotechnological Relevance

OPEN ACCESS

Edited by:
Ludmila Chistoserdova,
University of Washington,
United States

Reviewed by:
Lisa Y. Stein,
University of Alberta, Canada
Andrew Crombie,
University of East Anglia,
United Kingdom

*Correspondence:
Svetlana N. Dedysh
dedysh@mail.ru

Igor Y. Oshkin[1], Olga V. Danilova[1], Sergey Y. But[1,2], Kirill K. Miroshnikov[1],
Ruslan Z. Suleimanov[1], Svetlana E. Belova[1], Ekaterina N. Tikhonova[1],
Nikolai N. Kuznetsov[1], Valentina N. Khmelenina[2], Nikolai V. Pimenov[1] and
Svetlana N. Dedysh[1]*

[1]*Winogradsky Institute of Microbiology, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, Russia,*
[2]*G. K. Skryabin Institute of Biochemistry and Physiology of Microorganisms, Pushchino Scientific Center for Biological*
Research, Russian Academy of Sciences, Pushchino, Russia

The bacterial genus *Methylococcus*, which comprises aerobic thermotolerant methanotrophic cocci, was described half-a-century ago. Over the years, a member of this genus, *Methylococcus capsulatus* Bath, has become a major model organism to study genomic and metabolic basis of obligate methanotrophy. High biotechnological potential of fast-growing *Methylococcus* species, mainly as a promising source of feed protein, has also been recognized. Despite this big research attention, the currently cultured *Methylococcus* diversity is represented by members of the two species, *M. capsulatus* and *M. geothermalis*, while finished genome sequences are available only for two strains of these methanotrophs. This study extends the pool of phenotypically characterized *Methylococcus* strains with good-quality genome sequences by contributing four novel isolates of these bacteria from activated sludge, landfill cover soil, and freshwater sediments. The determined genome sizes of novel isolates varied between 3.2 and 4.0 Mb. As revealed by the phylogenomic analysis, strains IO1, BH, and KN2 affiliate with *M. capsulatus*, while strain Mc7 may potentially represent a novel species. Highest temperature optima (45–50°C) and highest growth rates in bioreactor cultures (up to 0.3 h^{-1}) were recorded for strains obtained from activated sludge. The comparative analysis of all complete genomes of *Methylococcus* species revealed 4,485 gene clusters. Of these, pan-genome core comprised 2,331 genes (on average 51.9% of each genome), with the accessory genome containing 846 and 1,308 genes in the shell and the cloud, respectively. Independently of the isolation source, all strains of *M. capsulatus* displayed

surprisingly high genome synteny and a striking similarity in gene content. Strain Mc7 from a landfill cover soil differed from other isolates by the high content of mobile genetic elements in the genome and a number of genome-encoded features missing in *M. capsulatus*, such as sucrose biosynthesis and the ability to scavenge phosphorus and sulfur from the environment.

INTRODUCTION

The genus *Methylococcus* was described in 1966 by Foster and Davis, who isolated a culture of coccus-shaped bacteria capable of growth in a mineral medium with methane (CH_4) as the only carbon source from a sludge sample of the Austin municipal sewage plant, Texas (Foster and Davis, 1966). After description of "*Pseudomonas methanica*" (Dworkin and Foster, 1958; currently known as *Methylomonas methanica*), this was the second documented case of an organism, which consumes CH_4 for growth. The original description of the genus *Methylococcus* characterized its members as non-motile, obligately aerobic, Gram-negative cocci, approximately 1 μm in diameter, with a characteristic diplococcoid arrangement, which utilize methane or methanol as the only carbon sources. The species epithet *capsulatus* was assigned to these bacteria in order to denote the presence of capsules revealed by staining cell specimens with the India ink. The originally described isolate, strain Texas, was defined as the type strain of the species *Methylococcus capsulatus*. Several later proposed species of this genus, such as "*M. bovis*," "*M. chroococcus*," "*M. luteus*," "*M. vinelandii*," and "*M. whittenburyi*" (Romanovskaya et al., 1981), were transferred to the genus *Methylobacter* after detailed re-examination of their characteristics (Bowman et al., 1993). The species *M. thermophilus* (Malashenko et al., 1975) has been retained in the genus *Methylococcus* (Bowman et al., 1993), but the type strain of this species is no longer available from culture collections. The single strain making up the species *Methylococcus mobilis* (Hazeu et al., 1980) has also been lost. Apparently, all *Methylococcus* species are difficult to preserve (Bowman, 2015), which is the main reason behind poor availability of these bacteria from both public and laboratory culture collections. Recently described *Methylococcus geothermalis* (Awala et al., 2020) is the second described species of the genus, which is currently available in culture besides *M. capsulatus*.

Members of the genus *Methylococcus* are thermotolerant or moderately thermophilic bacteria, with optimal growth between 40–60°C (Bowman, 2015). They have been isolated from sewage sludge, sediments of rivers and ponds, wastewater of coal mines, and geothermal fields (Malashenko et al., 1975; Bowman et al., 1993). As suggested by culture-independent studies, *Methylococcus*-like methanotrophs can also be detected in landfill cover soils (Gebert et al., 2009), geothermal soils (Gagliano et al., 2014), and hot springs (Kizilova et al., 2014; Houghton and Stewart, 2020). Cells of these bacteria possess both known types of methane monooxygenase (MMO), which catalyzes the oxidation of methane to methanol, i.e., particulate (pMMO) and soluble (sMMO) forms of this enzyme. Taxonomically, the genus *Methylococcus* belongs to the gammaproteobacterial family *Methylococcaceae*, which accommodates the so-called type I methanotrophs. The latter utilize the ribulose monophosphate pathway (RuMP) for formaldehyde assimilation, while alphaproteobacterial type II methanotrophs employ the serine pathway. One specific feature of the metabolic organization of *Methylococcus* and *Methylococcus*-related methanotrophs is that, in addition to RuMP pathway, they also possess the Calvin–Benson–Bassham (CBB) cycle. This was one of the key reasons to consider these bacteria as representing a separate, type X, methanotrophs (Whittenbury, 1981; Whittenbury and Dalton, 1981). Most cultured representatives of the genus *Methylococcus* did not receive much research attention. One clear exception is *M. capsulatus* strain Bath, which was extensively studied and gradually became probably the best characterized of the aerobic methanotrophs in terms of genetics and physiology (Anthony, 1983; Dalton, 2005; Murrell, 2010). All major insights into MMO structure and function were made by using *M. capsulatus* Bath. Those include studies on resolving the structure of soluble and particulate MMO (Colby and Dalton, 1978; Rosenzweig et al., 1993; Zahn and DiSpirito, 1996; Müller et al., 2002; Chatwood et al., 2004; DiSpirito et al., 2004; Lieberman and Rosenzweig, 2005; Balasubramanian et al., 2010; Ross et al., 2019) as well as control of the MMOs expression based on copper-to-biomass ratios known as a "copper switch" (Stanley et al., 1983; Kao et al., 2004; Larsen and Karlsen, 2016). *Methylococcus capulatus* Bath was also the first methanotrophic organism for which a genome was published (Ward et al., 2004). This bacterium has a relatively small genome (3.3 Mb) compared to other methanotrophs. The genome contains two copies of the gene clusters encoding pMMO and one copy of the sMMO gene cluster as well as *mxa* and *xox* operons encoding two alternative types of methanol dehydrogenases (MDHs). Primary route for carbon assimilation is the RuMP pathway although genes for the serine pathway, and the CBB cycle are also present. Routes for nitrogen assimilation include ammonia assimilation enzymes (glutamine synthetase, glutamate synthase, and alanine dehydrogenase) and nitrogenase. Genome contains genes encoding all tricarboxylic acid (TCA) cycle enzymes, but the absence of 2-oxoglutarate dehydrogenase activity *in vitro* still suggests that the Krebs cycle cannot operate in *M. capsulatus* Bath (Murrell, 2010). To date, two metabolic models were published based on the genome sequence of

M. capsulatus Bath (Gupta et al., 2018; Lieven et al., 2018). Experimental verification of metabolic features of this methanotroph was performed in transcriptomic and proteomic studies (Kao et al., 2004; Larsen and Karlsen, 2016).

High biotechnological potential of fast-growing *Methylococcus* species has also been recognized. In particular, the conversion of methane to biomass by *M. capsulatus* has been exploited for large-scale commercial production of microbial proteins by fermentation (Skrede et al., 1998). The biomass of *M. capsulatus* Bath was approved as a promising source of protein based on criteria such as amino acid composition, digestibility, and animal performance and health (Skrede et al., 2009). Single-cell protein based on *M. capsulatus* is suitable for fish feeding as besides good nutritional value it was proved to prevent soybean meal-induced enteritis in Atlantic salmon (Romarheim et al., 2011). *Methylococcus capsulatus* Bath was also explored for the potential of generating electricity directly from methane (Jawaharraj et al., 2021).

Despite the big research interest in methanotrophs of the genus *Methylococcus*, the number of currently available good-quality genome assemblies of *Methylococcus* species is limited to those of *M. capulatus* Bath and *M. geothermalis* IM1^T. Besides these two finished genomes, 24 draft genome assemblies for members of the genus *Methylococcus* have been deposited in GenBank. Of these, 21 assemblies represent metagenome-assembled genomes (MAGs). This lack of good-quality genome sequences limits the potential of comparative genomic studies as well as our understanding of the variability of genome-encoded features within the genus *Methylococcus*.

This study was initiated in order to expand the narrow range of currently available *Methylococcus* cultures by isolating novel strains of these bacteria from various habitats. Our work resulted in obtaining four new strains of the genus *Methylococcus* and the corresponding complete genomes. These newly obtained genomes as well as all available complete genome sequences of other *Methylococcus* strains were compared in order to examine variability of genome-encoded features within this genus. Despite being isolated from geographically remote habitats, all strains of the species *M. capsulatus* displayed surprisingly high genome synteny and a striking similarity in gene content. By contrast, one novel isolate from a landfill cover soil possessed a number of genome-encoded features missing in *M. capsulatus*, such as sucrose biosynthesis and the ability to scavenge phosphorus and sulfur from the environment. This isolate may potentially represent a novel species of the genus *Methylococcus*.

MATERIALS AND METHODS

Isolation Procedures

In this study, activated sludge, freshwater sediment, and landfill cover soil were used as sources for isolation of new *Methylococcus* strains. Activated sludge samples were collected in bottles (500 ml each) from two municipal wastewater treatment plants in Moscow, Russia. Freshwater sediment samples were collected in 50 ml falcon tubes from beneath shallow water in an unnamed lake in Krasnodar region, Russia. Surface (depth of 0–5 cm)

soil samples were collected from a landfill site in Khanty-Mansiysk region. Aliquots of collected samples were used as inoculum to obtain enrichment cultures of methanotrophic bacteria. The latter were obtained using modified AMS medium (mAMS), containing (in grams per liter of distilled water) NH_4Cl, 0.1; $MgSO_4 \times 7H_2O$, 0.2; $CaCl_2 \times 2H_2O$, 0.02; 100 mM phosphate buffer, pH 5.8, 1% (vol/vol); and trace element solution 0.1% (vol/vol), containing the following (g/L): EDTA, (in grams per liter) EDTA, 5; $FeSO_4 \times 7H_2O$, 2; $ZnSO_4 \times 7H_2O$, 0.1; $MnCl_2 \times 4H_2O$, 0.03; $CoCl_2 \times 6H_2O$, 0.2; $CuSO_4 \times 5H_2O$, 0.1; $NiCl_2 \times 6H_2O$, 0.02; and Na_2MoO_4, 0.03. After inoculation, 500 ml bottles were sealed with silicone rubber septa, and methane was added aseptically using a syringe equipped with a disposable filter (0.22 μm) to achieve a 10–20% mixing ratio in the headspace. Bottles were incubated on a rotary shaker (150 r.p.m.) at 42°C. After 1 week of incubation, the cultures enriched with methanotrophic bacteria were subjected to serial dilutions. After several serial dilution steps, cell suspensions were plated on agar-solidified mAMS medium. The plates were incubated at 42°C in desiccators containing approximately 30% methane in air. The colonies appearing on the plates were picked and restreaked on the same agar medium. The set of finally selected colonies was subjected to several additional serial dilution steps in a liquid mAMS medium at 42°C until isolates of methanotrophic bacteria were obtained. Culture purity was verified by examination using phase-contrast microscopy and by plating on 10-fold diluted Luria–Bertani agar (1.0% tryptone, 0.5% yeast extract, 1.0% NaCl).

Morphological Characterization and Growth Tests

Prior to cell size measurements and growth experiments, the isolates were maintained in liquid mAMS medium with 20% (*v*/*v*) CH_4 at 42°C for 10 days, with regular transfers each 2–3 days. *Methylococcus capulatus* Bath was used as a reference organism in all experiments. Morphological observations and cell-size measurements were made with a Zeiss Axioplan 2 microscope and Axiovision 4.2 software (Zeiss). Cell sizes were measured for 20 randomly selected cells of all strains. Comparative analysis of growth characteristics of the isolates was performed by monitoring their growth dynamics in liquid mineral medium mAMS with 20% methane in the headspace within the temperature range of 25–55°C. Variations in the pH were achieved by mixing 0.1 M solutions of H_3PO_4, KH_2PO_4, and K_2HPO_4. All incubations were performed in triplicate. Growth dynamics was determined by measuring OD_{600} of the cultures on an Eppendorf Biophotometer AG spectrophotometer. The ability to grow on methanol was tested in mAMS medium containing 0.01–3% (*v*/*v*) CH_3OH. Nitrogen sources were tested by replacing NH_4Cl in mAMS with 0.01% (*w*/*v*) $NaNO_3$, $NaNO_2$, urea, methylamine, glutamine, glycine, alanine, peptone, and yeast extract. Growth was examined after 3 days of incubation. Salt tolerance was examined by adding NaCl to mAMS medium in concentrations of 0–2% (*w*/*v*).

Experiments on continuous cultivation were conducted in a 1.5 L bioreactor filled with 1 L of mineral medium. The inlet gases were methane with a flow rate of 100 ml min^{-1} and air

pumped in with an air compressor with a flow rate of 500 ml min^{-1}. The pH level of 5.6 was controlled by titration with 0.8% NH_4OH solution. Agitation was kept constant at 1,000 r.p.m. All gases were filtered with a 0.22 μm sterile membrane. Each of the growth experiments started as a batch cultivation. When the culture reached exponential phase, bioreactor was switched to a continuous mode at a dilution rate of 0.05 h^{-1}. The later was increased with an increment of 0.05 h^{-1} every 1–2 days until a culture washout was observed. After that, the dilution rate was decreased by 0.05 h^{-1} and bioreactor was operated in a continuous mode for the next 5 days.

Hydrogen Utilization

Cells of *Methylococcus* strains were pre-grown to the exponential phase in 120-ml serum bottles containing 20 ml of mAMS medium. Hydrogen utilization was tested with and without the addition of methane under fully aerobic conditions. Headspace of experimental bottles was filled with either 2% CO_2, 2% H_2 and 20% CH_4 or only 2% H_2 and 2% CO_2. Strains were cultivated for 24 h using the same cultivation conditions as described above. Control treatments contained 2% CO_2 or 20% CH_4 and 2% CO_2 in the headspace. Methane and hydrogen were measured by gas chromatography at the beginning of the experiment and after 24 h of cultivation. The optical density was measured at 600 nm wave length (OD_{600}) using Eppendorf Biophotometer UV/Vis spectrophotometer (Eppendorf, Germany).

DNA Extraction

Cultures of new isolates were grown in the liquid mAMS as described above. The cells were harvested after incubation at 42°C on a rotary shaker at 150 rpm for 2 days. Genomic DNA extraction was done using the standard CTAB and phenol-chloroform protocol (Wilson, 2001).

Genome Sequencing and Annotation

Genomic libraries suitable for MiSeq sequencing were prepared with a NEBNext ultra II DNA Library kit (New England Biolabs). On average, a total of 1.68 million paired-end reads (2 × 300; 250 nt) were obtained for each genome. Nanopore sequencing library was prepared using the 1D ligation sequencing kit (SQK-LSK108, Oxford Nanopore, United Kingdom). Sequencing was performed on an R9.4 flow cell (FLO-MIN106) using MinION device. Hybrid assembly of short and long reads was performed using Unicycler v.0.4.8 (Wick et al., 2017). Assemblies were evaluated with Quast 5.0 (Gurevich et al., 2013) and Busco 5.1.2 (Simão et al., 2015).

Annotation was performed using the RAST server (Aziz et al., 2008) and Prokka (Seemann, 2014). In addition, the presence of genes encoding enzymes of the primary and central metabolism as well as some secondary metabolic pathways discussed below was verified manually by web version of NCBI blastp using 35% identity and 50% coverage of amino acid sequence as a cut-off. The same criteria were used for searching genes of interest in the genomes of other methylotrophs. Assembled genomes have been deposited in NCBI GenBank under the accession numbers CP079095–CP079098.

Phylogenomic Analysis

A genome-based tree for members of the *Methylococcus* group was reconstructed using the Genome Taxonomy Database and GTDB-toolkit,[1] release 04-RS89. The maximum likelihood phylogenetic tree was constructed using MegaX software (Kumar et al., 2018). The Pyani program was used to estimate average nucleotide identities across *Methylococcus* genomes (Pritchard et al., 2016). The resulting distance matrices were further visualized as a heatmap in R (R Core Team, 2020).

Pan-Genome Analysis

The pan-genome was reconstructed using microbial pangenomics workflow in Anvi'o (Eren et al., 2015). The genomes were annotated in Prokka, after which the genes were organized using MCL algorithm into core, shell, and singleton clusters (Distance: Euclidean; Linkage: Ward). The core, shell, and singleton genes were separately annotated by BLASTp against the NCBI COG database using eggNOG-mapper (Cantalapiedra et al., 2021). Heatmap based on the annotated COG functions of the core and singleton gene clusters was then plotted in R. The Tettelin best-fit curves of the core- and pan-genomes were constructed using OMCL v1.4 implemented in GET_HOMOLOGUES pipeline (Contreras-Moreira and Vinuesa, 2013).

Whole-genome synteny was computed using Sibelia (Minkin et al., 2013). Synteny blocks were visualized using Circos (Krzywinski et al., 2009).

Identification of the Mobile Genetic Elements and Prophage Regions

Insertion sequences (IS) were identified and classified into IS families using ISsaga pipeline (Varani et al., 2011) with IS finder database (Siguier et al., 2006) and ISEScaner pipeline (Xie and Tang, 2017). ICEfinder and ICEBerg v2.0 were used for IME detection (Liu et al., 2018b). Integron Finder (Cury et al., 2016) was used to identify regions containing integrons. Potential prophage regions were searched with PHASTER server[2] using the settings described in Arndt et al. (2016). Gene functions were determined by homology with known viral proteins in the NCBI GenBank database and the VirFam package[3] using the settings described in Lopes et al. (2014).

RESULTS AND DISCUSSION

Identification and Characterization of New *Methylococcus* Strains

In total, four novel isolates of *Methylococcus*-like methanotrophs were obtained and characterized in this study. Strains IO1 and KN2 were isolated from activated sludge samples of the Moscow municipal sewage plant (**Table 1**). These isolates displayed nearly identical 16S rRNA gene sequences and were closely related to *M. capsulatus* Bath (99.93–99.97% 16S rRNA gene

---

sequence similarity). Strain BH was obtained from a sediment of an unnamed freshwater lake in Krasnodar region, South Russia and was also closely related to *M. capsulatus* Bath (99.02% 16S rRNA gene sequence similarity). Finally, strain Mc7 was isolated from a landfill cover soil in Khanty-Mansiysk region, West Siberia, Russia. In contrast to other three isolates, strain Mc7 displayed highest 16S rRNA gene sequence similarity (98.56%) to *M. geothermalis* IM1[T].

Four novel isolates were represented by non-motile, Gram-negative cocci, with a characteristic diplococcoid arrangement (**Figure 1**). Cell sizes of strains IO1, KN2, and BH were the same as those reported for *M. capsulatus*, approximately 1 μm in diameter (**Table 1; Figure 1A**). Cell sizes of strain Mc7 (1.6 μm in diameter; **Figure 1B**) were larger than those of other three strains. The colonies formed by these isolates on agar mineral medium after 2 weeks of incubation with 20% (*v/v*) methane were small (1–2 mm in diameter), round, milk-white (for strain Mc7) or cream colored (for strains IO1, KN2, and BH).

Methane and methanol were used as growth substrates, although the isolates differed in their ability to grow on methanol. Similar to *M. capsulatus* Bath, strains IO1 and Mc7 displayed only a trace growth (up to OD_{600} 0.10–0.12) in a relatively

narrow range of CH_3OH concentrations (**Table 1**). By contrast, strains KN2 and BH were capable of a relatively good growth (up to OD_{600} 0.3) on methanol in a wide range of concentrations. The isolates were quite uniform with regard to their pH preferences (**Table 1**). Most strains grew in the pH range of 5.5–8.0. Strain KN2 demonstrated slightly better adaptation to moderately acidic conditions growing between pH 4.6 and 8.5. All strains used nitrate and ammonium as nitrogen sources. In addition, strains BH and KN2 were able to utilize glutamine, while strain IO1 showed weak growth on peptone as a nitrogen source. Strain Mc7 displayed highest tolerance to NaCl, up to 1% (*w/v*), while other isolates tolerated up to 0.75% (*w/v*) NaCl. For comparison, NaCl tolerance determined in our experiments for *M. capsulatus* Bath was 0.5% (*w/v*).

The ability to utilize hydrogen was tested using incubations with either H_2 and CO_2 or H_2, CO_2 and CH_4 in the headspace (see Materials and Methods). No growth as well as no H_2 consumption were observed in the absence of CH_4, with H_2 as the only energy source. With the only exception of strain KN2, the presence of H_2 in addition to CH_4 increased the growth yield of all strains examined in this study by 2–20% (**Supplementary Figure S1A**) as well as the corresponding CH_4 consumption (**Supplementary Figure S1B**). H_2 utilization

TABLE 1 | Isolation sources and some characteristics of new *Methylococcus* isolates.

Strain	Isolation source	Cell size, μm	Growth temperature range (optimum), °C	Growth on methanol, % CH_3OH	NaCl tolerance, %	pH growth range (optimum)	Growth rate in batch culture/ bioreactor, h^{-1}
IO1	Activated sludge, Moscow, Russia	1.14 ± 0.02	28–52 (48)	0.05–0.1 (trace)	0.75	5.5–7.5 (6.5–7.0)	0.19/0.26
KN2	Activated sludge, Moscow, Russia	1.12 ± 0.03	25–53 (48–50)	0.05–3.0	0.75	4.6–8.5 (6.5–8.0)	0.22/0.30
BH	Lake sediment, Krasnodar region, Russia	1.12 ± 0.02	25–52 (42)	0.05–1.5	0.75	5.5–8.0 (6.5–7.2)	0.18/0.28
Mc7	Landfill cover soil, Khanty-Mansiysk, Russia	1.59 ± 0.03	28–53 (40)	0.05–0.1 (trace)	1.0	5.5–7.5 (6.5–7.0)	0.23/0.27
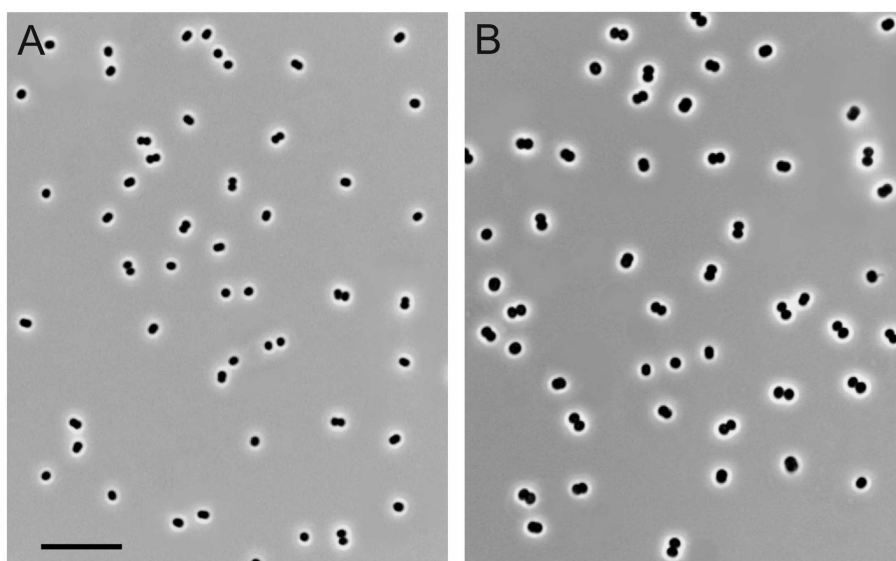
FIGURE 1 | Cell morphology of strains KN2 **(A)** and Mc7 **(B)** grown for 1 day in liquid medium mAMS with methane. Marker, 10 μm.

in the presence of CH_4 was observed for all studied strains, including strain KN2 (**Supplementary Figure S1C**). These results agree well with the previous report of the absence of autotrophic growth of *M. capsulatus* Bath in liquid medium with H_2 and CO_2 (Baxter et al., 2002; Henard et al., 2021). Apparently, *Methylococcus* strains are capable of using H_2 as an alternative energy source during their growth on methane. Recently, the ability to grow mixotrophically on H_2 and CH_4 was also reported for verrucomicrobial (*Methylacidiphilum* sp. RTK17.1) and proteobacterial (*Methylocystis* sp. strain SC2) methanotrophs (Carere et al., 2017; Mohammadi et al., 2019; Hakobyan et al., 2020).

The special attention in our study was given to assessing growth characteristics of novel *Methylococcus* isolates. The results of measuring specific growth rates of novel isolates and *M. capsulatus* Bath within the temperature range of 25–55°C are shown in **Figure 2**. The two strains obtained from sludge samples, IO1 and KN2, were clearly more thermotolerant than *M. capsulatus* Bath and displayed growth optima at 48–52°C. The lowest temperature growth optimum of 40°C was revealed for strain Mc7, which is reasonable given that it was isolated from a landfill soil in West Siberia. Highest specific growth rates, 0.22–0.23 h^{-1}, which were comparable to that of *M. capsulatus* Bath, were recorded for strains KN2 and Mc7. We also performed additional experiments in order to examine specific growth rates of novel isolates during continuous cultivation on methane in 1.5 L bioreactor. The highest specific growth rate in a fermenter culture (0.3 h^{-1}) was recorded for strain KN2. Other strains also demonstrated stable growth under conditions of continuous cultivation. Their growth rates, however, were below that demonstrated by strain KN2 (**Table 1**).

Genome Sequencing and Assembly

The genomes of four novel isolates were sequenced using a hybrid approach. Oxford Nanopore sequencing yielded 192,565–278,437 reads with a total length of 1.05–1.59 Gb (**Supplementary Table S1**). Sequencing on Illumina MiSeq platform generated a total of 415,244–3,477,384 paired-end reads, with a mean read length of 250 bp. Both short and long reads were combined to perform a hybrid assembly using Unicycler, resulting in circular genomes. Genome characteristics of new isolates are summarized in **Table 2**.

Genome sizes varied from 3.2 Mb in strain BH to 4.0 Mb in strain Mc7. The DNA G + C content was highly similar in all examined genomes and constituted 63.44–63.52%. Each genome contained two copies of rRNA operon, two copies of pMMO, and one copy of sMMO. The number of protein-coding genes varied between 2,912 and 3,752. No plasmids were detected. Lowest number (17) of insertion (IS) elements was detected in strain KN2, while substantially higher number of IS elements (78) was observed in strain Mc7.

Genome-Based Phylogeny and Genome-to-Genome Comparison

The genome-based phylogeny of four novel *Methylococcus*-affiliated isolates was determined using the comparative sequence analysis of 120 ubiquitous single-copy proteins (**Figure 3**). Three isolates (strains IO1, KN2, BH) were clustered together with two strains of *M. capsulatus*, Texas[T], and Bath, whereas strain Mc7 and *M. geothermalis* IM1[T] constituted two separate branches within the genus *Methylococcus*. The complete genome of strain EFPC2, which is annotated in the GenBank as belonging to *Methylococcus* species, was also included in the analysis.
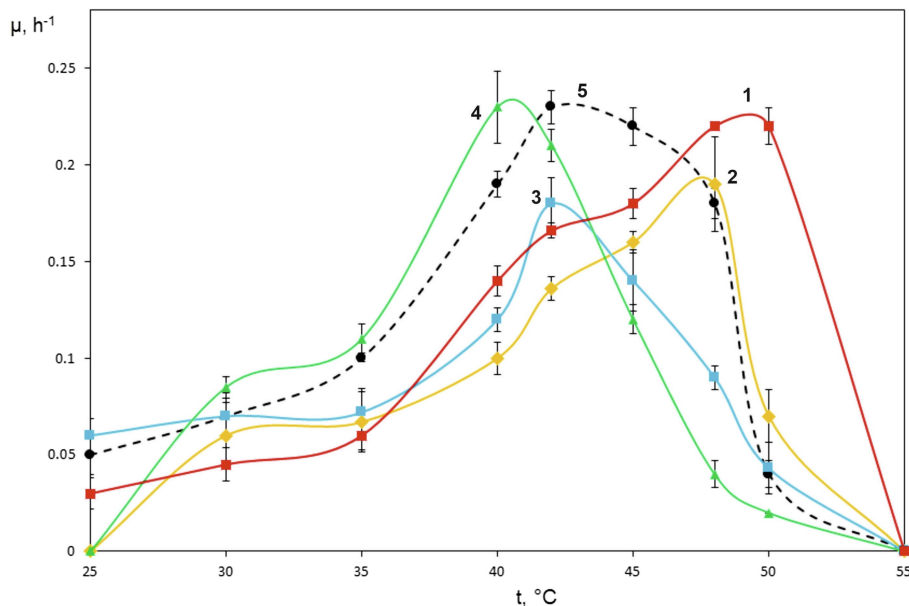
FIGURE 2 | Specific growth rates of novel isolates on methane as dependent on incubation temperature. *Methylococcus capsulatus* Bath was used as a reference organism. 1, strain KN2, 2, strain IO1, 3, BH, 4, Mc7, and 5, *Methylococcus capsulatus* Bath.

	Strain KN2	**Strain BH**	**Strain Mc7**	**Strain IO1**	***Methylococcus capsulatus* Bath**	***Methylococcus geothermalis* IM1[T]**
Genome size (Mb)	3.6	3.2	4.0	3.3	3.3	3.4
Contigs	1	1	1	1	1	1
G+C content (mol %)	63.49	63.52	63.44	63.49	63.60	63.25
CDS	3,289	2,912	3,752	3,008	3,043	3,091
Repeat region	2	2	6	2	2	3
tRNA	48	50	48	50	49	48
5S, 16S, 23S	2, 2, 2	2, 2, 2	2, 2, 2	2, 2, 2	2, 2, 2	2, 2, 2
pMMO operon	2	2	2	2	2	2
sMMO operon	1	1	1	1	1	1
Complete IS elements	17	23	78	26	33	45
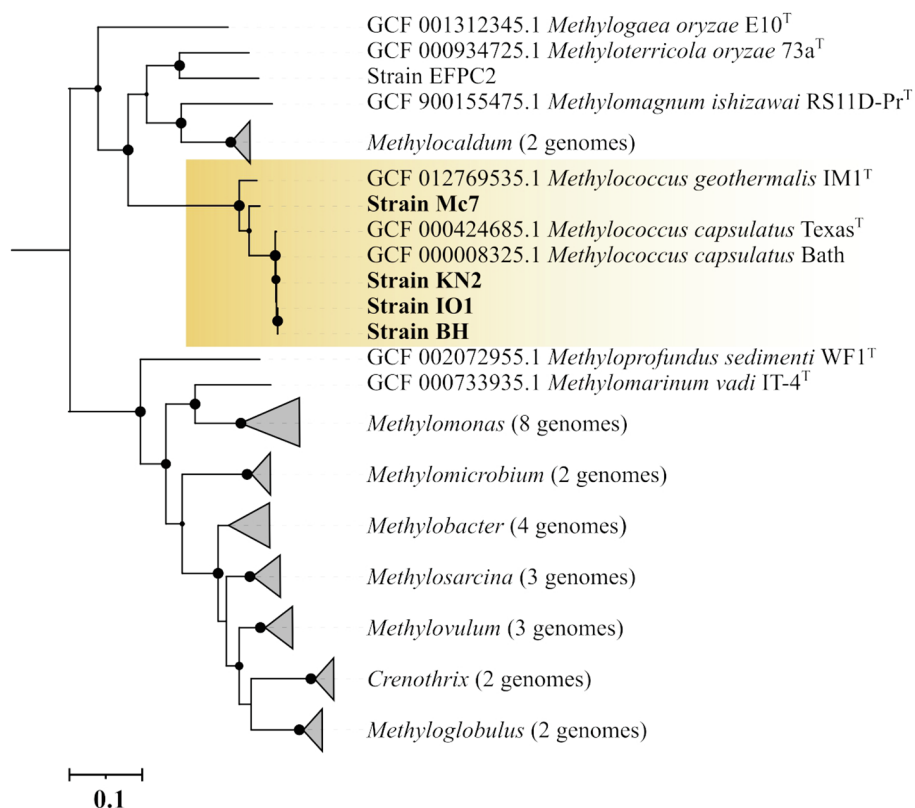
FIGURE 3 | Genome-based phylogeny showing the position of new isolates in relation to other gammaproteobacterial methanotrophs based on the comparative sequence analysis of 120 ubiquitous single-copy proteins. The clade of *Methylococcus* methanotrophs is highlighted by orange. The tree was constructed using the Genome Taxonomy Database toolkit (Parks et al., 2018), release 04-RS89. The significance levels of interior branch points obtained in maximum-likelihood analysis were determined by bootstrap analysis (100 data re-samplings). Bootstrap values of >70% are shown. The root (not shown) is composed of all genomes available in GTDB for methanotrophs of the genera *Methylosinus* and *Methylocystis*. Bar, 0.1 substitutions per amino acid position.

However, as seen from **Figure 3**, strain EFPC2 clustered together with *Methyloterricola oryzae* 73a[T] and did not affiliate with *Methylococcus* clade. The genome of strain EFPC2, therefore, was excluded from further comparative analysis of finished genomes available for *Methylococcus* species.

The average nucleotide identity (ANI) values were estimated for each pair of genomes (**Supplementary Figure S2**). ANI values calculated for isolates IO, KN2, BH, and *M. capsulatus* Bath were within a range of 98.75–99.73% which indicates that these organisms belong to the same species as intra-species level is

defined at ≥95% ANI (Konstantinidis and Tiedje, 2005; Goris et al., 2007). ANI value determined for separately clustered strain Mc7 and *M. geothermalis* IM1[T] was 88.56%. Thus, strain Mc7 may potentially represent a novel species of the genus *Methylococcus*.

The genomes of novel isolates, *M. capsulatus* Bath and *M. geothermalis* IM1[T], were included in the whole-genome synteny analysis in order to explore the evolution of genome structure. Overall, 216 shared synteny regions were identified. Among those, 86 regions were common for all six genomes. **Figure 4** depicts only the synteny blocks affiliated with *M. capsulatus* Bath.
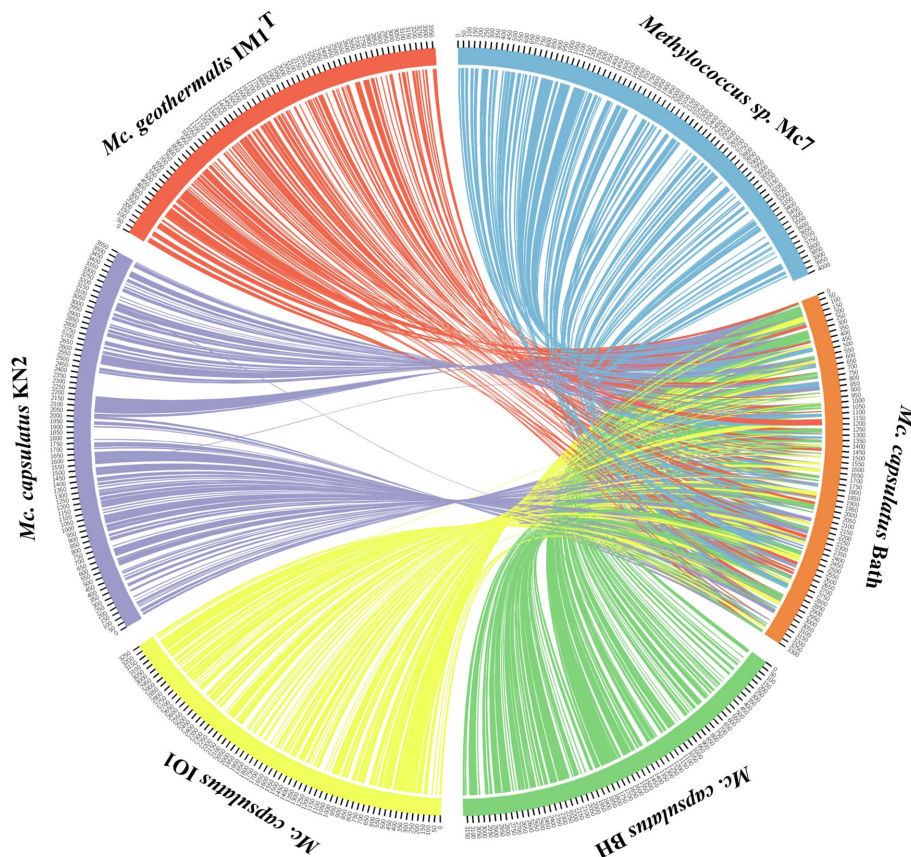
FIGURE 4 | Visualization of synteny blocks between *Methylococcus capsualtus* Bath and other examined *Methylococcus* genomes. Each link represents a single synteny block. Circle is divided into the individual arcs representing genomes of *Methylococcus capsulatus* Bath (orange), *Methylococcus capsulatus* BH (green), *Methylococcus capsulatus* IO1 (yellow), *Methylococcus capsulatus* KN2 (purple), *Methylococcus geothermalis* IM1ᵀ (red), and *Methylococcus* sp. Mc7 (blue).

The number of synteny blocks revealed between *M. capsulatus* Bath and each of the isolates IO1, KN2, and BH was 102–103. Strain Mc7 and *M. geothermalis* IM1ᵀ exhibited slightly lower levels of synteny to *M. capsulatus* Bath by displaying 97 and 93 synteny blocks, respectively. One of *M. capsulatus*-related isolates, strain KN2, and *Methylococcus* sp. Mc7 were chosen for more detailed analysis of whole-genome synteny patterns (**Figure 5**). As expected, the calculations made for two genomes resulted in lower number of synteny blocks, which were larger in size. Synteny regions shared between strain KN2 and *M. capsulatus* Bath covered 88.9 and 94.9% genomes of these bacteria, respectively. The largest synteny block in the genome of strain KN2 spanned about 1.1 Mb and contained 978 genes. Two relatively large synteny break regions in the genome of strain KN2 spanned 224 and 76 kb. Synteny breaks are caused by rearrangements, the insertion of novel genes, or the presence of genes that are too diverged to establish an orthologous relationship or have undergone expansion or loss (Liu et al., 2018a). Of 224 genes in the first synteny break region, 187 genes encoded hypothetical proteins. This region also included several genes coding for transposases, formate dehydrogenase displaying 95.5% amino acid sequence identity to NAD-dependent formate dehydrogenase from *Methylocaldum marinum*, flagellar

transcriptional regulator FlhC, lysozyme RrrD, several helicases, and endonucleases. Of 84 genes in the second synteny break region in the genome of strain KN2, 76 genes encoded hypothetical proteins, while others coded for transport proteins. Comparison of the genomes of strain Mc7 and *M. capsulatus* Bath revealed 81 synteny blocks, which corresponded to nearly 50% of their genome length (**Figure 5**). The largest synteny break regions spanned 98.9, 98.5, 79.7, and 74.4 kb. Approximately half of the genes within these regions encoded hypothetical proteins. Interestingly, 98.9 kb region contained genes encoding small and large subunits of ribulose bisphosphate carboxylase.

Pan-Genome Analysis

The Anvi'o pangenomics workflow was used to cluster protein-coding sequences into core, shell, and singleton genomes. Of 4,485 identified gene clusters, the *Methylococcus* pan-genome core comprised 2,331 genes (on average 51.9% of each genome), with the shell genome containing 846 gene clusters (18.9% of total gene clusters) and the cloud containing 1,308 gene clusters (29.2% of total gene clusters; **Figure 6**).

Pan-genome of the genus *Methylococcus* is open. **Supplementary Figure S3** displays the number of core genes (A) and the pan-genome size (B) as a function of the

FIGURE 5 | Visualization of synteny blocks between *Methylococcus capsulatus* Bath, *Methylococcus capsualtus* KN2, and *Methylococcus* sp. Mc7. Each link represents a single synteny block. Each circle is divided into two parts, which represent two compared genomes. The genome of *Methylococcus capsulatus* Bath is shown in orange, while the genomes of *Methylococcus capsulatus* KN2 and *Methylococcus* sp. Mc7 are displayed in purple and blue, respectively.

number of included genomes. Fitting the curve in **Supplementary Figure S3B** to a power law ($3{,}146.3n^{0.171}$) allows extrapolating calculations to *n* genomes. According to approximation, sequencing of one additional genome would provide 100 new genes to pan-genome. If the number of available genomes constituted 50 or 100 genomes, sequencing of one additional genome would contribute 21 and 12 genes, respectively. The core, shell, and cloud gene clusters were further annotated into COG classes (**Supplementary Figure S4**). The core genome was mostly conserved in the following: energy production and conversion (11.5% of total core gene clusters); translation, ribosomal structure, and biogenesis (8.6%); cell wall/membrane/envelope biogenesis (8.3%); coenzyme transport and metabolism (7.3%); inorganic ion transport and metabolism (6.3%); amino acid transport and metabolism (5.3%); posttranslational modification; protein turnover, chaperones (4.9%). However, the most abundant category was the one with no functional prediction (19.9%). The most abundantly represented functional categories in shell genome were replication, recombination, and repair, coenzyme transport and metabolism, signal transduction mechanisms, transcription, inorganic ion transport and metabolism, cell wall/ membrane/envelope biogenesis, energy production and conversion. The majority of cloud genes were concentrated in the genomes of strain Mc7, *M. geothermalis* IM1[T], and *M. capsulatus* KN2 and represented replication, transcription, cell wall/membrane/ envelope biogenesis, and signal transduction mechanisms. These functional categories were also affiliated with few cloud genes found in *M. capsulatus* Bath and strain BH. A number of studies demonstrated that core genes systematically represented the smallest fraction of pan-genome (Bazinet, 2017;

Livingstone et al., 2018; Kumar et al., 2020; Oshkin et al., 2020). The large proportion of core genes in *Methylococcus* pan-genome may be explained by both the small genome size of individual *Methylococcus* strains and the low number of available genomes that could be taken for analysis. However, approximation for core- and pan-genome (**Supplementary Figure S3**) suggests that even if 30 genomes were included into pan-genome analysis, the fraction of core genes would still be high (41.1% of the total gene clusters).

Methane Oxidation Enzymes

Two complete copies of the *pmoCAB* gene cluster encoding pMMO and one additional copy of the *pmoC* (*pmoC3*) gene were revealed in the genomes of all novel *Methylococcus* strains. The corresponding *pmoA* sequences were nearly identical to each other and to the *pmoA* sequences from *M. capsulatus* Bath and *M. geothermalis* IM1[T]. The sMMO-encoding gene cluster *mmoXYBZDC* was also present in all examined genomes, with MmoX sequences exhibiting high level of homology (100% identity between MmoX from *M. capsulatus* Bath and strains KN2, IO1, BH, and 97% identity with that from strain Mc7). This chromosomal locus in *M. capsulatus* Bath and strain IO1 contained also one gene encoding a hypothetic protein between *mmoB* and *mmoZ* genes. The gene cluster *mmoGQSR* coding for the large subunit of bacterial chaperonin GroEL (*mmoG*), the two-component sensory/regulatory system (*mmoQ* and *mmoS*), and the transcription activator (*mmoR*; Csáki et al., 2003) was found in all examined genomes.

Two PQQ-dependent MDHs catalyze the second stage of C_1-oxidation. The structural components of heterodimeric

FIGURE 6 | Pan-genome analysis. Clustering of genomes based on the presence/absence patterns of 4,485 pan-genomic clusters. The genomes are organized in radial layers as core, singleton, and shell gene clusters (Euclidean distance; Ward linkage) which are defined by the gene tree in the center. Genomes of new isolates are colored by blue. Genomes of *Methylococcus capsulatus* Bath and *Methylococcus geothermalis* IM1^T are colored by dark blue. Line segments indicate single gene clusters. The outermost circles carry the information regarding the number of genomes containing specific gene cluster, number of genes in gene clusters (gray), functional, geometric, and combined homogeneity indexes (purple), distribution of single-copy gene clusters (orange circle). The layers below the heatmap display from the bottom to top: total length of the genomes, GC content, genomes completion, redundancy of each genome, number of genes per kbp, number of singleton gene clusters, total number of gene clusters. Heatmap denotes correlation between the genomes based on average nucleotide identity values calculated using pyANI.

Ca^{2+}-dependent MDH (MxaFI type) and the proteins required for its catalytic activity are encoded by the *mxaFJGIRACKLD* cluster with gene organization completely identical in all *Methylococcus* strains. Namely, the *mxaG* gene encodes the specific electron acceptor cytochrome c$_L$, the *mxaACKL* gene cluster encodes the proteins required for Ca^{2+} incorporation into the active site, while the *mxaRS* gene cluster encodes two proteins of unknown functions (Chistoserdova, 2011; Vuilleumier et al., 2012). The four strains harbor gene cluster *xoxFJ*, which encodes an alternative MDH (XoxF) containing a rare earth element in the active site (Hibi et al., 2011), and a periplasmic solute-binding protein (XoxJ). No other genes relevant to methanol oxidation were present in this locus. The *xoxG* gene coding for a putative electron acceptor from XoxF and being

only distantly related to the *mxaG* gene (Keltjens et al., 2014; Yu et al., 2017; Zheng et al., 2018) was found in all genomes and was located separately from *xoxFJ*.

Oxidative transformations of formaldehyde are mediated by the enzymes of the tetrahydrofolate (THF)-based pathway whose amino acid sequences in the novel isolates are almost identical to those in *M. capsulatus* Bath. The complete sets of genes for the tetrahydromethanopterin (THMP)-dependent pathway of formaldehyde oxidation are present in all *Methylococcus* genomes. Each strain possesses the three-subunit formate dehydrogenase, which catalyzes the last stage of methane oxidation, and a five-gene operon encoding the cytoplasmic NAD$^+$-dependent formate dehydrogenase composed of (*abc*)$_2$ subunits that catalyze the reversible reaction of formate oxidation to CO$_2$

(Hartmann and Leimkühler, 2013). All strains of *M. capsulatus* (Bath, KN2, IO1, and BH) possess also the two-subunit formate dehydrogenase, whereas this enzyme is missing in strain Mc7 and *M. geotermalis* IM1[T]. The genomes of three strains, Mc7, KN2, and *M. geothermalis* IM1[T], encode the one-subunit formate dehydrogenase, which is not present in *M. capsulatus* Bath, IO1 and BH. Thus, strain KN2 possesses four formate dehydrogenases in comparison with other *Methylococcus* strains possessing three orthologs each. Since no other evident differences in the arrays of C1 oxidizing enzymes of these strains were found, one may hypothesize that the extended array of formate dehydrogenases helps increasing removal of C1 metabolites *via* their oxidation to CO_2. However, the exact role of different formate dehydrogenases in metabolism of methanotrophic bacteria remains to be elucidated.

Carbon Assimilation Pathways

Ribulose Monophosphate Cycle

Genes encoding enzymes of the RuMP cycle are linearly duplicated in all studied *Methylococcus* genomes. The central part of this duplicated DNA locus (~5,000 bp) contains the transaldolase gene, which is surrounded by the genes of phosphohexulose isomerase (*phi*), hexulose phosphate synthase (*hps*), fructose bisphosphate aldolase and transketolase (Ward et al., 2004). Each of the studied strains has a gene encoding the fusion protein hexulose phosphate synthase/isomerase. Further research is needed to clarify if this gene multiplicity for the RuMP cycle affects growth characteristics of methanotrophs.

Serine Pathway

All examined *Methylococcus* genomes contain the *sga–hpr–gck3* gene cluster coding for the serine-glyoxylate aminotransferase, hydroxypyruvate reductase, and 3-glycerate kinase, which are the key enzymes of the serine pathway. They also harbor genes for malyl-CoA lyase and malate thiokinase responsible for glyoxylate biosynthesis but lack PEP carboxylase. Pyruvate carboxylase that can fulfill the function of replenishing C_4 TCA intermediates is present only in strain Mc7. Although this function can be assigned to oxaloacetate decarboxylase (Ward et al., 2004) or to malic enzyme, these enzymes catalyze predominantly irreversible decarboxylation reactions (Rozova et al., 2019; Xu et al., 2020). The gene for PPi-dependent PEP carboxykinase was not detected in the examined *Methylococcus* genomes. The absence of genes coding for any of the currently known C_3 carboxylation enzymes in most *Methylococcus* representatives requires further efforts to decipher mechanisms for replenishing TCA cycle intermediates. In addition, glyoxylate can be generated from phosphoglycolate, which is the product of the oxygenase reaction of ribulose bisphosphate carboxylase (RuBisCo) as previously evidenced for *M. capsulatus* Bath (Taylor et al., 1981; Baxter et al., 2002). The genes for phosphoglycolate phosphatase and glycolate oxidase are present in the genomes of all *Methylococcus* representatives.

Calvin Cycle

Genomes of all *Methylococcus* strains possess *cbbL* and *cbbS* genes for the large and the small subunits of RuBisCo, as well

as the *cbbQ* gene encoding a polypeptide putatively acting as a posttranslational RuBisCo activator (Baxter et al., 2002). Additional copies of these genes are present only in the genome of strain Mc7. Six genes encoding carbonic anhydrases were found in strain KN2, five genes in other *M. capsulatus* strains, and only three enzymes are encoded by Mc7 and IM1. Recent study suggests that RuBisCO is essential for *M. capsulatus* Bath growth and central metabolites derived from CO_2 enter core intermediary metabolic pathways, including the Embden–Meyerhof–Parnas (EMP) glycolytic pathway, the pentose phosphate pathway, and the TCA cycle (Henard et al., 2021).

Central Metabolism

Several pathways for C_6-phosphosugars degradation are encoded in the genomes of four novel *Methylococcus* isolates. These include the modified Embden–Meyerhof–Parnas pathway, where PPi-dependent phosphofructokinase catalyzes the reaction of fructose-1,6-phosphate synthesis from fructose-6-phosphate (Reshetnikov et al., 2008), the Entner–Doudoroff pathway, and the phosphoketolase pathway. The genomes of strain Mc7 and *M. geothermalis* IM1[T] encode four isozymes of glucose-6-phosphate dehydrogenase, while all strains of *M. capsulatus* (Bath, KN2, IO1, BH) possess only two of these isozymes. The genomes of novel isolates encode phosphoketolase catalyzing cleavage of fructose-6-phosphate or xylulose-5-phosphate into glyceraldehyde-3-phosphate/erythrose-4-phosphate and acetyl phosphate; the genes coding for phosphoketolase and acetate kinase comprise a single cluster. These strains, therefore, are predicted to be able to convert acetylphosphate to acetate and to produce ATP as previously suggested for other gammaproteobacterial methanotrophs (Rozova et al., 2015; Henard et al., 2017). In strain Mc7 and *M. geothermalis* IM1[T], acetyl phosphate can be directly converted into acetyl-CoA without ATP consumption, while strains of *M. capsulatus* lack the respective acetyl phosphotransferase. In all *Methylococcus* species, the synthesis of acetyl-CoA can also proceed from acetate by ATP-dependent acetyl-CoA synthetase forming AMP. Interestingly, two strains, KN2 and Bath, code for two enolases, while other strains possess only one enzyme. The relative excess of enolase, which directs the primary C3 intermediates of the Calvin cycle and the serine pathway to the central metabolism, most likely, can increase the impact of these "minor" pathways in overall carbon assimilation of these methanotrophs.

It is known that gammaproteobacterial methanotrophs are capable of fermentation under micro-oxic conditions (Kalyuzhnaya et al., 2013). The genomes of novel *Methylococcus* isolates contain the same genes potentially involved in fermentation (pyruvate formate lyase, alcohol dehydrogenase, acetate kinase) as *M. capsulatus* Bath (Ward et al., 2004). However, similar to strain Bath, these methanotrophs lack the lactate dehydrogenase gene.

Nitrogen Metabolism

All *Methylococcus* strains possess the structural genes for nitrogenase (*nifHDK*) apparently constituting an operon together with *nifENX*; the latter being involved in synthesis of the

nitrogenase iron-molybdenum cofactor. One gene of unknown function is present between *nifK* and *nifE* in strains Bath, BH, KN2, and IO1, whereas the same locus in strain Mc7 contains three genes. All strains possess the gene cluster responsible for nitrate reduction, i.e., nitrate reductase (NasA), nitrite reductase (NirBD), protein kinase and nitrate transporter. All strains contain the genes *haoAB* for putative hydroxylamine dehydrogenase that transforms hydroxylamine to nitric oxide. One potential source of hydroxylamine is ammonium oxidation by pMMO since these methanotrophs lack ammonia monooxygenase. The gene cluster *norBC* responsible for reduction in nitric oxide to nitrous oxide is also present in the genomes. Like *M. capsulatus* Bath (Ward et al., 2004), novel strains encode three predicted hydrogenases: a multisubunit formate hydrogen lyase, most likely involved in the conversion of formate to dihydrogen and carbon dioxide; a soluble cytoplasmic NAD-reducing hydrogenase, which transfers electrons to NAD; and a membrane-bound Ni-Fe hydrogenase.

Secondary Metabolism

The genomes of novel isolates as well as the genome of *M. capsulatus* Bath encode enzymes of glycogen synthesis, i.e., 4-α-glucanotransferase (amylomaltase), glucose-1-phosphate adenylyltransferase, 1,4-α-glucan-branching enzyme GlgB, and glycogen synthase. The corresponding genes are organized in one cluster similar to that in *M. capsulatus* Bath. All strains, including *M. capsulatus* Bath, also harbor an additional gene cluster, which encodes glycogen synthase 2 and α-amylase. Among examined *Methylococcus* representatives, only strain Mc7 possesses genes coding for sucrose biosynthesis and degradation. The genes for sucrose phosphate synthase, sucrose phosphate phosphatase, and sucrose synthase are assembled in one gene cluster in strain Mc7 similar to that in *Methylocaldum szegediense* O12 (But et al., 2018). Like the latter, strain Mc7 possesses the gene coding for mannose-1-phosphate

guanyltransferase; however, it lacks fructokinase encoding gene. The ability to synthesize sucrose may be one of the reasons behind the enhanced salt tolerance observed for strain Mc7 in comparison with other *Methylococcus* isolates (**Table 1**).

The genomes of the studied methanotrophs encode enzymes homologous (40–42% sequence identity) to the previously characterized acetolactate decarboxylase from *Klebsiella aerogenes* (Blomqvist et al., 1993). Interestingly, in *Methylococcus* species, the acetolactate decarboxylase gene forms a cluster together with genes of the malic enzyme and acetolactate synthase; the latter known to produce acetolactate from two molecules of pyruvate with CO_2 release (Blomqvist et al., 1993). Given that acetolactate decarboxylase degrades acetolactate into acetoin and CO_2, functioning of a fermentation pathway that converts malate to acetoin through pyruvate and acetolactate appears to be possible. The BLAST search revealed that, besides methanotrophs of the genus *Methylococcus*, acetolactate decarboxylase is also encoded in the genomes of *Methyloterricola oryzae*, *Methylogaea oryzae* as well as in *Methylovulum* and *Methyloprofundus* representatives. In microorganisms, the pathway of acetoin synthesis is regarded as mechanism preventing over-acidification of the intracellular environment, and also as an energy-storing strategy (Xiao and Xu, 2007).

In strain Mc7, DNA locus of ~14,000 bp is represented by the 13-gene cluster *phnFFCEGHIJKLMNP*, which displays high structural and sequence similarity (27–54% amino acid sequence identity) to the *Escherichia coli* operon encoding the carbon-phosphorus (CP) lyase (**Figure 7**). In *E. coli*, this gene cluster was found to determine degradation of organophosphonates (Adams et al., 2008; Jochimsen et al., 2011) and we may assume the same function of the corresponding genes in strain Mc7. BLAST search revealed that, among gammaproteobacterial methanotrophs, these genes are present only in the genome of marine methanotroph *Methyloprofundus sedimenti* WF1[T]. Organophosphonates are widely distributed in nature
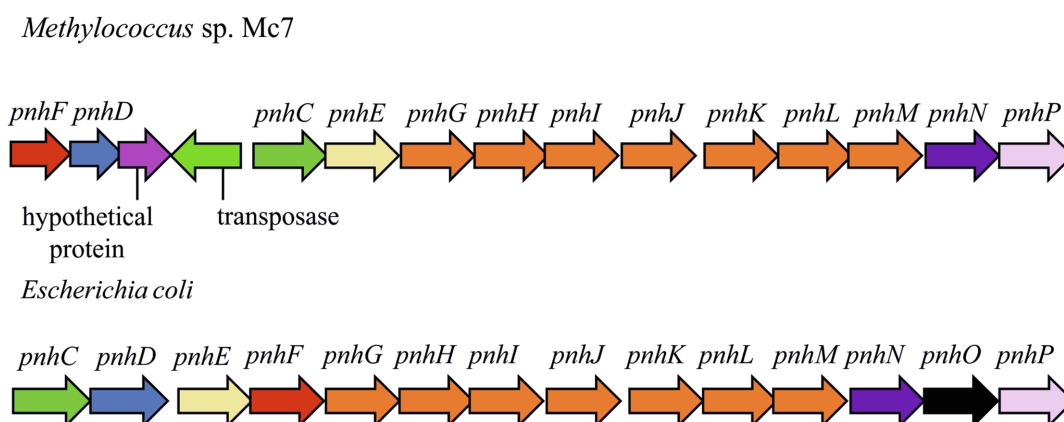
FIGURE 7 | The gene clusters carrying genes of pathway for organophosphonates degradation in methylotrophs and *Escherichia coli*. The products of the *phn* genes were proposed based on the comparison with the respective sequences in the *Escherichia coli* operon (Jochimsen et al., 2011): phosphonate ABC transporter ATP binding subunit (*phnC*); phosphonate ABC transporter periplasmic binding protein (*phnD*); phosphonate transport system permease protein (*phnE*); putative transcriptional regulator (*phnF*); carbon-phosphorus lyase core complex subunits (*phnGHIJK*); methylphosphonate degradation complex subunit (*phnL*); RPnTP hydrolase (*phnM*); ribose-1,5-bisphosphate phosphokinase (*phnN*); aminoalkylphosphonate N-acetyltransferase (phnO); 5-phospho-alpha-D-ribosyl-1,2-cyclic phosphate phosphodiesterase (*phnP*).

and are represented by various naturally occurring compounds (phosphonopyruvate, 2-aminoethylphosphonate, and phosphonoacetate; Ternan et al., 1998; Seto and Kuzuyama, 1999) as well as by synthetic xenobiotics such as herbicide glyphosate ([N-(phosphonomethyl) glycine]). Obviously, the potential substrates of CP lyase from strain Mc7 and *M. sedimenti* WF1[T] are naturally occurring organophosphonates. The ability to decompose a highly stable carbon-phosphorous bond allows survival in habitats with low contents of phosphates.

The genomes of strain Mc7 and *M. geothermalis* IM1[T] contain the genes for haloalkane dehalogenase, which catalyzes the hydrolytic cleavage of carbon–halogen bonds (EC 3.8.1.5). It is absent in *M. capsulatus* strains, whereas its homologues were found in the genomes of *Methylocaldum* spp. (with ~80% identity), *Methylobacter tundripaludum* and *Methylovulum myaconense* (<70%) as well as in alphaproteobacterial methanotrophs (<50%). The corresponding protein shares 51% identity with haloalkane dehalogenase characterized from *Bradyrhizobium japonicum* (Sato et al., 2005).

The genome of strain Mc7 contains the gene for putative methanethiol S-methyltransferase, which catalyzes the transmethylation between methanethiol (MeSH) and S-adenosyl-L-methionine into dimethylsulfide (DMS) and S-adenosyl-L-homocysteine as well as the genes for FMNH$_2$-dependent dimethylsulfone monooxygenase and FMNH$_2$-dependent monooxygenase SfnG. MeSH can be formed by sulfide methylation in anaerobic environments or by degradation of sulfur-containing amino acids, *via* the cleavage of dimethylsulfopropionate, which is osmoprotector in marine microalgae, or by demethiolation of sulfhydryl groups (Lomans et al., 2001, 2002; Bentley and Chasteen, 2004). Further metabolism of DMS can proceed *via* chemical or biochemical oxidation into dimethylsulfone. The latter compound, being substrate for FMNH2-dependent monooxygenases, is converted into inorganic sulfite and formaldehyde. Obviously, this mechanism of scavenging sulfur from organosulfur compounds enables bacterial survival in sulfate-depleted environments (Boden et al., 2011; Carrión et al., 2015). All strains under study possess genes for the sulfur oxidizing (Sox) system that allows utilization of inorganic sulfur compounds in energy metabolism. In these methanotrophs, the Sox system is represented by the SoxYZ complex that presumably carries the intermediates of the pathway on a cysteine residue near the C terminus of SoxY as well as by sulfane dehydrogenase SoxCD; the latter has been found to catalyze a six-electron oxidation reaction (Friedrich et al., 2005). However, all *Methylococcus* strains lack SoxB, SoxA, and SoxX components needed for the full Sox system. Functionality and role of the truncated pathway of sulfur oxidative metabolism in the methanotrophs is not clear.

Integrons

Integrons are gene-capturing platforms that play a significant role in generating phenotypic diversity and shaping adaptive responses in microorganisms, including the spread of antibiotic resistance genes (Mazel, 2006; Gillings, 2014). The integron structure includes variable gene cassette array (recombination site attC with corresponding genes) and a stable platform containing integrase (IntI), recombination site (attI) and promoter (Boucher et al., 2007). Of all novel *Methylococcus* isolates, integron-like elements were detected only in the genome of strain Mc7. This strain contains one complete integron with one attC site, one separate attC site, large attC-cassette (nine sites), and stand-alone integrase. The gene encoding 6-phosphogluconate phosphatase was found in a large attC-cassette; the functions of other genes near attC sites could not be predicted. A complete integron containing three attC sites was also revealed in the genome of *M. geothermalis* IM1[T].

Integrative and Conjugative Elements/ Integrated Mobilizable Elements

Integrative and conjugative elements (ICEs) are widespread mobile DNA that transmit both vertically, in a host-integrated state, and horizontally, through excision and transfer to new recipients. Recent findings indicated that the main actors of conjugative transfer are not the well- known conjugative or mobilizable plasmids but the integrated mobilizable elements (IMEs; Delavat et al., 2017; Guédon et al., 2017). IMEs encode their own excision and integration and use the conjugation machinery of unrelated co-resident conjugative element for their own transfer (Guédon et al., 2017). One putative IME was found in the genome of strain Mc7; unfortunately, the ends of the element could not be identified. A putative ICE region of 123,456 bp was revealed in the genome of strain KN2. This region contained the genes of the type IV secretion system, which plays a key role in conjugation (Lawley et al., 2003). Interestingly, some fragments of a prophage were also found in this ICE region (**Figure 8**).

Phages

One to six prophage regions were found in the genomes of the examined *Methylococcus* strains (**Table 3**). Potentially complete prophages were detected in the genomes of strains IO1 and KN2. In the prophage region of strain IO1, the attL and attR sites were displaced, and the lysozyme genes were also absent (**Figure 7**). This prophage is structurally highly similar to the prophage 2 from *M. capsulatus* Bath. The prophage region in the genome of strain KN2 was, most likely, divided into large fragments as a result of integrase activity. This suggests that the conversion of these prophages to the lytic cycle is unlikely.

Insertion Sequences

Insertion sequences are transposable DNA segments ranging in length from 0.7 to 3.5 kb, generally including a transposase gene encoding the enzyme that catalyzes IS movement. The number of complete IS in the genomes of studied strains varied from 17 in KN2 to 78 in Mc7 (**Figure 9**). Similar to *M. capsulatus* Bath, the genomes of strains IO1 and BH contained IS elements belonging to IS256, IS3, IS5, and ISAs1 families (Siguier et al., 2014). The total number of complete IS elements in these three strains was also similar. By contrast, the genomes of strains IO1 and IM1[T] did not contain IS elements of the ISAs1
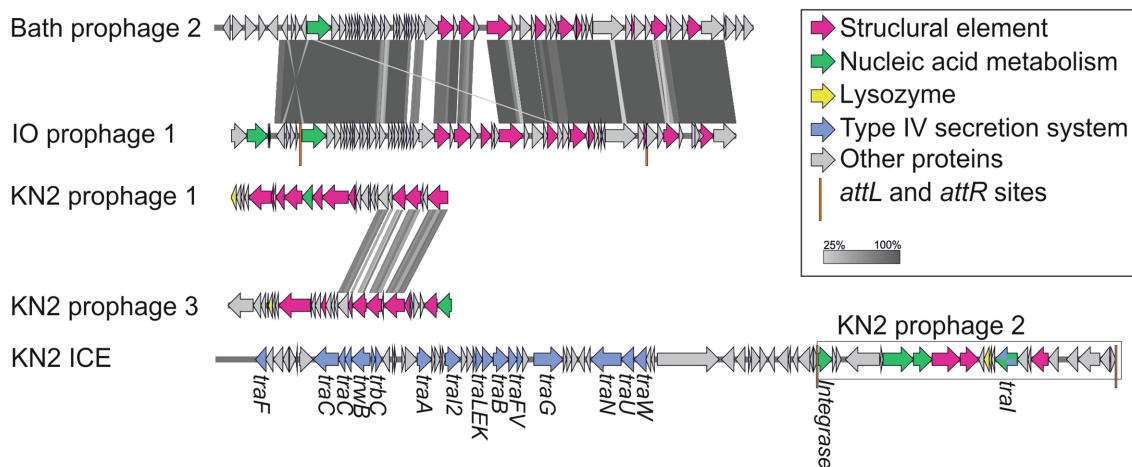
FIGURE 8 | The structure of integrative and conjugative element (ICE) region revealed in the genome of strain KN2 and prophage-associated regions (≥20 Kb) detected in the examined *Methylococcus* genomes.

TABLE 3 | Prophage regions revealed in the examined *Methylococcus* genomes.

Strain	Region no.	Region length, kb	Region position	tRNA	Total proteins	Number of phage protein hits	att_site present
Bath	1	36.1	2,819,640–2,855,795	0	40	20	No
	2	51.7	3,091,049–3,142,838	0	58	44	Yes
IO1	1	48.4	2,508,110–2,556,599	2	64	45	Yes
KN2	1	20.8	1,847,198–1,868,022	0	26	19	No
	2	29	2,258,604–2,287,649	0	18	12	Yes
	3	21.4	2,313,376–2,334,848	0	27	16	No
	4	15.9	3,029,181–3,045,103	0	21	12	No
Mc7	1	13	1,444,044–1,457,136	0	8	6	Yes
	2	7.4	3,652,375–3,659,823	0	7	6	No
IM1^T	1	11.6	672,463–684,108	1	12	7	No
	2	11.3	2,433,262–2,444,593	0	8	7	No
BH	1	11.3	501,048–512,403	0	8	7	No
	2	7.8	650,531–658,368	0	8	6	No
	3	8.5	1,088,481–1,096,982	0	10	6	No
	4	4.1	1,607,407–1,611,579	0	8	7	No
	5	7.7	2,024,931–2,032,721	0	9	6	No
	6	8.1	2,164,268–2,172,463	0	8	6	No

family but contained one IS of the IS110 and IS630 families. Strain IM1^T also contained IS of IS91, ISNCY, IS200_IS605 families, which were not revealed in other studied genomes. Among the studied strains, the Mc7 genome was noticeably distinguished by both the total number and the variety of IS elements. IS massive expansion is accompanied by gene inactivation and decay, genome rearrangement, and genome reduction (Siguier et al., 2014). Some IS elements, for example, IS30 (Dalrymple and Arber, 1985) are capable of reactivating the expression of nearby genes. Therefore, IS elements can play an important role in the adaptation of the host cell to new lifestyle, such as continuous cultivation for the needs of industry (Gaffé et al., 2011). High similarity of the IS family composition between strains Bath, IO1, and BH suggests low influence of a lateral gene transfer on the evolution of their genomes. The change in the IS composition in other strains may be associated with the activity of other mobile elements,

i.e., integrons in the genomes of strains Mc7 and *M. geothermalis* IM1^T, and ICE in the genome of strain KN2.

In summary, this study expanded the pool of phenotypically characterized *Methylococcus* strains with good-quality genome sequences by contributing four novel isolates of these bacteria from various habitats. Three novel isolates, strains KN2, IO1, and BH, were assigned to the species *M. capsulatus* and displayed surprisingly high similarity in gene content to that in *M. capsulatus* Bath. As evidenced by the original study of Foster and Davis (1966) and our work, methanotrophs of this species tend to inhabit nutrient-rich freshwater habitats, such as wastewater, sewage or sediments. They possess relatively small (3.2–3.6 Mb) genomes and display high growth rates on methane in the temperature range between 42 and 52°C. The ability to grow on methanol, however, may vary in different members of this species. Thus, *M. capsulatus* Bath displayed only a trace growth on methanol. The same was
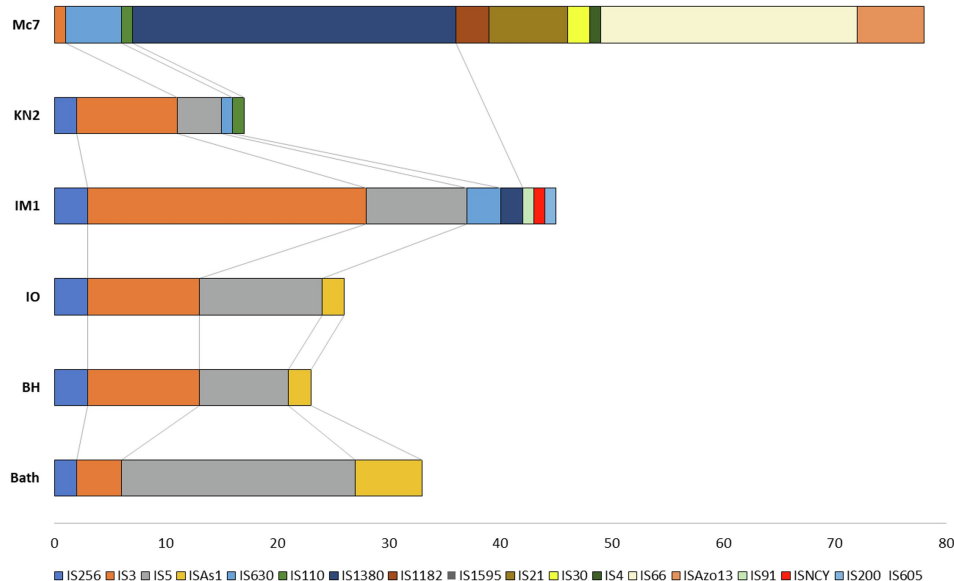
FIGURE 9 | Distribution of insertion sequences (IS) element families in *Methylococcus* genomes.

true for one of the isolates obtained in our study, strain IO1, which showed poor growth only at low methanol concentrations (0.05–0.1%, *v/v*). By contrast, two other strains, KN2 and BH, were capable of growth in a wide range of methanol concentrations, up to 3.0% CH₃OH in case of strain KN2. Notably, high resistance of strain KN2 to methanol correlates with the occurrence of four different isozymes of formate dehydrogenase in this methanotroph. This array of formate dehydrogenases may potentially be responsible for the efficient oxidation of C1 metabolites to CO_2, thus determining high resistance to methanol. However, the exact role of different formate dehydrogenases in metabolism of methanotrophic bacteria remains to be clarified. Strain KN2 was also distinct with regard to its temperature optimum (48–52°C), which was higher than that in other isolates and strain Bath. Finally, strain KN2 showed the best performance during its continuous cultivation in a bioreactor. Taken together, *M. capsulatus* KN2 has the highest potential for the biotechnologies implying high growth rates on methane and/ or abilities to grow on methanol.

The fourth isolate obtained in our study from a landfill cover soil, strain Mc7, may potentially represent a novel species of the genus *Methylococcus*. This is suggested by the results of phylogenomic analysis, comparative genome analysis as well as by the phenotypic difference of strain Mc7 and two previously described *Methylococcus* species. Thus, cells of strain Mc7 (~1.6 μm in diameter) were larger that cells of *M. capsulatus* (1.0–1.1 μm) or *M. geothermalis* (0.7–1.0 μm; Awala et al., 2020). The genome size of strain Mc7 (4.0 Mb) also exceeded that in other described *Methylococcus* species. Notably, the genome of this methanotroph from a landfill cover soil contained high number and variety of mobile elements, which may have played an important role in the adaptation of strain Mc7 to highly

variable conditions of its natural habitat. Thus, strain Mc7 possessed a number of genome-encoded features, which were absent in strains of *M. capsulatus*, such as sucrose biosynthesis and the ability to scavenge phosphorus and sulfur from the environment. These metabolic capabilities may represent important components of the genome-determined environmental adaptations of this methanotroph. Further examination of phenotypic and chemotaxonomic features is needed to establish the taxonomic position of strain Mc7.

Thus, our study contributes to the current knowledge of *Methylococcus*-like methanotrophs, the bacteria of high environmental and biotechnological relevance. As suggested by the pangenome analysis, the metabolic diversity within this genus remains underestimated and calls for further cultivation- and genome-based studies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm. nih.gov/genbank/, CP079095–CP079098.

AUTHOR CONTRIBUTIONS

SD and NP obtained funding and designed the study. OD and IO isolated novel *Methylococcus* strains. IO obtained and annotated the genome sequences. IO, KM, SBu, and VK performed the comparative genome analysis. IO, RS, KM, ET, and SBe conducted physiology and growth tests. ET and NK assisted with experiments on continuous cultivation in bioreactor.

IO, KM, VK, and SD wrote the manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.756830/full#supplementary-material

REFERENCES

Adams, M. A., Luo, Y., Hove-Jensen, B., He, S. M., Van Staalduinen, L. M., Zechel, D. L., et al. (2008). Crystal structure of PhnH: an essential component of carbon-phosphorus lyase in *Escherichia coli*. *J. Bacteriol.* 190, 1072–1083. doi: 10.1128/JB.01274-07

Anthony, C. (1983). *The Biochemistry of Methylotrophs*. London, New York: Academic Press. doi: 10.1016/0968-0004(83)90116-0

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387

Awala, S. I., Bellosillo, L. A., Gwak, J.-H., Nguyen, N.-L., Kim, S.-J., Lee, B.-H., et al. (2020). *Methylococcus geothermalis* sp. nov., a methanotroph isolated from a geothermal field in the Republic of Korea. *Int. J. Syst. Evol. Microbiol.* 70, 5520–5530. doi: 10.1099/ijsem.0.004442

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75

Balasubramanian, R., Smith, S. M., Rawat, S., Yatsunyk, L. A., Stemmler, T. L., and Rosenzweig, A. C. (2010). Oxidation of methane by a biological dicopper Centre. *Nature* 465, 115–119. doi: 10.1038/nature08992

Baxter, N. J., Hirt, R. P., Bodrossy, L., Kovacs, K. L., Embley, M. T., Prosser, J. I., et al. (2002). The ribulose-1,5-bisphosphate carboxylase/oxygenase gene cluster of *Methylococcus capsulatus* (Bath). *Arch. Microbiol.* 177, 279–289. doi: 10.1007/s00203-001-0387-x

Bazinet, A. L. (2017). Pan-genome and phylogeny of *Bacillus cereus* sensu lato. *BMC Evol. Biol.* 17:176. doi: 10.1186/s12862-017-1020-1

Bentley, R., and Chasteen, T. G. (2004). Environmental VOSCs–formation and degradation of dimethyl sulfide, methanethiol and related materials. *Chemosphere* 55, 291–317. doi: 10.1016/j.chemosphere.2003.12.017

Blomqvist, K., Nikkola, M., Lehtovaara, P., Suihko, M. L., Airaksinen, U., Stråby, K. B., et al. (1993). Characterization of the genes of the 2,3-butanediol operons from *Klebsiella terrigena* and *Enterobacter aerogenes*. *J. Bacteriol.* 175, 1392–1404. doi: 10.1128/jb.175.5.1392-1404.1993

Boden, R., Borodina, E., Wood, A. P., Kelly, D. P., Murrell, J. C., and Schäfer, H. (2011). Purification and characterization of dimethylsulfide monooxygenase from *Hyphomicrobium sulfonivorans*. *J. Bacteriol.* 193, 1250–1258. doi: 10.1128/JB.00977-10

Boucher, Y., Labbate, M., Koenig, J. E., and Stokes, H. W. (2007). Integrons: mobilizable platforms that promote genetic diversity in bacteria. *Trends Microbiol.* 15, 301–309. doi: 10.1016/j.tim.2007.05.004

Bowman, J. P. (2015). "Methylococcus," in *Bergey's Manual of Systematics of Archaea and Bacteria*. eds. M. E. Trujillo, S. Dedysh, P. DeVos, B. Hedlund, P. Kampfer, F. A. Rainey et al. (John Wiley & Sons, Inc.).

Bowman, J. P., Sly, L. T., Nichols, P. D., and Hayward, A. C. (1993). Erratum: revised taxonomy of the methanotrophs: description of *Methylobacter* gen. Nov., emendation of *Methylococcus*, validation of *Methylosinus* and *Methylocystis* species, and a proposal that the family *Methylococcaceae* includes only the group I methanotrophs. *Int. J. Syst. Bacteriol.* 43, 735–753. doi: 10.1099/00207713-44-2-375

But, S. Y., Solntseva, N. P., Egorova, S. V., Mustakhimov, I. I., Khmelenina, V. N., Reshetnikov, A., et al. (2018). The genes and enzymes of sucrose metabolism in moderately thermophilic methanotroph *Methylocaldum szegediense* O12. *Extremophiles* 22, 433–445. doi: 10.1007/s00792-018-1006-y

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. bioRxiv [Preprint]. doi: 10.1101/2021.06.03.446934

Carere, C. R., Hards, K., Houghton, K. M., Power, J. F., McDonald, B., Collet, C., et al. (2017). Mixotrophy drives niche expansion of verrucomicrobial methanotrophs. *ISME J.* 11, 2599–2610. doi: 10.1038/ismej.2017.112

Carrión, O., Curson, A. R. J., Kumaresan, D., Fu, Y., Lang, A. S., Mercadé, E., et al. (2015). A novel pathway producing dimethylsulphide in bacteria is widespread in soil environments. *Nat. Commun.* 6:6579. doi: 10.1038/ncomms7579

Chatwood, L. L., Müller, J., Gross, J. D., Wagner, G., and Lippard, S. J. (2004). NMR structure of the flavin domain from soluble methane monooxygenase reductase from *Methylococcus capsulatus* (Bath). *Biochemistry* 43, 11983–11991. doi: 10.1021/bi049066n

Chistoserdova, L. (2011). Modularity of methylotrophy, revisited. *Environ. Microbiol.* 13, 2603–2622. doi: 10.1111/j.1462-2920.2011.02464.x

Colby, J., and Dalton, H. (1978). Resolution of the methane mono-oxygenase of *Methylococcus capsulatus* (Bath) into three components. Purification and properties of component C, a flavoprotein. *Biochem. J.* 171, 461–468. doi: 10.1042/bj1710461

Contreras-Moreira, B., and Vinuesa, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl. Environ. Microbiol.* 79, 7696–7701. doi: 10.1128/AEM.02411-13

Csáki, R., Bodrossy, L., Klem, J., Murrell, J. C., and Kovács, K. L. (2003). Genes involved in the copper-dependent regulation of soluble methane monooxygenase of *Methylococcus capsulatus* (Bath): cloning, sequencing and mutational analysis. *Microbiology* 149, 1785–1795. doi: 10.1099/mic.0.26061-0

Cury, J., Jové, T., Touchon, M., Néron, B., and Rocha, E. P. (2016). Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res.* 44, 4539–4550. doi: 10.1093/nar/gkw319

Dalrymple, B., and Arber, W. (1985). Promotion of RNA transcription on the insertion element IS30 of *E. coli* K12. *EMBO J.* 4, 2687–2693. doi: 10.1002/j.1460-2075.1985.tb03988.x

Dalton, H. (2005). The Leeuwenhoek lecture 2000 The natural and unnatural history of methane-oxidizing bacteria. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1207–1222. doi: 10.1098/rstb.2005.1657

Delavat, F., Miyazaki, R., Carraro, N., Pradervand, N., and van der Meer, J. R. (2017). The hidden life of integrative and conjugative elements. *FEMS Microbiol. Rev.* 41, 512–537. doi: 10.1093/femsre/fux008

DiSpirito, A. A., Kunz, R. C., Choi, D.-W., and Zahn, J. A. (2004). Chapter 7: respiration in methanotrophs. 149–168. doi: 10.1007/978-1-4020-3163-2_7

Dworkin, M., and Foster, J. W. (1958). Experiments with some microorganisms which utilize ethane and hydrogen. *J. Bacteriol.* 75, 592–603. doi: 10.1128/jb.75.5.592-603.1958

Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 2015, 1–29. doi: 10.7717/peerj.1319

Foster, J. W., and Davis, R. H. (1966). A methane-dependent coccus, with notes on classification and nomenclature of obligate, methane-utilizing bacteria. *J. Bacteriol.* 91, 1924–1931. doi: 10.1128/jb.91.5.1924-1931.1966

Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A., and Fischer, J. (2005). Prokaryotic sulfur oxidation. *Curr. Opin. Microbiol.* 8, 253–259. doi: 10.1016/j.mib.2005.04.005

Gaffé, J., McKenzie, C., Maharjan, R. P., Coursange, E., Ferenci, T., and Schneider, D. (2011). Insertion sequence-driven evolution of *Escherichia coli* in chemostats. *J. Mol. Evol.* 72, 398–412. doi: 10.1007/s00239-011-9439-2

Gagliano, A. L., D'Alessandro, W., Tagliavia, M., Parello, F., and Quatrini, P. (2014). Methanotrophic activity and diversity of methanotrophs in volcanic geothermal soils at Pantelleria (Italy). *Biogeosciences* 11, 5865–5875. doi: 10.5194/bg-11-5865-2014

Gebert, J., Singh, B. K., Pan, Y., and Bodrossy, L. (2009). Activity and structure of methanotrophic communities in landfill cover soils. *Environ. Microbiol. Rep.* 1, 414–423. doi: 10.1111/j.1758-2229.2009.00061.x

Gillings, M. R. (2014). Integrons: past, present, and future. *Microbiol. Mol. Biol. Rev.* 78, 257–277. doi: 10.1128/MMBR.00056-13

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0

Guédon, G., Libante, V., Coluzzi, C., Payot, S., and Leblond-Bourget, N. (2017). The obscure world of integrative and mobilizable elements, highly widespread elements that pirate bacterial conjugative systems. *Genes* 8:337. doi: 10.3390/genes8110337

Gupta, A., Ahmad, A., Chothwe, D., Madhu, M. K., Srivastava, S., and Sharma, V. K. (2018). Genome-scale metabolic reconstruction and metabolic versatility of an obligate methanotroph *Methylococcus capsulatus* str. Bath. *PeerJ* 7:e6685. doi: 10.7717/peerj.6685

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Hakobyan, A., Zhu, J., Glatter, T., Paczia, N., and Liesack, W. (2020). Hydrogen utilization by *Methylocystis* sp. strain SC2 expands the known metabolic versatility of type IIa methanotrophs. *Metab. Eng.* 61, 181–196. doi: 10.1016/j.ymben.2020.05.003

Hartmann, T., and Leimkühler, S. (2013). The oxygen-tolerant and NAD+-dependent formate dehydrogenase from *Rhodobacter capsulatus* is able to catalyze the reduction of CO_2 to formate. *FEBS J.* 280, 6083–6096. doi: 10.1111/febs.12528

Hazeu, W., Batenburg-van der Vegte, W. H., and de Bruyn, J. C. (1980). Some characteristics of *Methylococcus mobilis* sp. nov. *Arch. Microbiol.* 124, 211–220. doi: 10.1007/BF00427729

Henard, C. A., Smith, H. K., and Guarnieri, M. T. (2017). Phosphoketolase overexpression increases biomass and lipid yield from methane in an obligate methanotrophic biocatalyst. *Metab. Eng.* 41, 152–158. doi: 10.1016/j.ymben.2017.03.007

Henard, C. A., Wu, C., Xiong, W., Henard, J. M., Davidheiser-Kroll, B., Orata, F. D., et al. (2021). Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) is essential for growth of the methanotroph *Methylococcus capsulatus* strain Bath. *Appl. Environ. Microbiol.* 87:e00881-21. doi: 10.1128/AEM.00881-21

Hibi, Y., Asai, K., Arafuka, H., Hamajima, M., Iwama, T., and Kawai, K. (2011). Molecular structure of La3+-induced methanol dehydrogenase-like protein in *Methylobacterium radiotolerans*. *J. Biosci. Bioeng.* 111, 547–549. doi: 10.1016/j.jbiosc.2010.12.017

Houghton, K. M., and Stewart, L. C. (2020). Temperature-gradient incubation isolates multiple competitive species from a single environmental sample. *Access Microbiol.* 2:acmi000081. doi: 10.1099/acmi.0.000081

Jawaharraj, K., Sudha Dhiman, S., Bedwell, S., Vemuri, B., Islam, J., Sani, R. K., et al. (2021). Electricity from methane by *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* OB3b. *Bioresour. Technol.* 321:124398. doi: 10.1016/j.biortech.2020.124398

Jochimsen, B., Lolle, S., McSorley, F. R., Nabi, M., Stougaard, J., Zechel, D. L., et al. (2011). Five phosphonate operon gene products as components of a multi-subunit complex of the carbon-phosphorus lyase pathway. *Proc. Natl. Acad. Sci. U. S. A.* 108, 11393–11398. doi: 10.1073/pnas.1104922108

Kalyuzhnaya, M. G., Yang, S., Rozova, O. N., Smalley, N. E., Clubb, J., Lamb, A., et al. (2013). Highly efficient methane biocatalysis revealed in a methanotrophic bacterium. *Nat. Commun.* 4:2785. doi: 10.1038/ncomms3785

Kao, W. C., Chen, Y. R., Yi, E. C., Lee, H., Tian, Q., Wu, K. M., et al. (2004). Quantitative proteomic analysis of metabolic regulation by copper ions in *Methylococcus capsulatus* (Bath). *J. Biol. Chem.* 279, 51554–51560. doi: 10.1074/jbc.M408013200

Keltjens, J. T., Pol, A., Reimann, J., and Op Den Camp, H. J. M. (2014). PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Appl. Microbiol. Biotechnol.* 98, 6163–6183. doi: 10.1007/s00253-014-5766-8

Kizilova, A. K., Sukhacheva, M. V., Pimenov, N. V., Yurkov, A. M., and Kravchenko, I. K. (2014). Methane oxidation activity and diversity of aerobic methanotrophs in pH-neutral and semi-neutral thermal springs of the Kunashir Island, Russian Far East. *Extremophiles* 18, 207–218. doi: 10.1007/s00792-013-0603-z

Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes Konstantinos. *Proc. Natl. Acad. Sci. U. S. A.* 102, 2567–2572. doi: 10.1016/S0040-4020(01)97190-X

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Kumar, R., Bröms, J. E., and Sjöstedt, A. (2020). Exploring the diversity within the genus *Francisella* – an integrated pan-genome and genome-mining approach. *Front. Microbiol.* 11:1928. doi: 10.3389/fmicb.2020.01928

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Larsen, Ø., and Karlsen, O. A. (2016). Transcriptomic profiling of *Methylococcus capsulatus* (Bath) during growth with two different methane monooxygenases. *Microbiol. Open* 5, 254–267. doi: 10.1002/mbo3.324

Lawley, T. D., Klimke, W. A., Gubbins, M. J., and Frost, L. S. (2003). F factor conjugation is a true type IV secretion system. *FEMS Microbiol. Lett.* 224, 1–15. doi: 10.1016/S0378-1097(03)00430-0

Lieberman, R. L., and Rosenzweig, A. C. (2005). Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* 434, 177–182. doi: 10.1038/nature03311

Lieven, C., Petersen, L. A. H., Jørgensen, S. B., Gernaey, K. V., Herrgard, M. J., and Sonnenschein, N. (2018). A genome-scale metabolic model for *Methylococcus capsulatus* predicts reduced efficiency uphill electron transfer to pMMO. bioRxiv, 329714 [Preprint]. doi:10.1101/329714

Liu, D., Hunt, M., and Tsai, I. J. (2018a). Inferring synteny between genome assemblies: a systematic evaluation. *BMC Bioinformatics* 19:26. doi: 10.1186/s12859-018-2026-4

Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., et al. (2018b). ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic Acids Res.* 47, D660–D665. doi: 10.1093/nar/gky1123

Livingstone, P. G., Morphew, R. M., and Whitworth, D. E. (2018). Genome sequencing and pan-genome analysis of 23 *Corallococcus* spp. strains reveal unexpected diversity, with particular plasticity of predatory gene sets. *Front. Microbiol.* 9:3187. doi: 10.3389/fmicb.2018.03187

Lomans, B. P., Leijdekkers, P., Wesselink, J. J., Bakkes, P., Pol, A., van der Drift, C., et al. (2001). Obligate sulfide-dependent degradation of methoxylated aromatic compounds and formation of methanethiol and dimethyl sulfide by a freshwater sediment isolate, *Parasporobacterium paucivorans* gen. Nov., sp. nov. *Appl. Environ. Microbiol.* 67, 4017–4023. doi: 10.1128/AEM.67.9.4017-4023.2001

Lomans, B. P., van der Drift, C., Pol, A., and Op den Camp, H. J. M. (2002). Microbial cycling of volatile organic sulfur compounds. *Cell. Mol. Life Sci.* 59, 575–588. doi: 10.1007/s00018-002-8450-6

Lopes, A., Tavares, P., Petit, M. A., Guérois, R., and Zinn-Justin, S. (2014). Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics* 15:1027. doi: 10.1186/1471-2164-15-1027

Malashenko, I. R., Romanovskaia, V. A., Bogachenko, V. N., and Shved, A. D. (1975). Thermophilic and thermotolerant bacteria that assimilate methane. *Mikrobiologiia* 44, 855–862.

Mazel, D. (2006). Integrons: agents of bacterial evolution. *Nat. Rev. Microbiol.* 4, 608–620. doi: 10.1038/nrmicro1462

Minkin, I., Patel, A., Kolmogorov, M., Vyahhi, N., and Pham, S. (2013). "Sibelia: a scalable and comprehensive synteny block generation tool for closely related microbial genomes," in *Algorithms in Bioinformatics*. eds. A. Darling and J. Stoye (Berlin, Heidelberg: Springer Berlin Heidelberg), 215–229.

Mohammadi, S. S., Schmitz, R. A., Pol, A., Berben, T., Jetten, M. S. M., and Op Den Camp, H. J. M. (2019). The acidophilic methanotroph *Methylacidimicrobium tartarophylax* 4ac grows as autotroph on h2 under microoxic conditions. *Front. Microbiol.* 10:2352. doi: 10.3389/fmicb.2019.02352

Müller, J., Lugovskoy, A. A., Wagner, G., and Lippard, S. J. (2002). NMR structure of the [2Fe-2S] ferredoxin domain from soluble methane monooxygenase reductase and interaction with its hydroxylase. *Biochemistry* 41, 42–51. doi: 10.1021/bi015668k

Murrell, J. C. (2010). "Genomics of *Methylococcus capsulatus*," in *Handbook of Hydrocarbon and Lipid Microbiology*. ed. K. N. Timmis (Berlin, Heidelberg: Springer).

Oshkin, I. Y., Miroshnikov, K. K., Grouzdev, D. S., and Dedysh, S. N. (2020). Pan-genome-based analysis as a framework for demarcating two closely related methanotroph genera *Methylocystis* and *Methylosinus*. *Microorganisms* 8:768. doi: 10.3390/microorganisms8050768

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36:996. doi: 10.1038/nbt.4229

Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., and Toth, I. K. (2016). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods* 8, 12–24. doi: 10.1039/C5AY02550H

R Core Team (2020). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.r-project.org/

Reshetnikov, A. S., Rozova, O. N., Khmelenina, V. N., Mustakhimov, I. I., Beschastny, A. P., Murrell, J. C., et al. (2008). Characterization of the pyrophosphate-dependent 6-phosphofructokinase from *Methylococcus capsulatus* Bath. *FEMS Microbiol. Lett.* 288, 202–210. doi: 10.1111/j.1574-6968.2008.01366.x

Romanovskaya, V. A., Malashenko, Y. R., and GrishChenko, N. I. (1981). Diagnosis of methane-oxidizing bacteria by numerical methods based on cell fatty acid composition. *Mikrobiologiya* 49, 762–764.

Romarheim, O. H., Øverland, M., Mydland, L. T., Skrede, A., and Landsverk, T. (2011). Bacteria grown on natural gas prevent soybean meal-induced enteritis in Atlantic salmon1-3. *J. Nutr.* 141, 124–130. doi: 10.3945/jn.110.128900

Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993). Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane. *Nature* 366, 537–543. doi: 10.1038/366537a0

Ross, M. O., MacMillan, F., Wang, J., Nisthal, A., Lawton, T. J., Olafson, B. D., et al. (2019). Particulate methane monooxygenase contains only mononuclear copper centers. *Science* 364, 566–570. doi: 10.1126/science.aav2572

Rozova, O. N., Khmelenina, V. N., Gavletdinova, J. Z., Mustakhimov, I. I., and Trotsenko, Y. A. (2015). Acetate kinase-an enzyme of the postulated phosphoketolase pathway in *Methylomicrobium alcaliphilum* 20Z. *Antonie Van Leeuwenhoek* 108, 965–974. doi: 10.1007/s10482-015-0549-5

Rozova, O. N., Mustakhimov, I. I., But, S. Y., Reshetnikov, A. S., and Khmelenina, V. N. (2019). Role of the malic enzyme in metabolism of the halotolerant methanotroph *Methylotuvimicrobium alcaliphilum* 20Z. *PLoS One* 14:e0225054. doi: 10.1371/journal.pone.0225054

Sato, Y., Monincová, M., Chaloupková, R., Prokop, Z., Ohtsubo, Y., Minamisawa, K., et al. (2005). Two rhizobial strains, *Mesorhizobium loti* MAFF303099 and *Bradyrhizobium japonicum* USDA110, encode haloalkane dehalogenases with novel structures and substrate specificities. *Appl. Environ. Microbiol.* 71, 4372–4379. doi: 10.1128/AEM.71.8.4372-4379.2005

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Seto, H., and Kuzuyama, T. (1999). Bioactive natural products with carbon-phosphorus bonds and their biosynthesis. *Nat. Prod. Rep.* 16, 589–596. doi: 10.1039/a809398i

Siguier, P., Gourbeyre, E., and Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* 38, 865–891. doi: 10.1111/1574-6976.12067

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference Centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi: 10.1093/nar/gkj014

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351

Skrede, A., Berge, G. M., Storebakken, T., Herstad, O., Aarstad, K. G., and Sundstøl, F. (1998). Digestibility of bacterial protein grown on natural gas in mink, pigs, chicken and Atlantic salmon. *Anim. Feed Sci. Technol.* 76, 103–116. doi: 10.1016/S0377-8401(98)00208-9

Skrede, A., Mydland, L. T., and Øverland, M. (2009). Effects of growth substrate and partial removal of nucleic acids in the production of bacterial protein meal on amino acid profile and digestibility in mink. *J. Anim. Feed Sci.* 18, 689–698. doi: 10.22358/jafs/66442/2009

Stanley, S. H., Prior, S. D., Leak, D. J., and Dalton, H. (1983). Copper stress underlies the fundamental change in intracellular location of methane mono-oxygenase in methane-oxidizing organisms: studies in batch and continuous cultures. *Biotechnol. Lett.* 5, 487–492. doi: 10.1007/BF00132233

Taylor, S. C., Dalton, H., and Dow, C. S. (1981). Ribulose-1,5-bisphosphate carboxylase/oxygenase and carbon assimilation in *Methylococcus capsulatus* (Bath). *Microbiology* 122, 89–94. doi: 10.1099/00221287-122-1-89

Ternan, N. G., McGrath, J. W., and Quinn, J. P. (1998). Phosphoenolpyruvate phosphomutase activity in an L-phosphonoalanine- mineralizing strain of *Burkholderia cepacia*. *Appl. Environ. Microbiol.* 64, 2291–2294. doi: 10.1128/AEM.64.6.2291-2294.1998

Varani, A. M., Siguier, P., Gourbeyre, E., Charneau, V., and Chandler, M. (2011). ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol.* 12:R30. doi: 10.1186/gb-2011-12-3-r30

Vuilleumier, S., Khmelenina, V. N., Bringel, F., Reshetnikov, A. S., Lajus, A., Mangenot, S., et al. (2012). Genome sequence of the haloalkaliphilic methanotrophic bacterium *Methylomicrobium alcaliphilum* 20Z. *J. Bacteriol.* 194, 551–552. doi: 10.1128/JB.06392-11

Ward, N., Larsen, Ø., Sakwa, J., Bruseth, L., Khouri, H., Durkin, A. S., et al. (2004). Genomic insights into methanotrophy: the complete genome sequence of *Methylococcus capsulatus* (Bath). *PLoS Biol.* 2:e303. doi: 10.1371/journal.pbio.0020303

Whittenbury, R. (1981). "The interrelationship of autotrophy and methylotrophy as seen in *Methylococcus capsulatus* (Bath)," in *Microbiol Growth on C1 Compounds*. ed. H. Dalton (London, UK: Heyden & Son Ltd.), 181–190.

Whittenbury, R., and Dalton, H. (1981). "The methylotrophic bacteria," in *The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria*. eds. M. P. Starr, H. Stolp, H. G. Trüper, A. Balows and H. G. Schlegel (Berlin, Heidelberg: Springer Berlin Heidelberg), 894–902.

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595

Wilson, K. (2001). Preparation of genomic DNA from bacteria. *Curr. Protoc. Mol. Biol.* 56, 2–4. doi: 10.1016/B978-0-12-418687-3.00011-2

Xiao, Z., and Xu, P. (2007). Acetoin metabolism in bacteria. *Crit. Rev. Microbiol.* 33, 127–140. doi: 10.1080/10408410701364604

Xie, Z., and Tang, H. (2017). ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics* 33, 3340–3347. doi: 10.1093/bioinformatics/btx433

Xu, X., Shi, H., Gong, X., Chen, P., Gao, Y., Zhang, X., et al. (2020). Structural insights into sodium transport by the oxaloacetate decarboxylase sodium pump. *eLife* 9:e53853. doi: 10.7554/eLife.53853

Yu, Z., Beck, D. A. C., and Chistoserdova, L. (2017). Natural selection in synthetic communities highlights the roles of *Methylococcaceae* and *Methylophilaceae* and suggests differential roles for alternative methanol dehydrogenases in methane consumption. *Front. Microbiol.* 8:2392. doi: 10.3389/fmicb.2017.02392

Zahn, J. A., and DiSpirito, A. A. (1996). Membrane-associated methane monooxygenase from *Methylococcus capsulatus* (Bath). *J. Bacteriol.* 178, 1018–1029. doi: 10.1128/jb.178.4.1018-1029.1996

Zheng, Y., Huang, J., Zhao, F., Chistoserdova, L., and Giovannoni, S. (2018). Physiological effect of XoxG(4) on lanthanide-dependent methanotrophy. *mBio* 9:e02430-17. doi: 10.1128/mBio.02430-17

Check for updates

A Nanoparticle-Based Biosensor Combined With Multiple Cross Displacement Amplification for the Rapid and Visual Diagnosis of *Neisseria gonorrhoeae* in Clinical Application

Xu Chen[1,2,3*†], Liming Huang[1†], Qingxue Zhou[4], Yan Tan[5], Xuhong Tan[1] and Shilei Dong[6*]

[1] The Second Clinical College, Guizhou University of Traditional Chinese Medicine, Guiyang, China, [2] Clinical Medical Laboratory of the Second Affiliated Hospital, Guizhou University of Traditional Chinese Medicine, Guiyang, China, [3] Guizhou Provincial Center for Disease Control and Prevention, Guiyang, China, [4] Clinical Laboratory, Hangzhou Women's Hospital, Hangzhou, China, [5] Guizhou Provincial Center for Clinical Laboratory, Guiyang, China, [6] Department of Clinical Laboratory, Zhejiang Hospital, Hangzhou, China

Gonorrhea is a sexually transmitted disease caused by the host-adapted human pathogen, *Neisseria gonorrhoeae*. The morbidity is increasing and poses a major public health concern, especially in resource-scarce regions. Therefore, a rapid, visual, sensitive, specific, cost-saving, and simple assay for *N. gonorrhoeae* detection is critical for prompt treatment and the prevention of further transmission. Here, for the first time, we report a novel assay called the multiple cross displacement amplification combined with gold nanoparticle-based lateral flow biosensor (MCDA-LFB), which we constructed for the rapid and visual identification of *N. gonorrhoeae* in clinical samples. We successfully devised a set of MCDA primers based on the *N. gonorrhoeae*-specific gene, *orf1*. Optimal assay conditions were determined at 67°C, including genomic DNA preparation (~15 min), MCDA amplification (30 min), and LFB reading (~2 min), which can be completed within 50 min. The limit of detection (LoD) of the assay was 20 copies/test (in a 25-µl reaction mixture). Assay specificity was 100%, with no cross-reactions with other pathogens. Thus, our *N. gonorrhoeae*-MCDA-LFB is a rapid, specific, visual, cost-saving, and easy-to-use assay for *N. gonorrhoeae* diagnostics, and may have great potential for point-of-care (POC) testing in clinical settings, especially in resource-limited regions.

Keywords: gonorrhea, *Neisseria gonorrhoeae*, multiple cross displacement amplification, gold nanoparticle-based lateral flow biosensor, limit of detection, point-of-care testing

Abbreviations: WHO, World Health Organization; MCDA, multiple cross displacement amplification; LFB, gold nanoparticle-based lateral flow biosensor; PCR, polymerase chain reaction; LoD, limit of detection; ATCC, American type culture collection; MG, malachite green; 2nd GZUTCM, Second Affiliated Hospital, Guizhou University of Traditional Chinese Medicine; GZCCL, Guizhou Provincial Center for Clinical Laboratory; FAM, carboxyfluorescein; nt, nucleotide; mer, monomeric unit; TL, test line; CL, control line; BC, blank control; NC, negative control; DW, distilled water.

INTRODUCTION

Gonorrhea is a sexually transmitted disease caused by the host-adapted human pathogen, *Neisseria gonorrhoeae*. According to the World Health Organization (WHO), approximately 87 million new cases occur globally, of these, more than 90% occur in underdeveloped regions (Rowley et al., 2019). With an increasing global morbidity, gonorrhea poses a major public health concern (Hook and Bernstein, 2019; Quillin and Seifert, 2018). *N. gonorrhoeae* easily infects the mucosa of exposed anatomic regions, including the urogenital tract, pharynx, conjunctivae, and rectum (Chan et al., 2016; Kirkcaldy et al., 2019), and causes serious disease if not treated in a timely and appropriate manner. In females, infection leads to pelvic inflammatory disease, ectopic pregnancy, chronic pelvic pain, and tubal factor infertility. Also, infections during pregnancy may lead to the premature rupture of membranes, spontaneous abortion, preterm birth, and low birth weight (Stevens and Criss, 2018; Hook and Bernstein, 2019; Młynarczyk-Bonikowska et al., 2020). Maternal transmission to infants during birth may also lead to hyper-acute conjunctivitis, corneal perforation, and blindness (Stevens and Criss, 2018). Gonococcus infection in males can cause prostatitis and epididymo-orchitis (Abraha et al., 2018; Jefferson et al., 2021). Gonorrhea is also related to an increased risk of human immunodeficiency virus (HIV) infection and transmission (Jefferson et al., 2021). Hence, the development of an assay for *N. gonorrhoeae* diagnosis is critical for prompt treatment and the prevention of further transmission.

Traditional laboratory-based gold standard gonococcus diagnostic methods were based on cultivation. However, sensitivity was low and was largely due to poor sample collection, transport, and storage (Meyer and Buder, 2020). In addition, cultivation was time consuming and labor intensive. Nowadays, nucleic acid amplification tests (NAATs), including polymerase chain reaction (PCR) and real-time-PCR, are considered new diagnostic gold standards for *N. gonorrhoeae* detection, owing to improved sensitivity, rapidity, and automation (Golparian et al., 2015; Perera et al., 2017). Nevertheless, NAATs are often inaccessible and unaffordable in resource-constrained settings due to expensive thermo-cycling instrumentation and the requirement for highly trained personnel.

Multiple cross displacement amplification (MCDA) is an innovative nucleic acid isothermal amplification technique proposed as an attractive alternative to traditional PCR-related techniques (Wang et al., 2017). The strategy has potential applications as a point-of-care (POC) assay due to its rapidity, simplicity, and easy-to-use operations. Critically, it has been used to detect several pathogens such as SARS-CoV-2, hepatitis B virus, and *Streptococcus agalactiae* (Cheng et al., 2020; Li et al., 2020; Chen et al., 2021). The gold nanoparticle-based lateral flow biosensor (LFB) is a paper-based platform, which facilitates a low-cost rapid diagnosis; it is robust, visual, sensitive, specific, and has a low limit of detection (LoD) (Zeng et al., 2013; Urusov et al., 2019). Importantly, the technique can be used in clinical settings to detect specific biomarkers, including antibodies, antigens, and DNA; therefore it is considered a new and easy-to-use POC device.

In our study, MCDA combined with gold nanoparticle-based LFB (MCDA-LFB) was constructed to detect the *N. gonorrhoeae orf1* gene (Chaudhry and Saluja, 2002; Edwards et al., 2014), which appeared to be uniquely present in *N. gonorrhoeae* as it showed no homology with other microbial genomes at GenBank by BLAST searches The assay principle and study workflow is outlined in **Figures 1**, **2**, and we validated its feasibility using patient clinical samples. The assay was completed with high accuracy within 50 min. Therefore, the MCDA-LFB assay can be considered a valuable POC testing device for diagnosing gonorrhea, even in financially impoverished settings.

MATERIALS AND METHODS

Reagents

Nucleic acid releasing agents were obtained from Sansure Biotech Inc. (Changsha, China). Thayer–Martin (TM) chocolate agar plates were purchased from Autobio Biotechnology Co., Ltd. (Zhengzhou, China). The colorimetric indicator, malachite green (MG) and universal isothermal amplification kits were obtained from HuiDeXin Bio-Technique (Tianjin, China). Biotinylated bovine serum albumin (Biotin-BSA) and rabbit anti-fluorescein antibody (anti-FAM) were purchased from Abcam Co., Ltd. (Shanghai, China). Streptavidin dye-coated gold nanoparticles (Crimson red) were obtained from Bangs Laboratories, Inc. (IN, United States). Gold nanoparticle-based LFB materials, including sample pad, conjugate pad, absorbent pad, nitrocellulose membranes, and backing cards were manufactured by HuiDeXing Biotech Co., Ltd. (Tianjing, China) according to our design specifications. Gonococcus commercial PCR diagnosis kits were obtained from DaAn Gene Co., Ltd. (Guangzhou, China).

Clinical Sample Preparation

We collected 116 genital secretion samples from suspected *N. gonorrhoeae*-infected patients at the Hangzhou Women's Hospital between August 2020 and April 2021. Two samples were collected from patients using sterile swabs. One sample was immediately inoculated on Thayer-Martin (TM) chocolate agar plates at 37°C, 80% humidity, pH 6.8–7.5, and in a 5% CO2-enriched atmosphere for 2–3 d. The remaining sample was for nucleic acid extraction. *N. gonorrhoeae* genomic DNA was obtained using nucleic acid releasing agents in accordance with manufacturer's instructions. In brief, the collected genital secretion samples (500 μl) were centrifuged at 12000 rpm at 4°C for 5 min. The pellet was suspended in 50 μl nucleic acid releasing agents, and incubated at room temperature (25°C) for 10 min. Genomic DNA concentrations were measured using a Nano Drop ND-2000 (Thermo, United States) at A260/280 nm. The corresponding genome copy number was calculated from the weight of the *N. gonorrhoeae* genome, One *N. gonorrhoeae* genome is 2.45 fg (2.2×10^6 bp (Dempsey et al., 1991) $\times$ 665 Da/bp $\times 1.67 \times 10^{-24}$ g/Da) (Edwards et al., 2014).

FIGURE 1 | *Neisseria gonorrhoeae*-multiple cross displacement amplification-lateral flow biosensor (MCDA-LFB) workflow. *N. gonorrhoeae*-MCDA-LFB assay uses three closely linked steps: DNA extraction (step 1), multiple cross displacement amplification (MCDA) (step 2), and gold nanoparticle-based lateral flow biosensor (LFB) readout (step 3). The detection process takes approximately 50 min.

FIGURE 2 | Schematic of the gold nanoparticle-based LFB for *N. gonorrhoeae*-MCDA product visualization. **(A)** MCDA products (0.5 μl) and running buffer (100 μl) were deposited on the sample pad. **(B)** The running buffer containing MCDA products moves along the LFB due to capillary action, while streptavidin dye-coated gold nanoparticles (streptavidin-DPNs) are rehydrated in the conjugate region. **(C)** For positive samples, FAM/biotin-labeled *orf1*-MCDA amplicons are captured by anti-carboxyfluorescence (FAM) at the test line (TL), and streptavidin-DPNs are captured by biotin-BSA at the control line (CL). For negative outcomes, only streptavidin-DPNs are captured by biotinylated bovine serum albumin (biotin-BSA) at the CL. **(D)** Interpretation of *N. gonorrhoeae*-MCDA-LFB assay results; positive – CL and TL appear on the LFB, and negative – only the CL appears on the LFB.

Gold Nanoparticle-Based Lateral Flow Biosensor Preparations

The gold nanoparticle-based LFB design is shown in **Figure 2**. Briefly, the LFB was composed of the following sections: sample pad, conjugate pad, nitrocellulose membrane with immobilized anti-FAM and biotin-BSA, and an absorbent pad.

Dye streptavidin-coated gold nanoparticles were deposited on the conjugate pad. Biotin-BSA and anti-FAM were dispensed onto the nitrocellulose membrane to act as the control line (CL) and test line (TL) (*N. gonorrhoeae*), respectively. Each band was separated by 5 mm. The biosensor components were affixed to a backing card.

Primer name	Sequence and modifications	Length	Gene
F1	5′-ACGTCCACCAATCCATTTGG-3′	20 nt	*orf1*
F2	5′-GATGGAAGCGGAACGGTT-3′	18 nt	
CP1	5′-TGTAGTAGAGCGCGGTATCGGAACGGTCAAAACCTGTTCGC-3′	41 mer	
CP2	5′-ACCAACTCCTACAAACGCCTCGTTGGCGGAATAGGCCAATT-3′	41 mer	
C1	5′-TGTAGTAGAGCGCGGTATCGGA-3′	22 nt	
C1*	5′-Biotin-TGTAGTAGAGCGCGGTATCGGA-3′	22 nt	
C2	5′-ACCAACTCCTACAAACGCCTCG-3′	22 nt	
D1	5′-AAACCGGCATAGCCGTCG-3′	18 nt	
D1*	5′-FAM-AAACCGGCATAGCCGTCG-3′	18 nt	
D2	5′-ACTTTGAAGCACCGACC-3′	17 nt	
R1	5′-GCTTTGGCGTGTTTGAT-3′	17 nt	
R2	5′-TGAACGCGATTACCAAT-3′	17 nt	

C1*, 5′-labeled with biotin when used for MCDA-LFB detection; D1*, 5′-labeled with FAM when used for MCDA-LFB detection. FAM, 6-carboxy-fluorescein; nt, nucleotide; mer, monomeric unit.

No.	Pathogen	Source of pathogens[a]	No. of strains	*N. gonorrhoeae*-MCDA result[b]
1	*N. gonorrhoeae* (reference strain)	ATCC 49926	1	P
2	*N. gonorrhoeae* (clinical samples)	Hangzhou Women's Hospital	8	P
3	*Neisseria meningitides*	Hangzhou Women's Hospital	1	N
4	*Chlamydia trachomatis*	Hangzhou Women's Hospital	1	N
5	Enteropathogenic *Escherichia coli*	GZCCL	1	N
6	*Haemophilus influenzae*	ATCC49247	1	N
7	*Pseudomonas aeruginosa*	2nd GZUTCM	1	N
8	*Candida glabrata*	2nd GZUTCM	1	N
9	*Staphylococcus aureus*	2nd GZUTCM	1	N
10	*Acinetobacter baumannii*	2nd GZUTCM	1	N
11	*Streptococcus pyogenes*	2nd GZUTCM	1	N
12	*Bordetella pertussis*	GZCCL	1	N
13	*Mycobacterium tuberculosis*	GZCDC	1	N
14	*Cryptococcus neoformans*	ATCC13690	1	N
15	*Leptospira interrogans*	GZCDC	1	N
16	*Klebsiella pneumoniae*	GZCCL	1	N
17	*Shigella flexneri*	Hangzhou Women's Hospital	1	N
18	*Mycoplasma pneumonia*	Hangzhou Women's Hospital	1	N
19	*Haemophilus parainfluenzae*	GZCCL	1	N
20	Human rhinovirus	Hangzhou Women's Hospital	1	N
21	Respiratory syncytial virus type A	Hangzhou Women's Hospital	1	N
22	Adenoviruses	Hangzhou Women's Hospital	1	N

[a]ATCC, American type culture collection; 2nd GZUTCM, the Second Affiliated Hospital, Guizhou University of Traditional Chinese Medicine; GZCCL, Guizhou Provincial Center for Clinical Laboratory; GZCDC, Guizhou Provincial Center for Disease Control and Prevention. [b]P, positive; N, negative.

Neisseria gonorrhoeae-Multiple Cross Displacement Amplification Primer Design

Based on MCDA reaction mechanisms, five pairs of primers based on *orf1* (GenBank Accession No. M84113) were designed using primer Explorer V5 and PRIMER PREMIER 5.0 software (Eiken Chemical, Japan). A set of *N. gonorrhoeae*-MCDA primers, including displacement primers (F1 and F2), cross primers (CP1 and CP2), and amplification primers

(D1, D2, C1, C2, R1, and R2), was also generated (**Table 1**). *N. gonorrhoeae*-MCDA primer specificity was verified using the BLAST analysis tool. In addition, Oligo Analyzer online software (V3.1; Integrated DNA Technologies, Coralville, IA, United States) was used for primer secondary structure and dimer investigations. MCDA primer sequences are shown (**Table 1**). All primers were synthesized and purified at TsingKe Biotech Co., Ltd. (Beijing, China) using high performance liquid chromatography purification grade.

FIGURE 3 | Determination and confirmation of *N. gonorrhoeae*-MCDA products. *N. gonorrhoeae*-MCDA products were monitored using a colorimetric indicator **(A)** and a gold nanoparticle-based biosensor **(B)**. Tube 1/Biosensor 1: positive result for the *N. gonorrhoeae* reference strain (ATCC 49926); Tube 2/Biosensor 2: negative result for *Neisseria meningitides*; Tube 3/Biosensor 3: negative result for *Chlamydia trachomatis*; Tube 4/Biosensor 4: blank control (distilled water). TL, test line; CL, control line.

Multiple Cross Displacement Amplification and Detection

Multiple cross displacement amplification was performed using an isothermal amplification kit (HuiDeXing Biotech Co., Ltd., Tianjing, China). Briefly, a one-step 25-µl reaction mixture contained 12.5-µl 2 × reaction buffer [40 mM Tris–HCl (pH 8.8), 40 mM KCl, 16 mM $MgSO_4$, 20 mM $(NH_4)_2SO_4$, 2 M betaine, and 0.2% Tween-20]; 1 µl (8 U) *Bst* 2.0 DNA polymerase (New England Biolabs, United States); 1 µl (10 U) AMV Reverse Transcriptase (only used for RNA templates); 0.4 µM F1 and F2 primers; 1.6 µM CP1 and CP2 primers; 1.2 µM of D1*, D2, C1*, C2, R1, and R2 primers; 2-µl nucleic acid template; and doubly distilled water (DW) to 25 µl. The reaction was performed in heat-block at constant temperature (optimization outlined later).

Monitoring techniques, including the colorimetric indicator, MG, real-time turbidity (LA-500), and an LFB were used to analyze MCDA products. For colorimetric analysis, colorless reactions changed to light green suggesting a positive amplification. A colorless reaction was observed in negative and blank controls (BCs). For the real-time turbidity method, turbidity >0.1 was regarded as a positive result. For LFB detection, the CL and TL reacted simultaneously, indicating positive results. For negative and blank outcomes, only the CL was observed on the biosensor.

Optimization of *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplication-Lateral Flow Biosensor Reaction Conditions

Temperature is critical for isothermal amplification. The amplification temperature was optimized, ranging from 63 to 70°C (with 1°C intervals) using *N. gonorrhoeae* genomic DNA (2.0×10^3 copies/assay). MCDA amplicons were monitored using real-time turbidity (LA-500). Then, amplification times (20, 30, 40, and 50 min) of the assay were optimized. The *N. gonorrhoeae*-MCDA reaction was carried out with optimal amplification temperature, and the results were readout simultaneously with MG and LFB. Each test was conducted at least three times.

Sensitivity and Specificity of the *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Assay

To test assay sensitivity, reference *N. gonorrhoeae* [American type culture collection (ATCC); 49926] genomic DNA was 10-fold serial diluted from 2.0×10^4 to 2.0×10^{-1} copies. Then, *N. gonorrhoeae*-MCDA reactions were conducted under optimal conditions, and the results were analyzed using MG and LFB. The LoD of *N. gonorrhoeae*-MCDA was confirmed as the lowest dilution for which all three replicates were positive. Assay specificity was determined using nucleic acid (at least 2.0×10^4/reaction) from various bacteria, viruses, and fungi (**Table 2**). DW was used as a BC. MCDA results were monitored using an LFB and each assay was performed in triplicate.

Verification of *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Feasibility Using Clinical Samples

To confirm assay feasibility, the optimized detection system was verified using clinical samples. We collected 116 suspected *N. gonorrhoeae*-infection genital secretion samples from Hangzhou Women's Hospital (Hangzhou, China). Samples were simultaneously assayed using cultivation, quantitative PCR (qPCR), and MCDA-LFB methods.

Culturing was conducted as above described. qPCR detection was performed using the Gonococcus nucleic acid assay kit (DaAn Gene Co., Ltd., China) (Cat. #DA-D053) and detection was performed using the Applied Biosystems™ 7500 Real-Time PCR System (Life Technologies, Singapore). According to the instruction of the manufacturer, *N. gonorrhoeae* concentrations <500 copies were regarded as negative results The *N. gonorrhoeae*-MCDA-LFB operation was as described above. All assays were conducted at biosafety level 2 according to the WHO Laboratory Biosafety Manual, 3rd edition.
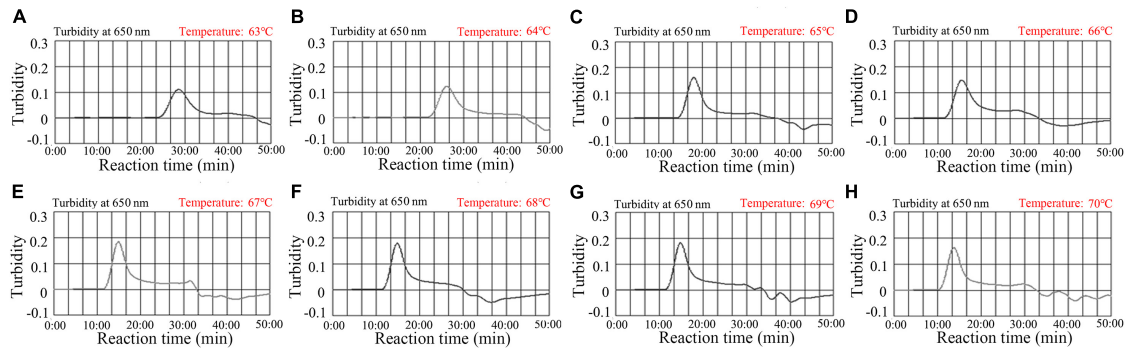
FIGURE 4 | Temperature optimization for *N. gonorrhoeae*-MCDA reactions. *N. gonorrhoeae*-MCDA reactions were monitored by real-time turbidity measurements (LA-500). Turbidity >0.1 was considered positive, and the threshold value was 0.1. Eight kinetic graphs **(A–H)** were derived at various reaction temperatures; 63–70°C at 1°C intervals, with *N. gonorrhoeae* reference strain (ATCC 49926) target DNA at 2×10^3 copies/reaction. The optimal MCDA reaction temperature was selected based on higher turbidity. The temperature, 67°C **(E)** showed robust amplification.
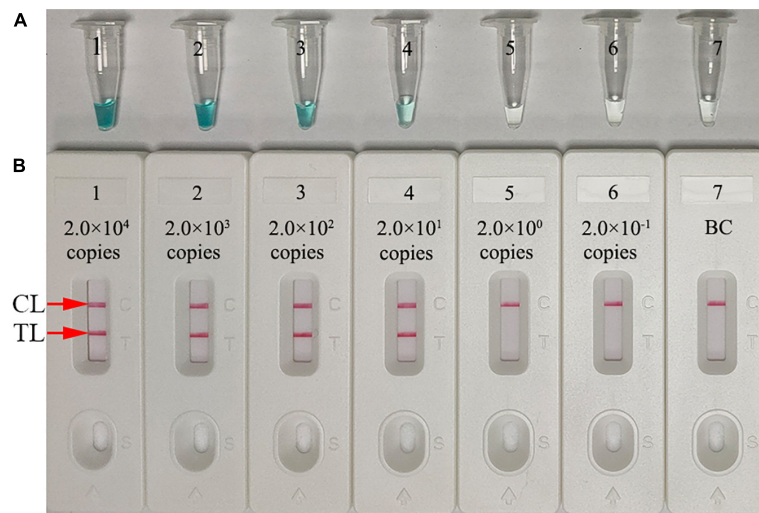
FIGURE 5 | Assay sensitivity using diluted genomic DNA templates. Malachite green (MG) **(A)** and gold nanoparticle-based LFB **(B)** techniques were simultaneously used for reporting *N. gonorrhoeae*-MCDA results. Tubes A1–A7 (Biosensors B1–B7) contained 2.0×10^4, 2.0×10^3, 2.0×10^2, 2.0×10^1, 2.0×10^0, and 2.0×10^{-1} DNA copies/reaction and blank control (distilled water). The limit of detection (LoD) of the *N. gonorrhoeae*-MCDA-LFB assay for *orf1* was 2.0×10^1 DNA copies/reaction. CL, control line; TL, test line.

Statistical Analysis

The χ^2 test was used to compare sensitivity differences between cultivation, qPCR, and MCDA-LFB assays. The SPSS 23.0 software was used for statistical analyses, and $p < 0.05$ was considered statistically significant.

RESULTS

Overview of the *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Detection System

The *N. gonorrhoeae*-MCDA-LFB system and workflow is shown in **Figures 1**, **2**. Briefly, released *N. gonorrhoeae* genomic DNA

was pre-amplified by MCDA at a constant temperature of 67°C for 30 min. We modified the C1 and D1 MCDA primers at the 5′-end with biotin and FAM for LFB detection. Amplicons were then labeled with biotin and FAM (**Figures 1**, **2**). For positive samples, FAM/biotin-labeled *orf1*-MCDA amplicons were captured by anti-FAM at the TL, and streptavidin-DPNs were captured by biotin-BSA at the CL. For a negative outcome, only streptavidin-DPNs were captured by biotin-BSA at the CL (**Figure 2**).

Demonstration of the *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification Assay

To validate *N. gonorrhoeae*-MCDA primers, MCDA reaction mixtures were incubated on a heat-block at a constant temperature of 65°C for 1 h using nucleic acid from purified

FIGURE 6 | Optimization of assay reaction times. Four different amplification times **(A)**, 20 min; **(B)**, 30 min; **(C)**, 40 min; and **(D)**, 50 min were tested at 67°C. Tube/Biosensors 1, 2, 3, 4, 5, 6, and 7 represented *N. gonorrhoeae* (ATCC 49926) genomic DNA at 2.0 × 10^4 to 2.0 × 10^{-1} copies/reaction and blank control (distilled water). Optimal sensitivity was indicated when the amplification lasted 30 min **(B)**. CL, control line; T, test line.

N. gonorrhoeae (ATCC 49926) cultures. MCDA amplicons were then analyzed using MG and LFB techniques. Our data showed that the color of the *N. gonorrhoeae*-MCDA tube went from colorlessness to bright green, while negative and BCs remained colorless (**Figure 3A**). Using LFB detection, TL and CL appeared simultaneously in the *N. gonorrhoeae*-MCDA reaction, while in negative and BCs, only the CL appeared (**Figure 3B**). These data suggested that the *N. gonorrhoeae*-MCDA primers for *orf1* detection were valid for ongoing assay development.

Determining the Optimal Amplification Temperature of the *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Assay

Reaction temperature is important for isothermal amplification. For the MCDA pre-amplification stage, temperatures from 63–70°C with 2.0 × 10^3 copies/reaction of *N. gonorrhoeae* DNA template were investigated. Using real-time turbidity (LA-500) (**Figure 4**), the robust amplification of *N. gonorrhoeae*-MCDA

was observed at 67–69°C. Therefore, 67°C was used as the optimal reaction temperature for subsequent studies.

Sensitivity of the *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Assay

Assay sensitivity was tested using serial dilutions of *N. gonorrhoeae* DNA (from 2.0 × 10^4 to 2.0 × 10^{-1} copies/test). The assay was conducted as described, and the results were analyzed using MG and LFB. The assay LoD was 20 copies/test (in a 25-μl reaction mixture) (**Figures 5A,B**).

Optimal Amplification Times for the *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Assay

Amplification assay times of 20, 30, 40, and 50 min were investigated and optimized at 67°C. MCDA products were analyzed using MG and LFB. As shown in **Figure 6**, the LoD

FIGURE 7 | Assay specificity using different pathogens. The
N. gonorrhoeae-MCDA-LFB assay was verified using different genomic
RNA/DNA as templates. Biosensor 1, *N. gonorrhoeae* (ATCC 49926);
Biosensors 2–9, *N. gonorrhoeae* (clinical samples); Biosensor 10, *Neisseria
meningitides*; Biosensor 11, *Chlamydia trachomatis*; Biosensor 12,
Enteropathogenic *Escherichia coli*; Biosensor 13, *Haemophilus influenzae*;
Biosensor 14, *Pseudomonas aeruginosa*; Biosensor 15, *Candida glabrata*;
Biosensor 16, *Staphylococcus aureus*; Biosensor 17, *Acinetobacter
baumannii*; Biosensor 18, *Streptococcus pyogenes*; Biosensor 19, *Bordetella
pertussis*; Biosensor 20, *Mycobacterium tuberculosis*; Biosensor 21,
Cryptococcus neoformans; Biosensor 22, *Leptospira interrogans*; Biosensor
23, *Klebsiella pneumoniae*; Biosensor 24, *Shigella flexneri*; Biosensor 25,
Mycoplasma pneumonia; Biosensor 26, *Haemophilus parainfluenzae*;
Biosensor 27, human rhinovirus; Biosensor 28, respiratory syncytial virus type
A; Biosensor 29, adenoviruses; Biosensor 30, blank control (distilled water).
CL, control line; TL, test line.

level of *N. gonorrhoeae* template (20 copies) was tested when
the amplification lasted for 30, 40, and 50 min. Hence, a
30-min reaction time was recommended for assay at the MCDA
pre-amplification stage.

Specificity of the *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Assay

Assay specificity was evaluated and verified using an
N. gonorrhoeae reference strain (ATCC 49926), eight
N. gonorrhoeae-positive clinical samples (confirmed with
cultivation), and 20 non-*N. gonorrhoeae* samples (bacteria,
viruses, and fungi) (**Table 2**). Optimal assay reaction
conditions were implemented as outlined, and the results
were read using an LFB. Positive results were only observed
from *N. gonorrhoeae* templates, while other pathogens
and BCs were negative (**Figure 7**). No cross-reactions
were observed. Combined, these data suggested that the

N. gonorrhoeae-MCDA-LFB assay was highly selective
for *orf1*.

Confirming *Neisseria gonorrhoeae*-Multiple Cross Displacement Amplification-Lateral Flow Biosensor Assay Feasibility Using Clinical Specimens

We used 116 suspected *N. gonorrhoeae*-infection samples
for analysis, and then, all of the samples were detected
simultaneously with culture-biotechnical, qPCR, and
MCDA-LFB. Our data showed that 47 of 116 samples were
N. gonorrhoeae-positive using the *N. gonorrhoeae*-MCDA-LFB
assay. Cultivation assay data were also consistent with this
result. For qPCR, only 45 samples were confirmed as positive
(**Table 3** and **Supplementary Table 1**). These results suggested
our *N. gonorrhoeae*-MCDA-LFB assay functioned as an advanced
gonorrhea diagnostic tool for clinical specimens.

DISCUSSION

The Gram-negative diplococcus, *N. gonorrhoeae* causes sexually
transmitted gonorrhea and is a major global public health
concern (Comunian-Carrasco et al., 2018; Vallely et al., 2021).
The overwhelming majority of gonococcal infections are in less-
developed regions, likely owing to a lack of well-functioning
public health infrastructure and poor availability of laboratory
diagnostics (Kirkcaldy et al., 2019). Therefore, a reliable, cost-
saving, rapid, sensitive, and easy-to-use assay is critical for
addressing the ever-increasing gonorrhea transmission rates. In
this study, a novel *N. gonorrhoeae*-MCDA-LFB assay, which
integrated gold nanoparticle-based LFB with MCDA was devised
and successfully applied to the rapid and visual identification of
N. gonorrhoeae in clinical specimens.

The ideal laboratory-based diagnostic method should be
accurate, sensitive, specific, rapid, cheap, and easy-to-use.
Traditionally, cultivation was considered the gold standard for
N. gonorrhoeae detection; however, the bacteria are fastidious
and require specific culture conditions, otherwise false negative
results may be generated (Meyer and Buder, 2020; Visser
et al., 2020). Furthermore, culturing is labor intensive and
time consuming (approximately 2 days). Compared with
cultivation, direct microscopy is relatively simple and rapid,
but is not recommended as a diagnostic method owing
to low sensitivity (Thorley and Radcliffe, 2015; Meyer and
Buder, 2020). NAATs are more sensitive and specific and are
considered primary diagnostic methods for *N. gonorrhoeae*
detection (Perera et al., 2017; Dona et al., 2018). However,
in underdeveloped regions, their use is confined due to the
requirement for trained operators and expensive thermal cyclers.
Here, we developed an *N. gonorrhoeae*-MCDA-LFB technique
requiring simple instruments, a heating block, water bath,
or even a thermos cup that can hold 67°C for 30 min.
Moreover, amplification products were objectively visualized
using an LFB. The detection process, including genomic DNA

TABLE 3 | Conventional cultivation, quantitative polymerase chain reaction (qPCR), and multiple cross displacement amplification-lateral flow biosensor (MCDA-LFB) comparisons for *N. gonorrhoeae* testing in clinical samples.

Detection assay	Results	Gold standard method (cultivation)		True positive rate (%)	True negative rate (%)
		+	−		
qPCR	+	45	0	95.7[a]	100
	−	2	69		
MCDA-LFB	+	47	0	100	100
	−	0	69		

[a]*Statistically significant (p < 0.05) when compared with MCDA-LFB.*

extraction (∼15 min), MCDA amplification (30 min), and result readout (∼2 min) can be accomplished within 50 min. In the current study, the method of template preparation (nucleic acid-releasing agents) contributed to the rapidity of the assay and simplified the detection process. It is a relatively crude extraction technique, and more investigation is needed to empirically confirm this method.

To improve *N. gonorrhoeae*-MCDA-LFB assay sensitivity, MCDA was used to pre-amplify *orf1*. MCDA is a novel isothermal amplification method first devised by Wang et al. (2015) and is more sensitive than PCR and loop-mediated isothermal amplification methods (Wang et al., 2015; Zhu et al., 2021). The isothermal amplification of specific DNA sequences is performed using a set of 10 primers spanning 10 distinct regions of the target fragment; displacement primers (F1 and F2), cross primers (CP1 and CP2), and amplification primers (D1, D2, C1, C2, R1, and R2). The amplification procedure requires a *Bst* DNA polymerase with strand displacement capability at a single temperature (between 60°C and 68nd (Wang et al., 2015). In our study, a set of *N. gonorrhoeae*-MCDA primers were specifically designed to identify 10 regions in *orf1*. *N. gonorrhoeae*-MCDA-LFB assay specificity was verified using *N. gonorrhoeae* strains and other microbes. Positive results were only identified from gonococci isolates, and no cross reactions were observed with non-*N. gonorrhoeae* microbes (**Table 2** and **Figure 7**). Hence, the *N. gonorrhoeae*-MCDA-LFB assay was highly specific for *N. gonorrhoeae* identification. Similarly, the assay detected as low as 20 *N. gonorrhoeae* copies/test. To identify *N. gonorrhoeae* in clinical samples, our assay displayed higher sensitivity than qPCR (**Table 3** and **Supplementary Table 1**) and correctly diagnosed 100% (47/47) of *N. gonorrhoeae* samples identified by cultivation.

In this study, a gold nanoparticle-based LFB was used to detect *N. gonorrhoeae*-MCDA *orf1* amplicons. This paper-based platform is widely used in clinical settings due to its high selectivity, low LoD, low sample volume, low cost, robustness, rapidity, and user-friendly format (Quesada-González and Merkoçi, 2015; Koczula and Gallotta, 2016; Urusov et al., 2019; Wang et al., 2021). The LFB rapidly and visually detected *N. gonorrhoeae*-MCDA products for labeling with BSA-biotin and anti-FAM on LFB strips. Positive results were indicated by two crimson red bands on the LFB, CL, and TL. For negative outcomes, only the CL was indicated on the biosensor. Although real-time turbidity and MG generated *N. gonorrhoeae*-MCDA results, the former is costly and requires expensive

equipment, while the latter is ambiguous when MCDA amplicon concentrations were low (**Figure 5**). The cost of each LFB in this study was calculated at approximately $2 US dollars (USD). Therefore, the total cost of each *N. gonorrhoeae*-MCDA-LFB test, including genomic DNA preparation (∼$1 USD), MCDA amplification (∼$3.5 USD), and LFB detection (∼$2 USD) was estimated at $6.5 USD. Isothermal amplification methods, including loop-mediated isothermal amplification (LAMP) and cross-priming amplification (CPA), have also been used to test *N. gonorrhoeae*. Edwards et al. (2014) utilized the LAMP assay for testing *N. gonorrhoeae* with a minimum of 20 copies per reaction. Yu et al. (2016) used CPA to test *N. gonorrhoeae* and *Chlamydia trachomatis* with the detection limits of 65 and 45 copies per reaction, respectively. However, these methods must rely on agarose gel electrophoresis and colorimetric indicator. In this study, we first combined MCDA amplification with LFB for diagnosis of *N. gonorrhoeae*, which is more convenience and rapid than that of methods.

The *N. gonorrhoeae*-MCDA-LFB also has some limitations. First, the molecular assays identify nucleic acid sequences specific for the pathogen. However, a positive result may indicate either an infectious or non-infectious agent. Second, the LFB detection must be taken off the lid of MCDA amplification tube. Also, there is a risk of contamination with post-reaction processing of MCDA amplicons. To control this, spraying a 10–15% sodium hypochlorite solution and 70% ethanol after LFB assay completion is effective in avoiding DNA contamination in laboratory conditions. We observed no cross contamination with non-*N. gonorrhoeae* isolates. It is indicated that the false positive rate has been effectively controlled in our laboratory.

CONCLUSION

We successfully integrated an LFB platform with MCDA pre-amplification technology to devise a novel *N. gonorrhoeae*-MCDA-LFB assay for the sensitive, specific, rapid, cheap, and visual diagnosis of *N. gonorrhoeae* in clinical specimens. Our assay detected 20 genomic DNA copies/test and exhibited no cross reaction with other microbes. The detection process was approximately 50 min and did not require expensive apparatus. Hence, our assay shows great potential as a POC test for the identification of *N. gonorrhoeae* in clinical settings, especially in resource-limited regions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The study was approved by the Human Ethics Committee of Hangzhou Women's Hospital [Approval No. (2021)-K (2)-8] and complied with the Declaration of Helsinki. Before clinical samples/isolates were obtained, all personal patient identifiers were removed. Patient informed consent was waived by the Committee.

AUTHOR CONTRIBUTIONS

XC, LH, and SD designed and conceived the study and wrote and revised the manuscript. XC, LH, and YT participated in the primer design. QZ, YT, XT, and SD collected the clinical samples. XC, YT, QZ, and SD performed all the laboratory work. XC and QZ performed the statistical analyses. All authors read and approved the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.747140/full#supplementary-material

REFERENCES

Abraha, M., Egli-Gany, D., and Low, N. (2018). Epidemiological, behavioural, and clinical factors associated with antimicrobial-resistant gonorrhoea: a review. *F1000Res.* 7:400. doi: 10.12688/f1000research.13600.1

Chan, P. A., Robinette, A., Montgomery, M., Almonte, A., Cu-Uvin, S., Lonks, J. R., et al. (2016). Extragenital infections caused by *Chlamydia trachomatis* and *Neisseria gonorrhoeae*: a review of the literature. *Infect. Dis. Obstet. Gynecol.* 2016:5758387. doi: 10.1155/2016/5758387

Chaudhry, U., and Saluja, D. (2002). Detection of *Neisseria gonorrhoeae* by PCR using orf1 gene as target. *Sex. Transm. Infect.* 78:72. doi: 10.1136/sti.78.1.72

Chen, X., Zhou, Q., Dong, S., Wang, S., Liu, R., Wu, X., et al. (2021). Multiple cross displacement amplification linked with nanoparticles-based lateral flow biosensor in screening of hepatitis B virus in clinical application. *Infect. Drug Resist.* 14, 1219–1229. doi: 10.2147/IDR.S297645

Cheng, X., Dou, Z., Yang, J., Gu, Y., Liu, D., Xie, L., et al. (2020). Highly sensitive and rapid identification of *Streptococcus agalactiae* based on multiple cross displacement amplification coupled with lateral flow biosensor assay. *Front. Microbiol.* 11:1926. doi: 10.3389/fmicb.2020.01926

Comunian-Carrasco, G., Pena-Marti, G. E., and Marti-Carvajal, A. J. (2018). Antibiotics for treating gonorrhoea in pregnancy. *Cochrane Database Syst. Rev.* 2:D11167. doi: 10.1002/14651858.CD011167.pub2

Dempsey, J. A., Litaker, W., Madhure, A., Snodgrass, T. L., and Cannon, J. G. (1991). Physical map of the chromosome of *Neisseria gonorrhoeae* FA1090 with locations of genetic markers, including opa and pil genes. *J. Bacteriol.* 173, 5476–5486. doi: 10.1128/jb.173.17.5476-5486.1991

Dona, V., Smid, J. H., Kasraian, S., Egli-Gany, D., Dost, F., Imeri, F., et al. (2018). Mismatch amplification mutation assay-based real-time PCR for rapid detection of *Neisseria gonorrhoeae* and antimicrobial resistance determinants in clinical specimens. *J. Clin. Microbiol.* 56, e00365–18. doi: 10.1128/JCM.00365-18

Edwards, T., Burke, P. A., Smalley, H. B., Gillies, L., and Hobbs, G. (2014). Loop-Mediated isothermal amplification test for detection of *Neisseria gonorrhoeae* in urine samples and tolerance of the assay to the presence of urea. *J. Clin. Microbiol.* 52, 2163–2165. doi: 10.1128/JCM.00314-14

Golparian, D., Borang, S., Sundqvist, M., and Unemo, M. (2015). Evaluation of the new BD max GC real-time PCR assay, analytically and clinically as a supplementary test for the BD ProbeTec GC qx amplified DNA assay, for molecular detection of *Neisseria gonorrhoeae*. *J. Clin. Microbiol.* 53, 3935–3937. doi: 10.1128/JCM.01962-15

Hook, E. W., and Bernstein, K. (2019). Kissing, saliva exchange, and transmission of *Neisseria gonorrhoeae*. *Lancet Infect. Dis.* 19, e367–e369. doi: 10.1016/S1473-3099(19)30306-8

Jefferson, A., Smith, A., Fasinu, P. S., and Thompson, D. K. (2021). Sexually transmitted *Neisseria gonorrhoeae* infections-update on drug treatment and vaccine development. *Medicines* 8:11. doi: 10.3390/medicines8020011

Kirkcaldy, R. D., Weston, E., Segurado, A. C., and Hughes, G. (2019). Epidemiology of gonorrhoea: a global perspective. *Sex. Health* 16:401. doi: 10.1071/SH19061

Koczula, K. M., and Gallotta, A. (2016). Lateral flow assays. *Essays Biochem.* 60, 111–120. doi: 10.1042/EBC20150012

Li, S., Jiang, W., Huang, J., Liu, Y., Ren, L., Zhuang, L., et al. (2020). Highly sensitive and specific diagnosis of COVID-19 by reverse transcription multiple cross-displacement amplification-labelled nanoparticles biosensor. *Eur. Respir. J.* 56:2002060. doi: 10.1183/13993003.02060-2020

Meyer, T., and Buder, S. (2020). The laboratory diagnosis of *neisseria gonorrhoeae*: current testing and future demands. *Pathogens* 9:91. doi: 10.3390/pathogens9020091

Młynarczyk-Bonikowska, B., Majewska, A., Malejczyk, M., Młynarczyk, G., and Majewski, S. (2020). Multiresistant *Neisseria gonorrhoeae*: a new threat in second decade of the XXI century. *Med. Microbiol. Immun.* 209:641. doi: 10.1007/s00430-019-00651-4

Perera, S. R., Khan, N. H., Martin, I., Taheri, A., Parti, R. P., Levett, P. N., et al. (2017). Multiplex real-time PCR assay for simultaneous identification of *Neisseria gonorrhoeae* and its ciprofloxacin susceptibility status. *J. Clin. Microbiol.* 55, 3201–3209. doi: 10.1128/JCM.00855-17

Quesada-González, D., and Merkoçi, A. (2015). Nanoparticle-based lateral flow biosensors. *Biosens. Bioelectron.* 73, 47–63. doi: 10.1016/j.bios.2015.05.050

Quillin, S. J., and Seifert, H. S. (2018). *Neisseria gonorrhoeae* host adaptation and pathogenesis. *Nat. Rev. Microbiol.* 16, 226–240. doi: 10.1038/nrmicro.2017.169

Rowley, J., Vander Hoorn, S., Korenromp, E., Low, N., Unemo, M., Abu-Raddad, L. J., et al. (2019). *Chlamydia*, gonorrhoea, trichomoniasis and syphilis: global

prevalence and incidence estimates, 2016. *Bull. World Health Organ.* 97, 548–562. doi: 10.2471/BLT.18.228486

Stevens, J. S., and Criss, A. K. (2018). Pathogenesis of *Neisseria gonorrhoeae* in the female reproductive tract. *Curr. Opin. Hematol.* 25, 13–21. doi: 10.1097/MOH.0000000000000394

Thorley, N., and Radcliffe, K. (2015). The performance and clinical utility of cervical microscopy for the diagnosis of gonorrhoea in women in the era of the NAAT. *Int. J. STD AIDS* 26, 656–660. doi: 10.1177/0956462414551233

Urusov, A. E., Zherdev, A. V., and Dzantiev, B. B. (2019). Towards lateral flow quantitative assays: detection approaches. *Biosensors (Basel)* 9:89. doi: 10.3390/bios9030089

Vallely, L. M., Egli-Gany, D., Wand, H., Pomat, W. S., Homer, C. S. E., Guy, R., et al. (2021). Adverse pregnancy and neonatal outcomes associated with *Neisseria gonorrhoeae*: systematic review and meta-analysis. *Sex. Transm. Infect.* 97, 104–111. doi: 10.1136/sextrans-2020-054653

Visser, M., van Westreenen, M., van Bergen, J., and van Benthem, B. H. B. (2020). Low gonorrhoea antimicrobial resistance and culture positivity rates in general practice: a pilot study. *Sex. Transm. Infect.* 96, 220–222. doi: 10.1136/sextrans-2019-054006

Wang, T., Chen, L., Chikkanna, A., Chen, S., Brusius, I., Sbuh, N., et al. (2021). Development of nucleic acid aptamer-based lateral flow assays: a robust platform for cost-effective point-of-care diagnosis. *Theranostics* 11, 5174–5196. doi: 10.7150/thno.56471

Wang, Y., Li, H., Wang, Y., Li, H., Luo, L., Xu, J., et al. (2017). Development of multiple cross displacement amplification label-based gold nanoparticles lateral flow biosensor for detection of *Listeria monocytogenes*. *Int. J. Nanomedicine* 12, 473–486. doi: 10.2147/IJN.S123625

Wang, Y., Wang, Y., Ma, A., Li, D., Luo, L., Liu, D., et al. (2015). Rapid and sensitive isothermal detection of nucleic-acid sequence by multiple cross displacement amplification. *Sci. Rep*. 5:11902. doi: 10.1038/srep11902

Yu, B., An, Y., Xu, G., and Shan, H. (2016). Detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* based on cross-priming amplification. *Lett. Appl. Microbiol.* 62, 399–403. doi: 10.1111/lam.12560

Zeng, L., Lie, P., Fang, Z., and Xiao, Z. (2013). Lateral flow biosensors for the detection of nucleic acid. *Methods Mol. Biol.* 1039, 161–167. doi: 10.1007/978-1-62703-535-4_14

Zhu, X., Wang, X., Li, S., Luo, W., Zhang, X., Wang, C., et al. (2021). Rapid, ultrasensitive, and highly specific diagnosis of COVID-19 by CRISPR-based detection. *ACS Sensors* 6, 881–888. doi: 10.1021/acssensors.0c01984

The Effect of Allicin on the Proteome of SARS-CoV-2 Infected Calu-3 Cells

Kirstin Mösbauer[1,2†], Verena Nadin Fritsch[3†], Lorenz Adrian[4,5], Jörg Bernhardt[6], Martin Clemens Horst Gruhlke[7], Alan John Slusarenko[7], Daniela Niemeyer[1,2] and Haike Antelmann[3]*

[1] Institute of Virology, Berlin Institute of Health, Charité-Universitätsmedizin Berlin, Freie Universität Berlin, Berlin, Germany, [2] German Centre for Infection Research (DZIF), Berlin, Germany, [3] Institute for Biology-Microbiology, Freie Universität Berlin, Berlin, Germany, [4] Department Environmental Biotechnology, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany, [5] Fachgebiet Geobiotechnologie, Technische Universität Berlin, Berlin, Germany, [6] Institute for Microbiology, University of Greifswald, Greifswald, Germany, [7] Department of Plant Physiology, RWTH Aachen University, Aachen, Germany

Allicin (diallyl thiosulfinate) is the major thiol-reactive organosulfur compound produced by garlic plants (*Allium sativum*) upon tissue damage. Allicin exerts its strong antimicrobial activity against bacteria and fungi via *S*-thioallylation of protein thiols and low molecular weight thiols. Here, we investigated the effect of allicin on SARS-CoV-2 infected Vero E6 and Calu-3 cells. Toxicity tests revealed that Calu-3 cells showed greater allicin tolerance, probably due to >4-fold higher GSH levels compared to the very sensitive Vero E6 cells. Exposure of infected Vero E6 and Calu-3 cells to biocompatible allicin doses led to a ~60–70% decrease of viral RNA and infectious viral particles. Label-free quantitative proteomics was used to investigate the changes in the Calu-3 proteome after SARS-CoV-2 infection and the effect of allicin on the host-virus proteome. SARS-CoV-2 infection of Calu-3 cells caused a strong induction of the antiviral interferon-stimulated gene (ISG) signature, including several antiviral effectors, such as cGAS, Mx1, IFIT, IFIH, IFI16, IFI44, OAS, and ISG15, pathways of vesicular transport, tight junctions (KIF5A/B/C, OSBPL2, CLTCL1, and ARHGAP17) and ubiquitin modification (UBE2L3/5), as well as reprogramming of host metabolism, transcription and translation. Allicin treatment of infected Calu-3 cells reduced the expression of IFN signaling pathways and ISG effectors and reverted several host pathways to levels of uninfected cells. Allicin further reduced the abundance of the structural viral proteins N, M, S and ORF3 in the host-virus proteome. In conclusion, our data demonstrate the antiviral and immunomodulatory activity of biocompatible doses of allicin in SARS-CoV-2-infected cell cultures. Future drug research should be directed to exploit the thiol-reactivity of allicin derivatives with increased stability and lower human cell toxicity as antiviral lead compounds.

Keywords: allicin, SARS-CoV-2, proteome, Vero E6, Calu-3

INTRODUCTION

The Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) causes Coronavirus disease (COVID-19), which represents a global health burden (Zhou et al., 2020). COVID-19 is often associated with immunopathology since severely ill patients had decreased levels of T lymphocytes, including regulatory T cells, cytotoxic and helper T cells, and natural killer cells (Qin et al., 2020; Wei et al., 2020). Patients with severe illness showed a cytokine storm syndrome associated with a dysregulated immune activation and hyperinflammation (Fara et al., 2020). High levels of pro-inflammatory cytokines IL-1ß, IL-2, IL-6, IL-7, IL-10, macrophage inflammatory protein-1A (MIP-1A), TNF-α, and INF-γ have been detected, connecting the uncontrolled inflammation and dysregulation of the immune response with the high mortality in severely ill COVID-19 patients (Fara et al., 2020; Qin et al., 2020; Wei et al., 2020). While mild infections were characterized by highly activated HLA-DR^{hi}CD11c^{hi} inflammatory monocytes with the interferon-stimulated gene (ISG) signature, severe illness was manifested by dysfunctional neutrophil precursors, and HLA-DR^{lo} monocytes with pro-inflammatory functions (Schulte-Schrepping et al., 2020). These immunological markers of pro-inflammatory cytokines and the dysfunctional myeloid compartment might help to identify drug targets to prevent progression to severe illness (Fara et al., 2020; Schulte-Schrepping et al., 2020).

While global vaccination campaigns are underway, the development of efficient therapies to prevent COVID-19 disease progression is an urgent need. Garlic plants (*Allium sativum*) produce volatile organosulfur compounds, such as diallyl thiosulfinate (allicin) and diallyl polysulfanes, which are known for their antimicrobial, antiviral, anticancer, anti-inflammatory and immunomodulatory effects (Borlinghaus et al., 2014, 2021; Schäfer and Kaschula, 2014). Garlic compounds showed broad-spectrum antimicrobial activity against several pathogenic bacteria, viruses, fungi, and parasites (Rabinkov et al., 1998; Münchberg et al., 2007; Block, 2010; Borlinghaus et al., 2014, 2021; Reiter et al., 2017; Arbach et al., 2019; Loi et al., 2019; Rouf et al., 2020).

Allicin is a thiol-reactive compound, which reacts with Cys thiols via thiol-disulfide exchange reactions, leading to *S*-thioallylations of proteins (Miron et al., 2000, 2010). Widespread *S*-thioallylations of redox-sensitive Cys residues in proteins were identified in the proteome of human Jurkat cells, *Escherichia coli*, *Staphylococcus aureus*, and *Bacillus subtilis* (Rabinkov et al., 1998; Miron et al., 2010; Müller et al., 2016; Chi et al., 2019; Gruhlke et al., 2019; Loi et al., 2019). In Jurkat cancer cells, 332 *S*-thioallylated proteins were identified 10 min after allicin treatment, including highly abundant cytoskeleton proteins, HSP90 chaperones, translation elongation factors and glycolytic enzymes. Allicin caused disruption of the actin cytoskeleton, enzymatic inactivation and Zn^{2+} release to stimulate the IL-1-dependent IL-2 secretion by T cells as an immunomodulatory effect (Gruhlke et al., 2019).

In addition, *S*-thioallylations deplete low molecular weight thiols, such as glutathione (GSH) and bacillithiol (BSH) in bacteria and yeast cells (Gruhlke et al., 2010, 2019;

Arbach et al., 2019). Thus, allicin leads to oxidative stress responses, inhibition of protein functions and an impaired cellular redox balance. Since SARS-CoV-2 is rich in Cys residues in its surface spike glycoprotein, a reduced state of the host cell cytoplasm is required for efficient virus entry and membrane fusion. Moreover, allicin is cell permeable and has been shown to cause transient pore formation in phospholipid membranes (Miron et al., 2000; Gruhlke et al., 2015). The antiviral effect of allicin has been previously investigated against several respiratory viruses, including influenza, SARS-CoV and rhinovirus (Rouf et al., 2020).

In this work, we show that allicin at biocompatible doses decreases infectious viral particles and viral RNA of SARS-CoV-2 in the primate kidney-derived cell line Vero E6 and the human lung cell line Calu-3. We further identified proteome changes caused by SARS-CoV-2 infection and the effect of allicin on these host pathways. While the interferon-stimulated gene (ISG) signature was most prominently upregulated in SARS-CoV-2 infected Calu-3 cells, the ISG response and several host cellular pathways were restored to levels of untreated cells by allicin. Thus, allicin exerts an antiviral and immunomodulatory effect when applied in infected cell cultures *in vitro*, which is supported at the proteome level.

MATERIALS AND METHODS

Cultivation of Cell Lines and Infection Experiments With SARS-CoV-2

Vero E6 (ATCC CRL-1586) and Calu-3 (ATCC HTB-55) cell lines were cultivated in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 1% non-essential amino acids and 1% sodium pyruvate (Gibco), and grown at 37°C and 5% CO$_2$. Cell lines were free of mycoplasma, authenticated based on morphology and growth properties and confirmed by PCR. The cell cultures were used for viability or infection assays below cell passage 20. No antibiotics have been used during cultivation of eukaryotic cells.

The infection experiments were performed with SARS-CoV-2 Munich isolate (CSpecVir985) under biosafety level 3 conditions with appropriate respiratory personal protection equipment. Vero E6 and Calu-3 cells were seeded at densities of 3.5×10^5 or 6×10^5 cells/ml in 12-well TC plates (TPP Techno Plastic Products AG), respectively. After 24 h, cells were infected at a MOI of 0.01 or 0.005, diluted in serum-free OptiPro medium for 1 h at 37°C. The medium was removed and cells were washed twice with phosphate-buffered saline (PBS) followed by addition of DMEM and supplements. Samples were taken at 16 and 24 h p.i. for further analysis.

Allicin Synthesis and Treatment

Allicin was synthesized by oxidation of 3-[(Prop-2-en-1-yl)disulfanyl]prop-1-ene (diallyl disulfide, Sigma-Aldrich, Germany) with peracetic-acid (glacial acetic acid/H$_2$O$_2$) as described previously (Gruhlke et al., 2010). To analyze the antiviral effect of allicin, SARS-CoV-2 infection experiments were performed with an allicin pre- and post-treatment of Vero

E6 cells. For the pre-treatment, either the cells or the virus dilution were incubated with 50 µM allicin for 30 min. We have chosen 50 µM allicin since this concentration was determined as sub-lethal for Vero E6 cells. Pre-treated cells were washed with PBS and infected according to the infection protocol as described above. Pre-treated virus was used in the infection experiment according to the protocol above. Post-treatment of SARS-CoV-2 infected cells was accomplished by adding the indicated concentration of allicin into the medium after infection. Thus, in the post-treatment protocol, the added allicin remained on the infected cells until sample collection after 16 and 24 h.

Cell Viability Assay

The cell viability of Vero E6 and Calu-3 cells was analyzed by quantification of ATP levels using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega) according to the instructions of the manufacturer. The cells were cultivated as described above in 96-well flat clear bottom black TC-treated microplates (Corning®) and exposed to different amounts of allicin for 24 h. Cell viability of treated cells was normalized to non-treated cells.

Determination of the Levels of Glutathione and Glutathione Disulfide in Vero E6 and Calu-3 Cells

Vero E6 and Calu-3 cells were cultivated as described above in 96-well white opaque flat bottom tissue culture plate (Falcon) and seeded at densities of 1×10^4 cells/well. After washing with PBS, the intracellular GSH and glutathione disulfide (GSSG) concentrations were determined using the GSH/GSSG-Glo™ assay (Promega) according to the instructions of the manufacturer for adherent cells. Briefly, total GSH levels were measured in one sample by reduction of GSSG to GSH using DTT. Total GSSG amounts were measured in a second sample by blocking reduced GSH with N-ethylmaleimide (NEM), followed by GSSG reduction with DTT. The GSH transferase (GST) uses GSH as cofactor to convert luciferin-NT to GSH-NT resulting in the release of luciferin. Luciferin is oxidized to oxyluciferin by the Ultra-Glo™ rLuciferase, leading to emission of chemiluminescence, which was measured using an integration time of 1 s/well by the CLARIOstar microplate reader (BMG Labtech). GSH levels were calculated based on GSH standard curves. For determination of the cellular GSH levels, the GSSG amounts were subtracted from the total GSH level.

Plaque Titration Assay

The number of infectious virus particles was determined by a plaque titration assay. Vero E6 monolayers were seeded in 24-well TC plates (TPP Techno Plastic Products AG) and infected with 200 µl of serial dilutions of SARS-CoV-2 containing cell culture supernatants of infected Vero E6 or Calu-3 cells, which were diluted in OptiPro serum-free medium. After 1 h adsorption, the supernatant was removed and cells overlaid with 1.2% Avicel (FMC BioPolymers) diluted in DMEM. After 72 h, the overlay was removed, cells were fixed in 6% formaldehyde and plaques were visualized by crystal violet staining.

Viral RNA Extraction and Real-Time Reverse-Transcription PCR

Viral RNA extraction was performed from 50 µl culture supernatant of SARS-CoV-2 infected Vero E6 and Calu-3 cells using the viral RNA kit (Macherey-Nagel) according to the instructions of the manufacturer. SARS-CoV-2 genome equivalents (GE) were detected by quantitative RT–PCR [LightCycler 480 Real-Time PCR System and Software version 1.5 (Roche)], targeting the SARS-CoV-2 E gene using the primers E gene-F (5′-ACAGGTACGTTAATAGTTAATAGCGT-3′) and E gene-R (5′-ATATTGCAGCAGTACGCACACA-3′). Absolute quantification was performed using SARS-CoV-2 specific *in vitro*-transcribed RNA standards as described previously (Corman et al., 2020).

Proteome Analysis of SARS-CoV-2 Infected Host Cells Using Orbitrap Fusion Mass Spectrometry

6×10^5 Calu-3 cells per sample were infected with SARS-CoV-2 as described above and treated with 150 µM allicin for 24 h. Calu-3 cells were harvested by centrifugation. The cell pellets were washed with PBS and alkylated for 15 min at room temperature (RT) under denaturing conditions in 200 µl of UCE-IAM buffer, consisting of 8 M urea, 1% (w/v) CHAPS, 1 mM EDTA, 200 mM Tris–HCl pH 8.0 and 100 mM IAM as described (Rossius et al., 2018). Subsequently, the alkylated protein extracts were precipitated with trizol and 96% ethanol and washed four times with 1 ml 70% ethanol. The protein pellets were separated by a short 15% non-reducing SDS-PAGE, which was running for 15 min and stained with Colloidal Coomassie Blue. The gel fractions were cut and in-gel tryptic digested as described previously (Rossius et al., 2018). The eluted peptides were desalted using ZipTip-µC18 material (Merck Millipore) and dissolved in 0.1% (v/v) formic acid before LC-MS/MS analysis. The peptide samples of non-infected Calu-3 cells (Mock) and SARS-CoV-2 infected Calu-3 cells with and without allicin treatment were subjected to nLC-MS/MS analysis using an Orbitrap Fusion (Thermo Fisher Scientific) coupled to a TriVersa NanoMate (Advion, Ltd.) as described previously (Kublik et al., 2016). Peptide identification of the human and SARS-CoV-2 proteome was performed by Proteome Discoverer (version 2.2, Thermo Fisher Scientific) using the SequestHT search engine as described (Seidel et al., 2018). Human and SARS-CoV-2 proteins were identified by searching all tandem MS/MS spectra against the human proteome protein sequence database (20,286 entries) extracted from UniprotKB release 12.7 (UniProt Consortium, Nucleic acids research 2007, 35, D193-197) as well as against the European Virus Archive Global # 026V-03883 sequence database. Peptides were considered to be identified with high confidence at a target false discovery rate of ≤0.01 and with a medium confidence at ≤0.05, based on the q-values. Identified proteins were quantified by the "Percursor Ions Quantifier" implemented in Proteome Discoverer 2.2 based on peak intensities to estimate the abundance of the human and SARS-CoV-2 proteins in the peptide samples. Error tolerance

for precursor ion and fragment ion *m/z* values was set to 3 ppm and 0.5 Da, respectively. Two missed cleavage sites were allowed. Methionine oxidation (+15.994915 Da), cysteine carbamidomethylation (+57.021464 Da) and cysteine *S*-thioallylation by allicin (+72.00337 Da for $C_3H_5S_1$) were set as variable modifications. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Perez-Riverol et al., 2019; Deutsch et al., 2020) with the dataset identifier PXD024375.

Statistical Analyses

Statistical analysis of the cell viability assays, the GSH and GSSG measurements as well as the determination of viral RNA and infectious particles were performed from 3–4 biological replicates with 1–3 technical replicates using the Student's unpaired two-tailed *t*-test for two samples with unequal variance. Proteomics analyses were performed from 3–4 biological replicates with 1–3 technical replicates. For calculation of the statistics of the proteomics data, the LFQ intensity values of each proteomics sample and every single treatment were tested for normality by using Jarque Bera (testing for kurtosis and skewness) (Jarque and Bera, 1980) and Anderson Darling (based on Kolmogorov-Smirnov) tests (Anderson and Darling, 1954). Accordingly, *p*-values for pairwise treatment comparisons were calculated by the Welsh test (Student's unpaired two-tailed *t*-test for two samples with unequal variance and heteroscedastic data). The *p*-values and significance levels are included in the figure and table legends.

RESULTS

Biocompatible Allicin Concentrations Correlate With the Intracellular Glutathione Levels in Vero E6 and Calu-3 Cells

While allicin has many beneficial effects for human health, crushed garlic is also toxic and harmful for human cells. Fresh garlic can cause severe cellular and tissue damage upon direct exposure to the epithelial cells and mucous membranes of the respiratory tract and the skin, such as garlic burns (Bautista et al., 2005; Al-Qattan, 2009; Vargo et al., 2017; Hitl et al., 2021; Muniz et al., 2021). Thus, we first assessed the toxicity of allicin in Calu-3 and Vero E6 cells, which are used here as cell culture models for SARS-CoV-2 infection. Using cell viability assays, the biocompatible, non-harmful doses of allicin in Calu-3 and Vero E6 cells were determined. Both cell lines differed strongly in their susceptibilities toward allicin. Calu-3 cells showed high viability rates of ∼85% after treatment with 200 µM allicin. Even concentrations of 300 µM allicin decreased the viability rate of Calu-3 cells only non-significantly to ∼70% (**Figure 1A**). Treatment of Vero E6 cells with 75 µM allicin led to a cell viability rate of 84% (**Figure 1B**), whereas 150 µM allicin resulted in killing of 99% of Vero E6 cells. Thus, the sub-lethal biocompatible doses of allicin were determined as 50–75 µM in Vero E6 cells and 100–200 µM in the more tolerant Calu-3 cells.

Previous studies already revealed strong variations in the susceptibilities of different cell lines toward allicin, which correlated with different intracellular GSH contents (Gruhlke et al., 2016, 2019). Thus, we measured the intracellular GSH and GSSG levels in Vero E6 and Calu-3 cells (**Figure 1C**). The GSH content of the more tolerant Calu-3 cells was determined as 3.2 µM, which was 4.2-fold higher compared to only 0.77 µM GSH as measured in Vero E6. As expected, the amounts of GSSG were very low with 0.05 µM and 0.009 µM in Calu-3 and Vero E6 cells, respectively. These data suggest that Calu-3 cells show greater allicin tolerance in part due to their higher GSH levels compared to Vero E6 cells.

Allicin Leads to Decreased Infectious Viral Particles and Viral RNA in SARS-CoV-2 Infected Vero E6 and Calu-3 Cells

The antiviral effect of allicin against SARS-CoV-2 was analyzed using pre- and post-treatment options for the more allicin-sensitive infected Vero E6 cells: (1) Cells were pre-exposed to 50 µM allicin for 30 min before SARS-CoV-2 infection. (2) The virus was treated with 50 µM allicin for 30 min prior to infection. (3) SARS-CoV-2 infected Vero E6 cells were treated with 50 µM allicin post infection (p.i.) (**Figure 2A**). We have chosen 50 µM allicin since this concentration did not affect viability of Vero E6 cells (**Figure 1B**). The number of infectious SARS-CoV-2 particles (PFU, plaque forming units) was determined 24 h p.i. by the plaque titration assay. However, only post-treatment with 50 µM allicin led to a significant 70% decrease in the amount of infectious virus particles, whereas the pre-treatment of cells or virus caused only a 16–21% reduction of viral plaques (**Figure 2A**). These results suggest that allicin might affect host-virus interactions by its antiviral and immunomodulatory activities.

In addition, viral RNA genome equivalents (GE) were determined from the supernatant of infected Vero E6 cells using quantitative RT-PCR. In agreement with the plaque assays, the qRT-PCR results revealed a 72% lower amount of viral RNA after addition of 50 µM allicin to SARS-CoV-2 infected Vero E6 cells (**Figures 2B,C**). Moreover, virus plaque assays and qRT-PCR results showed an almost complete >99% inhibition of SARS-CoV-2 replication after exposure to 75 µM allicin, supporting the strong antiviral activity of allicin in infected Vero E6 cells (**Figures 2B,C**).

The antiviral effects of biocompatible doses of allicin were further analyzed in the more allicin-resistant Calu-3 cells. After infection with SARS-CoV-2 at a multiplicity of infection (MOI) of 0.01 and 0.005, Calu-3 cells were treated with biocompatible doses of 100 and 200 µM allicin and analyzed 16 and 24 h p.i., respectively (**Figure 3**). Treatment of infected Calu-3 cells with 100 µM allicin did not significantly inhibit viral replication (**Figures 3A–D**). However, exposure of infected Calu-3 cells to 200 µM allicin led to a significant >60% decrease of viral RNA (**Figures 3A,B**) and a >65% reduction of infectious particles (**Figures 3C,D**).
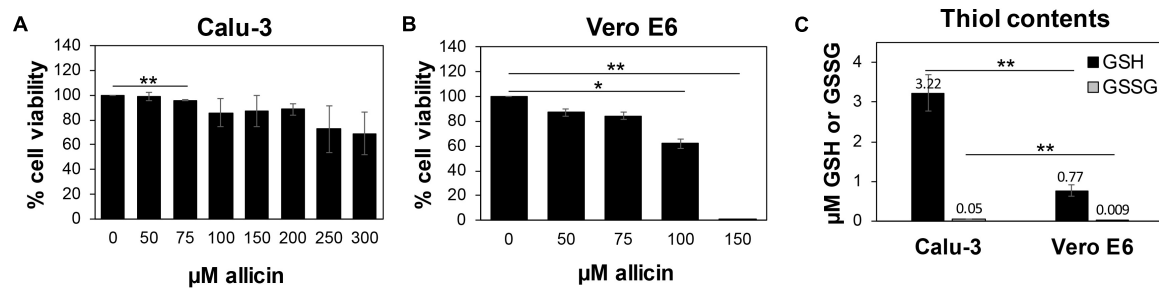
FIGURE 1 | Human Calu-3 cells are more resistant to allicin compared to Vero E6. **(A,B)** Cell viability of untreated and allicin treated Calu-3 **(A)** and Vero E6 cells **(B)** was measured after 24 h using the CellTiter-Glo® Luminescent Cell Viability Assay (Promega) according to the manufacturer's instructions. The viability of the control without allicin was set to 100%. The viability of Calu-3 cells was not significantly decreased upon exposure to 100–200 μM allicin, while concentrations of ≥100 μM allicin interfered with Vero E6 cell viability. **(C)** The levels of glutathione (GSH) and glutathione disulfide (GSSG) were determined in untreated Calu-3 and Vero E6 cells using the GSH/GSSG-Glo Assay (Promega) according to the manufacturer's instructions. The results are from 4 biological replicates with 3 technical replicates for **(C)**. Error bars represent the standard deviation (SD). *p*-values were calculated using an unpaired two-tailed *t*-test.*$p < 0.05$; **$p < 0.01$.
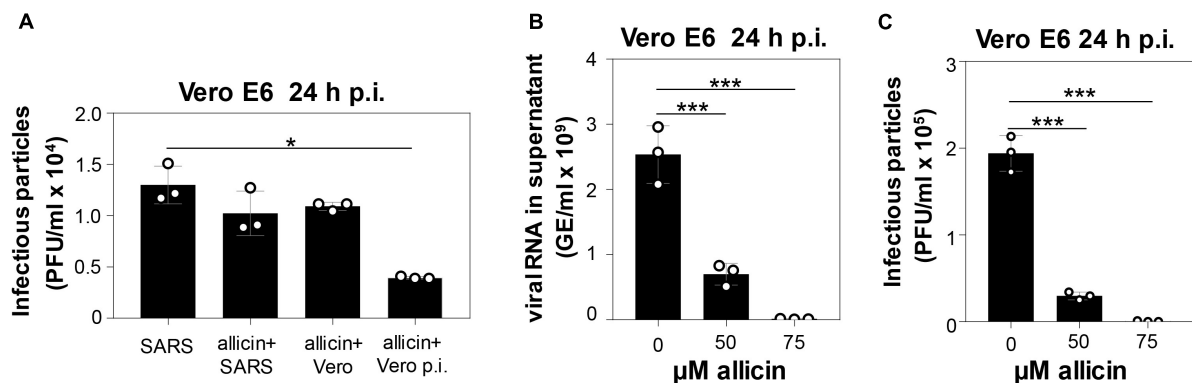
FIGURE 2 | Allicin treatment of SARS-CoV-2 infected Vero E6 cells reduced the levels of infectious viral particles and viral RNA. **(A–C)** Vero E6 cells were infected with SARS-CoV-2 at a MOI of 0.01. After 24 h p.i. viral replication was analyzed by determination of infectious viral particles or viral RNA in the supernatant. **(A)** Comparison of different allicin treatments: Untreated Vero E6 cells infected with SARS-CoV-2 as control (SARS); SARS-CoV-2 pre-treated with 50 μM allicin for 30 min prior to infection of host cells (allicin + SARS); host cells pre-treated with 50 μM allicin for 30 min prior to infection with SARS-CoV-2 (allicin + Vero); and SARS-CoV-2 infected host cells treated with 50 μM allicin p.i. (allicin + Vero p.i.). **(B,C)** The amount of viral RNA **(B)** and infectious viral particles **(C)** was determined after treatment of SARS-CoV-2 infected Vero E6 cells with 50 and 75 μM allicin p.i. The results **(A–C)** are from three biological replicates with two technical replicates for panel **(B)**. Error bars represent the SD. *p*-values were calculated using an unpaired two-tailed *t*-test.*$p < 0.05$; ***$p < 0.001$.

The antiviral effect of allicin on SARS-CoV-2 infected Calu-3 cells was further supported by microscopy imaging (**Figures 4A–C**). While SARS-CoV-2 infection at a MOI of 0.01 resulted in cellular damage of Calu-3 cells after 24 h p.i., the addition of 150 μM allicin partially protected the cells against this damage (**Figures 4B,C**). Taken together, our results indicate that biocompatible allicin doses exert an antiviral effect against SARS-CoV-2 in both Vero E6 and Calu-3 cells.

Changes in the Calu-3 Proteome After SARS-CoV-2 Infection

Label-free quantitative (LFQ) proteomics by Orbitrap Fusion LC-MS/MS analysis was used to investigate the changes in the proteome of Calu-3 cells after SARS-CoV-2 infection and the effect of 150 μM allicin. The concentration of 150 μM allicin was chosen since this was sub-lethal for Calu-3 cells (**Figure 1A**), and

protected the cells against SARS-CoV-2 damage (**Figure 4C**). The proteome samples of Calu-3 cells were analyzed before infection (Mock) and 24 h p.i. with SARS-CoV-2 at a MOI of 0.01 in the absence or presence of 150 μM allicin in 3–4 biological and 1–3 technical replicates. The total LFQ intensities of all proteins in each sample were normalized and represent 100% of the total protein abundance. Overall, we quantified 4,251 proteins, including 4,243 Calu-3 host proteins and 8 SARS-CoV-2 proteins in the total proteome (**Supplementary Tables 1,2**). After infection, about 207 and 329 proteins were ≥1.5-fold induced and <0.66-fold decreased, respectively (**Supplementary Table 3**). These 536 differentially expressed proteins contribute to only 2.73% of the total proteome abundance in SARS-CoV-2 infected Calu-3 cells (**Supplementary Table 3**). The proteins were sorted into KEGG Ontology (KO) or Uniprot categories and their fold-changes, *p*-values and averaged abundances were visualized in Voronoi treemaps as color gradients and
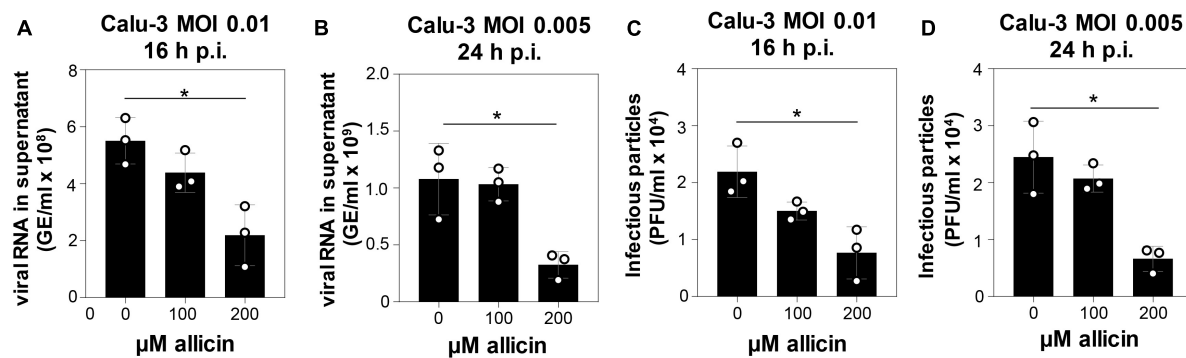
FIGURE 3 | Allicin treatment of SARS-CoV-2 infected Calu-3 cells leads to decreased amounts of infectious viral particles and viral RNA. **(A–D)** SARS-CoV-2 infected Calu-3 cells were treated with 100 and 200 μM allicin p.i. The amount of viral RNA **(A,B)** and infectious viral particles **(C,D)** was determined 16 h **(A,C)** and 24 h **(B,D)** p.i. with SARS-CoV-2 at a MOI of 0.01 **(A,C)** and 0.005 **(B,D)**. The results are from three biological replicates with two technical replicates for panels **(A,B)**. Error bars represent the SD. *p*-values were calculated using an unpaired two-tailed *t*-test. *$p < 0.05$.
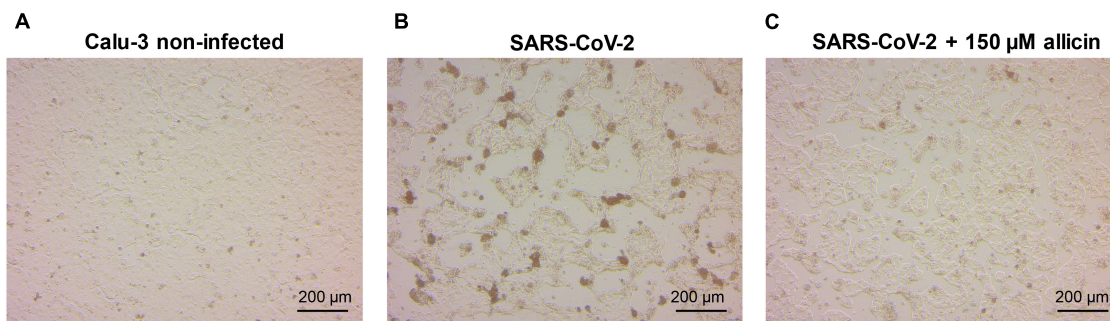
FIGURE 4 | Cellular effect of 150 μM allicin on SARS-CoV-2 infected Calu-3 cells. Non-infected Calu-3 cells were cultivated for 24 h and served as controls **(A)**. Calu-3 cells were infected with SARS-CoV-2 at a MOI of 0.01 and the SARS-CoV-2 induced cellular effects were studied at 24 h p.i. without **(B)** or with 150 μM allicin treatment **(C)** after infection as described in the "Materials and Methods" section. At 24 h p.i. infected Calu-3 cells showed cellular damages, including cell rounding, detachment and cell death **(B)**. Allicin treatment decreased some cellular damages **(C)**. Cells were imaged with a Nikon Ts2R-FL inverted microscope.

cell sizes, respectively (**Figures 5A–D** and **Supplementary Table 3**). A subset of the most strongly induced proteins in the Calu-3 proteome after SARS-CoV-2 infection is listed in **Table 1**.

The proteome after SARS-CoV-2 infection revealed altered expression of various cellular pathways, including the interferon-stimulated gene (ISG) signature, transcription, translation and protein degradation, the cytoskeleton, vesicular trafficking and tight junctions, apoptosis, signal transduction pathways as well as carbon, lipid and nucleotide metabolism (**Figure 5A**, **Table 1**, and **Supplementary Table 3**). In addition, the eight detected SARS-CoV-2 proteins were induced after 24 h p.i. of Calu-3 cells, with the ribonucleocapsid protein (N-protein) as one of the most abundant proteins in the proteome of infected Calu-3 cells with 0.35% of the total proteome. The N-protein was 29- and 21-fold higher expressed compared to the membrane protein (M-protein) (0.012%) and spike protein (S-protein) (0.016%), respectively (**Figure 6A** and **Supplementary Table 3**), confirming previous data with infected Vero E6 cells (Zecha et al., 2020). The viral proteins Nsp1, Nsp2, ORF3, ORF9b, and the papain-like

protease PLP were low abundant, contributing from 0.00022% (PLP) to 0.0043% (ORF3) to the total proteome (**Figure 6B** and **Supplementary Table 3**), while other viral proteins were not detected.

SARS coronaviruses have been shown to enter the cell via endocytosis and direct fusion with the cell membrane (Wang et al., 2008; Ou et al., 2020). In agreement with these reports, 18 proteins involved in vesicular transport and cytoskeleton regulation, such as formation of lysosomes, phagosomes, and exosomes were 1.5–5.2-fold higher expressed after infection in the Calu-3 proteome (**Figure 5A**, **Table 1**, and **Supplementary Table 3**). Among these proteins are the abundant and highly induced kinesins (KIF5A/B/C), clathrin (CLTCL1), and tubulin (TUBAL3), which are microtubule-associated proteins and participate in endocytosis and traffic of viral RNA and vesicles. The 1.7-fold induced highly abundant Rho GTPase-activating protein 17 (ARHGAP17) could be involved in the repair of tight junctions, which are often damaged in COVID-19 patients (De Maio et al., 2020; Tian et al., 2020).

About 21 proteins of the interferon (IFN) and ISG response were strongly induced, including sensors of viral RNA, the
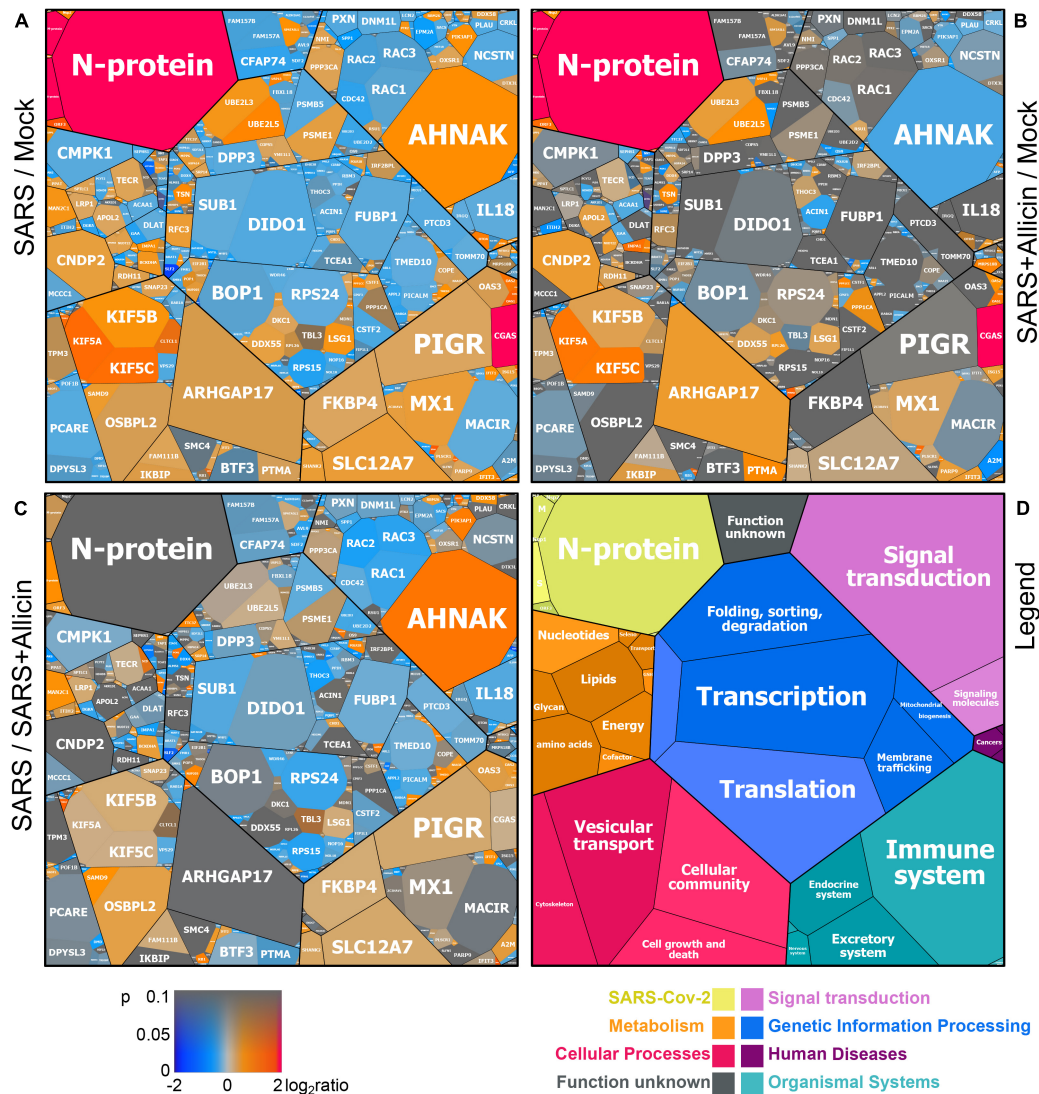
FIGURE 5 | The effect of allicin on the proteome of SARS-CoV-2 infected Calu-3 cells. The host-viral proteome treemaps **(A–C)** show the 536 differentially expressed proteins upon SARS-CoV-2 infection and were constructed by the Paver software (Mehlan et al., 2013). The treemaps visualize the following proteome changes: **(A)** SARS-CoV-2 infection/Mock, **(B)** SARS-CoV-2 infection + Allicin/Mock and **(C)** SARS-CoV-2 infection −/+ Allicin. The treemap **(D)** serves as legend for the functional KEGG categories displayed in different colors for level 1 and sublevel 2 as listed in **Supplementary Table 3**. The cell sizes in panels **(A–C)** denote the average abundances of 207 proteins with ≥1.5-fold inductions and 329 proteins with <0.66-fold decreased expression after SARS-CoV-2 infection. The log2 ratios of the proteins are shown by a red-blue color gradient (red – induction, blue – repression) **(A–C)**. p-values ($p < 0.05$; $p < 0.01$) were calculated using an unpaired two-tailed t-test from 3 to 4 biological replicates with 1–3 technical replicates (**Supplementary Table 3**).

JAK-STAT signal transduction pathway and antiviral effectors that interfere with the viral life cycle (**Figure 5A**, **Table 1**, and **Supplementary Table 3**) (Schneider et al., 2014). For a better understanding, the RNA sensing receptors, IFN and ISG signaling cascades and the previously described antiviral functions of the ISG effectors are displayed in a schematic (**Figures 7A,B**). The cyclic GMP-AMP (cGAMP) synthase (cGAS) was most strongly 98-fold upregulated upon infection, acting as sensor of viral RNA (**Table 1**, **Supplementary Table 3**, and **Figure 5A**) (Schneider et al., 2014). cGAMP activates the stimulator of interferon genes (STING) (**Figure 7A**).

The 2′-5′-oligoadenylate synthases (OAS1-3, OASL) were 1.6–7-fold induced upon infection to produce 2′-5′-adenylic acid as second messenger and activator of RNaseL for viral RNA degradation. The IFN-induced helicase C-domain-containing protein (IFIH) was 6.5–fold upregulated, which activates the mitochondrial antiviral signaling protein (MAVS) to induce the IFN response. Other IFN-induced effector proteins with tetratricopeptide repeats (IFIT1-3, IFIT5) were 1.6–3.5-fold induced after infection and function in RNA degradation and inhibition of translation. Further effectors are the Interferon-induced myxoma resistance protein 1

(MX1) and the Polymeric immunoglobulin receptor (PIGR), which represented 0.05 and 0.1% of the total proteome abundance and were 1.8 and 1.5-fold induced, respectively. MX1 is a dynamin-like GTPase, which forms ring-like structures and traps incoming ribonucleocapsids, thereby blocking uncoating and vesicular trafficking to direct them for degradation (**Figure 7B**) (Schneider et al., 2014). MX1 was also reported to be up-regulated in COVID-19 patients (Bizzotto et al., 2020).

Furthermore, the abundant cytokine IL18, the IL-1 receptor antagonist protein (IL1RN), the macrophage immunometabolism regulator MACIR and the Alpha-2-macroglobulin (A2M) were ∼0.6-fold lower expressed in infected cells. MACIR is implicated in the regulation of macrophages and autoimmune diseases (McGauran et al., 2020).

An important role in signal transduction and regulation of the antiviral response plays the abundant ISG15 effector, which was 1.8-fold induced after SARS-CoV-2 infection in the proteome. ISG15 functions amongst others as Ubiquitin-like modifier in ISGylation of RIG-I and IRF-3, which are targeted for degradation or activated to regulate IFN and ISG production (**Figure 7A**) (Masucci, 2020). Widespread ISGylation of newly synthesized viral proteins is proposed to inhibit viral replication and translation (Durfee et al., 2010).

Additionally, post-translational modification by polyubiquitination of host signaling factors, such as RIG-I, STING and MAVS is important for regulation of the IFN response upon SARS-CoV-2 infection (**Figure 7A**). Thus, several ubiquitin-conjugating E2 enzymes (UBE2L3 and UBE2L5), the E3 ubiquitin ligases (TRIM21, TRIM38, and ARIH2) and the ubiquitin specific protease or deconjugases (USP13) are 1.5–3.2-fold induced in the infected cells, while other E2, E3 enzymes and deconjugases (e.g., UBE2D2/3, RNF214, USP4, USP47, and USP48) are 0.2–0.62-fold lower expressed (**Table 1**, **Supplementary Table 3**, and **Figure 5A**). Since host and viral targets of ubiquitination and ISGylation are often directed to degradation, components of the proteasome, proteases, protein folding factors, and chaperones are 1.5–1.8-fold upregulated. The folding factors include the highly abundant peptidyl-prolyl *cis-trans* isomerase FKBP4, which functions as immunophilin and co-chaperone to interact with HSP90.

Apart from protein modification, the virus relies on protein synthesis and translation by the host machinery for its successful replication and infectivity. Accordingly, 24 proteins involved in translation were 1.5–2.8-fold upregulated under SARS-CoV-2 infection, including the translation factor EIF2B1, ribosomal proteins (RPL26, MRPS30, RRP8, PDCD11, and MRPL4), RNA helicases (DDX55 and DDX56), RNAses (POP1 and XRN1) and other regulatory factors, such as phosphatases (PPP1CC, PPP1CA, and PPP2R5A) (**Supplementary Table 3** and **Figure 5A**).

In addition, 16 proteins involved in transcription and the spliceosome were upregulated in infected Calu-3 cells, including the pre-mRNA splicing factors Slu7, PRPF40B, SCAF11, and the U1 small nuclear ribonucleoprotein C (SNRPC), which were 1.6–1.8-fold higher expressed. The transcription factors GABPA, ZNF579, SP110, and TSC22D2 were also induced

after infection. However, the majority of differentially expressed proteins involved in transcription (48) and translation (30) were repressed after SARS-CoV-2 infection, including the highly abundant proteins DIDO1, SUB1, FUBP1, TCEA1, BOP1, RPS24, and RPS15 (**Supplementary Table 3** and **Figure 5A**).

Moreover, virus replication and proliferation inside host cells requires reprogramming of the host metabolism, which was evident by the upregulation of 34 proteins and downregulation of 43 proteins involved mainly in lipid, energy, glycan, and nucleotide metabolism (**Table 1**, **Supplementary Table 3**, and **Figure 5A**). The induced proteins might function in the biosynthesis of the building blocks for viral phospholipid membranes, glycosylation of surface proteins and viral RNA genomes. Since the nucleotide pool is essential for coronavirus replication (Bojkova et al., 2020), some purine and pyrimidine biosynthesis proteins were 1.7–2.3-fold induced (NT5C2, UPP1, and PPAT), while others were 0.5-0.65-fold repressed (CMPK1, AK6, and ENPP4) (**Supplementary Table 3** and **Figure 5A**).
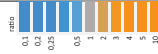
Furthermore, expression of several signaling pathways, including JAK-STAT, MAPK, Wnt, Ras, and Rap1 signaling were affected by SARS-CoV-2 infection. The JAK-STAT pathway senses and transduces IFN-signals via a phosphorylation cascade to activate ISG expression (**Figure 7A**). Thus, STAT2, N-myc interactor NMI and the RIG-I receptor were 1.6–1.8-fold induced upon infection (**Supplementary Table 3** and **Figure 5A**). Proteins of the MAPK signaling pathways were activated in response to infections with SARS-CoV (Bouhaddou et al., 2020) and 1.6–1.8-fold induced in the proteome of SARS-CoV-2 infected cells. Proteins of the PI3K/Akt signaling pathway were 2.3–2.6-fold upregulated in infected cells, controlling apoptosis of host cells for successful viral replication. The highly abundant neuroblast differentiation-associated protein AHNAK was 2.6-fold induced after SARS-CoV-2 infection. AHNAK is required for calcium signaling and might regulate the immune response (Matza et al., 2009). Proteins of the Ras-signaling pathway were 0.5–0.64-fold downregulated upon virus infection, including three Rac GTPases Rac1-3 that are implicated in the regulation of cell morphology, migration and invasion, by transducing signals from cell surface receptors to the actin and microtubule cytoskeletons (Wheeler et al., 2006). Similarly, other proteins involved in the cytoskeleton organization were 0.3–0.66-fold lower expressed, indicating re-organization of the cytoskeleton for transport of virus particles.

Allicin Leads to a Decreased Antiviral Interferon Response in the Proteome of Infected Calu-3 Cells

Next, we investigated the effect of allicin on the proteome changes upon SARS-CoV-2 infection. Quantification of the 8 viral proteins in infected Calu-3 cells after allicin treatment revealed a significantly 18–59% decreased abundance of the structural proteins N, M, and S and ORF3, supporting the antiviral effect of allicin in the proteome (**Figures 6A,B**).

Allicin treatment resulted in a diminished IFN-response in infected cells, since expression of innate immune receptors

TABLE 1 | Selected most strongly induced proteins in the Calu-3 proteome after SARS-CoV-2 infection and the effect of allicin treatment p.i.

Protein informations			Expression ratios			log2 Expression ratios						T-Test p-value					
Protein name	Accession number	Protein function	SARS/ Mock	SARS + All/ Mock	SARS/ SARS + All	SARS/ Mock	SARS + All/ Mock	SARS/ SARS + All	SARS/ Mock	SARS + All/ Mock	SARS/ SARS + All	T-Test SARS/ Mock	T-Test SARS + All/ Mock	T-Test SARS + All/ SARS	SARS/ Mock	SARS + All/ Mock	SARS/SARS + All
Protein folding and ubiquitination																	
UBE2L5	A0A1B0GUS4	Ubiquitin-conjugating enzyme E2 L5	3.02	2.45	1.23	1.59	1.29	0.30				0.000	0.000	0.036			
UBE2L3	P68036	Ubiquitin-conjugating enzyme E2 L3	2.06	1.79	1.15	1.04	0.84	0.20				0.000	0.000	0.139			
FKBP4	Q02790	Peptidyl-prolyl *cis-trans* isomerase FKBP4	1.51	1.10	1.37	0.59	0.14	0.45				0.003	0.550	0.032			
Tight junctions, endocytosis and phagosome																	
ARHGAP17	Q68EM7	Rho GTPase-activating protein 17	1.66	1.88	0.88	0.73	0.91	−0.18				0.092	0.028	0.462			
KIF5A	Q12840	Kinesin heavy chain isoform 5A	5.24	4.13	1.27	2.39	2.05	0.34				0.000	0.000	0.016			
KIF5C	O60282	Kinesin heavy chain isoform 5C	4.97	3.91	1.27	2.31	1.97	0.34				0.000	0.000	0.016			
KIF5B	P33176	Kinesin-1 heavy chain	2.20	1.61	1.36	1.14	0.69	0.45				0.001	0.009	0.036			
TUBAL3	A6NHL2	Tubulin alpha chain-like 3	5.24	0.73	7.14	2.39	−0.45	2.84				0.025	0.251	0.019			
Calcium signaling																	
AHNAK	Q09666	Neuroblast differentiation-associated protein	2.55	0.60	4.28	1.35	−0.74	2.10				0.012	0.055	0.004			
O- and N-glycan biosynthesis and degradation																	
GALNT4	Q8N4A0	Polypeptide N-acetylgalactosaminyltransferase 4	2.10	0.97	2.16	1.07	−0.05	1.11				0.003	0.801	0.002			
GALNT12	Q8IXK2	Polypeptide N-acetylgalactosaminyltransferase 12	1.67	1.00	1.66	0.74	0.01	0.73				0.003	0.215	0.020			
ALG3	Q92685	Dol-P-Man:Man(5)GlcNAc(2)-PP-Dol α-1,3-mannosyltransferase	1.67	0.85	1.97	0.74	−0.24	0.98				0.064	0.481	0.024			
FUT8	Q9BYC5	Alpha-(1,6)-fucosyltransferase	1.62	0.90	1.79	0.69	−0.15	0.84				0.027	0.704	0.015			
MAN2C1	Q9NTJ4	Alpha-mannosidase 2C1	2.08	0.92	2.25	1.06	−0.12	1.17				0.039	0.761	0.026			
RNA sensing, interferon (IFN) and ISG effectors																	
CGAS	Q8N884	Cyclic GMP-AMP synthase	97.90	80.16	1.22	6.61	6.32	0.29				0.000	0.000	0.028			
OAS1	P00973	2′–5′-oligoadenylate synthase 1	6.97	4.26	1.64	2.80	2.09	0.71				0.000	0.003	0.031			
OAS2	P29728	2′–5′-oligoadenylate synthase 2	4.32	2.31	1.87	2.11	1.21	0.90				0.000	0.038	0.010			
OAS3	Q9Y6K5	2′–5′-oligoadenylate synthase 3	1.61	0.95	1.70	0.69	−0.08	0.77				0.091	0.843	0.073			
OASL	Q15646	2′–5′-oligoadenylate synthase-like protein	6.13	3.49	1.76	2.62	1.80	0.81				0.000	0.014	0.015			
MX1	P20591	Interferon-induced GTP-binding protein	1.80	1.50	1.20	0.85	0.59	0.26				0.000	0.063	0.251			
IFI16	Q16666	Gamma-interferon-inducible protein 16	4.55	1.77	2.57	2.18	0.82	1.36				0.019	0.345	0.086			
IFI44	Q8TCB0	IFN-induced protein 44	8.00	5.56	1.44	3.00	2.48	0.52				0.000	0.009	0.145			
IFI44L	Q53G44	IFN-induced protein 44-like	5.77	3.19	1.81	2.53	1.67	0.85				0.000	0.037	0.033			
IFIH1	Q9BYX4	IFN-induced helicase C domain protein 1	6.59	2.28	2.89	2.72	1.19	1.53				0.011	0.139	0.029			
IFIT1	P09914	IFN-induced protein with tetratricopeptide repeats 1	2.53	1.05	2.41	1.34	0.07	1.27				0.003	0.237	0.008			
IFIT2	P09913	IFN-induced protein with tetratricopeptide repeats 2	1.69	1.02	1.66	0.76	0.03	0.73				0.026	0.298	0.041			
IFIT3	O14879	IFN-induced protein with tetratricopeptide repeats 3	2.33	1.85	1.26	1.22	0.89	0.33				0.002	0.004	0.189			
IFIT5	Q13325	IFN-induced protein with tetratricopeptide repeats 5	3.54	2.46	1.44	1.83	1.30	0.53				0.000	0.017	0.040			
ISG15	P05161	Ubiquitin-like protein ISG15	1.83	1.81	1.01	0.87	0.86	0.01				0.000	0.001	0.934			
PARP14	Q460N5	Protein mono-ADP-ribosyltransferase PARP14	3.59	0.42	8.54	1.84	−1.25	3.09				0.009	0.066	0.003			
PIGR	P01833	Polymeric immunoglobulin receptor	1.51	1.08	1.40	0.59	0.11	0.48				0.001	0.368	0.002			
SARS-CoV-2 proteins																	
S-protein	spike	protein surface glycoprotein SARS-CoV-2	440.36	180.28	2.44	8.78	7.49	1.29				0.002	0.000	0.033			
N-protein	Nucleocapsid	phosphoprotein SARS-CoV-2	130.33	106.28	1.23	7.03	6.73	0.29				0.000	0.000	0.684			
M-protein	Membrane	protein glycoprotein SARS-CoV-2	99.01	60.94	1.62	6.63	5.93	0.70				0.001	0.000	0.150			
ORF3	ORF3	structural protein SARS-CoV-2	9.80	5.18	1.89	3.29	2.37	0.92				0.000	0.000	0.036			
ORF9b	ORF9b	structural protein SARS-CoV-2	8.77	9.41	0.93	3.13	3.23	−0.10				0.001	0.000	0.251			
Nsp1	Nsp1	Non-structural protein1 SARS-CoV-2	6.54	7.09	0.92	2.71	2.83	−0.12				0.003	0.001	0.451			

*The proteome samples of Calu-3 Mock cells (Mock), SARS-CoV-2 infected Calu-3 cells (SARS) and SARS-CoV-2 infected Calu-3 cells treated with 150 μM allicin p.i. (SARS + All) were harvested after 24 h p.i. and separated by non-reducing SDS PAGE for prefractionation. Protein fractions were tryptic in-gel digested and peptides analyzed by Orbitrap Fusion LC-MS/MS analysis as described in the "Materials and Methods". The table lists 37 out of 207 identified proteins with >1.5-fold induction upon SARS-CoV-2 infection. These proteins are most strongly induced after SARS-CoV-2 infection, affected by allicin treatment and/or are present at high abundance in the Calu-3 proteome. The proteins were classified according to their KEGG ontologies and UniprotKB annotations. The full set of up- and downregulated proteins after SARS-CoV-2 infection is listed in **Supplementary Table 3**. The table includes protein names, accession numbers, protein functions, expression ratios, log2 expression ratios and p-values. The log2 ratios and p-values are visualized with a blue-orange and red-gray color code, respectively. p-values were calculated using an unpaired two-tailed t-test for two samples with unequal variance.*

FIGURE 6 | The effect of allicin on the abundance of SARS-CoV-2 proteins in the proteome of infected Calu-3 cells. **(A,B)** The abundance of the SARS-CoV-2 proteins (N, S, M, ORF3, Nsp2, ORF9b, Nsp1, and PLP) at 24 h p.i. relative to the total proteome abundance of infected Calu-3 cells was calculated in the absence (SARS) or presence of 150 μM allicin exposure p.i. (SARS + allicin). The N-protein showed a 21–29-fold higher abundance compared to the structural proteins S and M as shown in panel **(A)**. The lower abundant S, M, ORF3, Nsp2, ORF9b, Nsp1, and PLP proteins are separately displayed in panel **(B)**. The proteome results were obtained from 3–4 biological replicates with 1–3 technical replicates (**Supplementary Table 3**). Error bars represent the SD. p-values were calculated using an unpaired two-tailed t-test. $^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

FIGURE 7 | Schematic of viral RNA recognition, activation of the IFN and ISG signaling pathways **(A)** and antiviral functions of the identified ISG effectors **(B)** in this host-virus proteome study. **(A)** SARS-CoV-2 enters host cells via endocytosis. RIG-I is a cytosolic receptor to recognize viral RNA. cGAS and OAS are ISG effectors that function as RNA sensors. RIG-I, IFIH, and cGAS activate the mitochondrial antiviral-signaling protein (MAVS) and stimulator of IFN genes (STING), leading to phosphorylation of IFN responsive factors (e.g., IRF3), followed by IRF3 dimerization, translocation into the nucleus and transcriptional activation of IFN expression. IRF3 is negatively regulated by IFI44. RIG-I, MAVS, and STING can be regulated by ubiquitination (UBE2) and ISGylation (ISG15). Type-I IFN a/ß bind to the IFNAR receptor, resulting in phosphorylation of signal transducers and activators of transcription (STAT1/2) by the JAK and TYK kinases. Phosphorylated STAT1/2 form dimers and bind to IRF9, which triggers transcription of IFN-stimulated genes (ISGs) in the nucleus. **(B)** Antiviral ISG effectors affect different stages of the viral life cycle. Mx1 inhibits virus endocytosis and uncoating of the ribonucleocapsid. IFI16, OAS, and IFIT function in viral RNA degradation and block translation. IFIT, ISG15, TRIM, and UBE2 inhibit transcription, replication, translation or virus assembly. FKBP4 promotes protein folding. Kinesins (KIFA/B/C), Clathrin (CLTCL1), and TUBAL3 are involved in transport of virus vesicles. ARHGAP17 facilitates the formation or repair of tight junctions. The figure is adapted from reference (Schneider et al., 2014).

and ISG effectors of the JAK-STAT signaling pathways were decreased, including FKBP4, PIGR, MX1, cGAS, OAS1-3, and IFIT1-3 (**Table 1**, **Supplementary Table 3**, and **Figures 5B,C**). In addition, proteins involved in ubiquitination (UBE2L3/5) and the JAK-STAT, MAPK, PI3K/Akt, and Ras signaling pathways

showed lower expression changes after allicin treatment. The abundant calcium-signaling protein AHNAK was repressed after allicin exposure, while it was induced in infected cells. Allicin resulted in decreased expression of kinesins KIFA/B/C, clathrin CLTCL1 and tubulin TUBAL3, indicating reduced endocytosis

and traffic of vesicles. Moreover, prothymosin alpha (PTMA) was 2.6-fold upregulated after allicin exposure. PTMA showed antiviral activity to inhibit replication of human deficiency virus type-1 (Mosoian et al., 2006; Teixeira et al., 2015), indicating that PTMA induction by allicin might contribute to the antiviral effect against SARS-CoV-2.

Similarly, the expression of proteins involved in transcription, spliceosome and translation was reversed to the levels of uninfected (Mock) cells after allicin exposure of SARS-CoV-2 infected cells, including the abundant proteins DIDO1, SUB1, FUBP1, TCEA1, BOP1, RPS24 and RPS15. Finally, expression of metabolic enzymes involved in glycan, nucleotide, and lipid metabolism was restored by allicin, including GALNT4/12, ALG3, FUT8, MAN2C1, CMPK1, and TECR (**Table 1**, **Supplementary Table 3**, and **Figures 5B,C**). Overall, allicin showed antiviral effects in the host proteome as revealed by the diminished IFN-dependent antiviral response and the effects on signal transduction, transcription, translation, and metabolism.

DISCUSSION

Garlic organosulfur compounds showed antiviral activity against several enveloped viruses, including herpes simplex, parainfluenza, vaccinia, and rhinovirus (Weber et al., 1992; Rouf et al., 2020). These virucidal effects of garlic compounds were proposed to depend on the disruption of the viral envelope and inhibition of viral replication (Weber et al., 1992; Rouf et al., 2020).

In this work, we explored the antiviral effect of allicin on SARS-CoV-2 infected Vero E6 and Calu-3 cells. By determining >60–70% decreased levels of viral RNA and infectious viral particles, the antiviral effect of biocompatible allicin doses against SARS-CoV-2 was demonstrated in both cell lines. However, Calu-3 cells showed a greater allicin tolerance compared to the more sensitive Vero E6 cells. Different cell lines were previously shown to vary in their allicin susceptibilities, which correlated with their intracellular GSH levels (Gruhlke et al., 2016, 2019). Allicin leads to *S*-thioallylation of GSH and the formation of *S*-allylmercaptoglutathione (GSSA), which is accompanied by GSH depletion and an oxidative shift in the GSH redox potential (Gruhlke et al., 2010, 2019; Müller et al., 2016). The measurement of GSH levels confirmed that Calu-3 cells have 4.2-fold higher GSH levels compared to the allicin-sensitive Vero E6 cells.

Since Vero E6 cells are more sensitive toward allicin compared to Calu-3 cells, we first analyzed the effect of low levels of 50 μM allicin on SARS-CoV-2 infected Vero E6 cells using pre- and post-infection treatments of host cells or pre-exposure of the virus before infection. Interestingly, treatment of Vero E6 cells with 50 μM allicin showed only antiviral effects if applied after infection with SARS-CoV-2. The number of infectious viral particles was not affected if virus or Vero E6 host cells were pretreated with 50 μM allicin for 30 min before infection. This missing antiviral effect of allicin in the pre-treatment experiments might be explained by the washing step of Vero E6 cells after the 1 h infection, leading to the removal of the remaining allicin from the cells. During post-infection treatment, allicin remained in the cell culture until sample harvesting. These results suggest that allicin might disrupt early steps of viral replication in Vero E6 cells, which requires further investigation. Further experiments revealed the significant reduction in viral plaques and RNA copies in both SARS-CoV-2 infected Vero E6 and Calu-3 cell lines after exposure to sub-lethal doses allicin p.i.

To better understand the effect of allicin, we investigated the proteome changes of Calu-3 cells upon SARS-CoV-2 infection and the impact of 150 μM allicin on the host-virus proteome. In agreement with previous proteome studies, SARS-CoV-2 reprograms major host pathways, including signaling pathways, transcription, splicing, translation, protein modification and folding, lipid, glycan, and nucleotide metabolism (**Table 1**, **Supplementary Table 3**, and **Figures 5A, 7A,B**) (Bojkova et al., 2020; Bouhaddou et al., 2020; Zecha et al., 2020). The ribonucleocapsid protein was the most abundant viral protein in the infected cell, indicating that a large portion of the translation capacity goes to the N-protein for package of the viral RNA genome.

In addition, our proteome data highlight the importance of the IFN pathway and ISG effectors to prevent virus replication by interacting with various stages of the viral life cycle. Antiviral ISG effectors were among the most highly induced and abundant proteins in the infected host cells, such as MX1, cGAS, OAS1-3, IFIT1-3, ISG15, FKBP4, PIGR, and UBE2L3/5, which function in sensing and degradation of viral RNA, inhibition of ribonucleocapsid uncoating, translation and promote the innate immune response (**Table 1**, **Supplementary Table 3**, and **Figures 5A**, **7A,B**). Apart from IFN signaling, proteins involved in motility, tight junction and membrane trafficking are highly induced host proteins, supporting the importance of vesicular transport for virus endocytosis and exocytosis. Thus, our proteomics studies reflect all described host pathways known to be altered after viral infections, suggesting new host targets for SARS-CoV-2 interventions.

At the same time, the proteomic profiling gave the opportunity to monitor the responses of infected Calu-3 cells after allicin treatment. The proteome results of allicin-treated infected host cells revealed an 18–59% reduced abundance of the structural proteins N, M, S, and ORF3. Several expression changes dedicated to virus proliferation are reversed to Mock levels in allicin-treated cells. Allicin affected virus-responsive expression of JAK-STAT, MAPK, PI3K/Akt, and Ras signaling pathways, IFN and ISG effectors, transcription, splicing, translation, ubiquitination, vesicular transport, tight junctions as well as glycan, lipid, and nucleotide metabolism. Thus, our results confirm the antiviral effect of allicin in host cells in agreement with the infection assays.

The mode of action of allicin involves *S*-thioallylation of proteins and low molecular weight thiols in bacteria and human Jurkat cells, which was observed already 10–30 min after allicin treatment (Gruhlke et al., 2019; Loi et al., 2019). The majority of *S*-thioallylated Jurkat proteins were abundant cellular proteins, involved in the cytoskeleton, translation and protein folding, although also low abundant redox-sensitive transcription factors, such as MgrA, SarZ, OhrR, HypR, and YodB were targets for

allicin modification in *S. aureus* and *B. subtilis* cells (Chi et al., 2019; Gruhlke et al., 2019; Loi et al., 2019). In this study, we did not detect *S*-thioallylated Cys peptides in SARS-CoV-2 infected Calu-3 cells after 24 h of allicin exposure using label-free proteomics and MS/MS spectrum verification. In addition, no viral *S*-thioallylated Cys peptides were identified, although the spike protein is a Cys-rich glycoprotein exposed on the surface of the virus envelope (Ou et al., 2020). Given the long duration of allicin treatment, this is not surprising since cells have the capacity to reduce allicin and the majority of *S*-thioallylations within 24 h. In human Jurkat cells, allicin led to a rapid depletion of the cellular GSH pool within 10 min (Gruhlke et al., 2019). Efficient allicin detoxification and removal of *S*-thioallylations were confirmed in yeast and bacterial cells, as supported by fast recovery of growth after a short allicin-induced lag phase (Gruhlke et al., 2010, 2019; Müller et al., 2016; Loi et al., 2019). In *S. aureus*, allicin can be reduced by the disulfide reductase MerA, while the bacillithiol disulfide reductase YpdA enables the recycling of *S*-allylmercaptobacillithiol, formed in the reaction of BSH with allicin (Loi et al., 2019). Furthermore, YpdA was shown to function in the Brx/BSH/YpdA pathway in regeneration of *S*-thioallylated proteins (Loi et al., 2019). Similarly, Calu-3 cells should have reduced GSSA via the glutathione disulfide reductase and *S*-thioallylated proteins by the glutaredoxin/GSH/glutathione disulfide reductase system within 24 h.

Finally, the question arises about the antiviral mechanism of allicin on SARS-CoV-2 infected Calu-3 cells during host-pathogen interactions. Garlic organosulfur compounds were shown to exert their immunomodulatory activity via inhibition of the transcription factor NF-κB, leading to decreased levels of pro-inflammatory cytokines, such as TNF-α, IL-1β, IL-6, MCP-1, and IL-12 (Arreola et al., 2015). Allicin further stimulates the release of Zn^{2+} from proteins in murine EL-4 T-cells, possibly due to *S*-thioallylation of Zn^{2+} coordinating Cys thiolates (Gruhlke et al., 2019). Thus, the immunomodulatory effect of allicin on cytokine secretion in cell cultures could be mediated by elevated Zn^{2+} levels due to inactivation of host proteins by *S*-thioallylation.

On the other hand, allicin could also target Cys-containing virus proteins, such as the Cys-rich spike glycoprotein, the viral RNA-dependent RNA polymerase RdRp (Nsp12), the main protease M^{pro} (also termed as 3C-like protease) and the papain-like protease PLP. M^{pro} and PLP are both involved in proteolytic processing of the large pp1a and pp1ab polyproteins to produce functional polypeptides, which assemble into the replicase-transcriptase complex (Jin et al., 2020; Zhang et al., 2020; Amin et al., 2021). Since M^{pro} and RdRp are important for viral replication and transcription, these could be antiviral drug targets. M^{pro} has a catalytic active site motif consisting of His41 and Cys145 residues (Jin et al., 2020; Zhang et al., 2020; Amin et al., 2021). Several *in silico* docking studies with allicin revealed the formation of *S*-thioallylations at Cys145, Cys85 and Cys156 of M^{pro} and of Cys622 of RdRp, indicating the potential of allicin to attenuate SARS-CoV-2 replication (Bastikar et al., 2020; Shekh et al., 2020). Further docking studies with the garlic compounds alliin and ajoene revealed strong

ligand-protein binding stabilities and many interactions at the M^{pro} active site (Bastikar et al., 2020; Cheng and Li, 2020). In total, 17 garlic organosulfur compounds, accounting for 99.4% of substances found in garlic oil, showed interactions with ACE2 receptor and M^{pro} *in silico*, including the diallyl di- and trisulfides with promising docking scores (Thuy et al., 2020). Furthermore, elevated Zn^{2+} levels released from host proteins inhibited the RdRp of SARS coronavirus (te Velthuis et al., 2010) and the host ACE2 enzyme (Polak and Speth, 2021), indicating that allicin could target either host and viral proteins directly via *S*-thioallylation or via elevated Zn^{2+} levels to exert immunomodulatory and antiviral effects.

Thus, allicin might inhibit SARS-CoV-2 infection at different stages of the viral life cycle, preventing receptor binding, replication or transcription by *S*-thioallylation of host or viral proteins. However, in our host-virus proteome we could not identify the M^{pro} or RdRp proteins, indicating that these are low abundant proteins. Thus, possible *S*-thioallylations of M^{pro} and RdRp will be difficult to verify upon allicin exposure of infected host cells *in vivo*. Nevertheless, these *in silico* docking studies highlight the potential of garlic organosulfur compounds as inhibitors of viral Cys proteins, which could be further developed as possible future COVID-19 therapeutics.

Taken together, our results demonstrate that allicin shows antiviral and immunomodulatory activity in SARS-CoV-2 infected Vero E6 and Calu-3 cell cultures, supported on the proteome level by the decreased antiviral interferon response. However, allicin is unstable and quickly decomposes to polysulfanes, ajoene, and other sulfur compounds during heating (Block, 2010; Borlinghaus et al., 2021). The half-life of allicin is 30–40 days in water at 23°C, but decreases in garlic extracts with increased concentrations (Koch and Lawson, 1996). In the acidic stomach, the majority of allicin is degraded to 2-propenethiol and allyl methyl sulfide, which are excreted (Block, 2010; Borlinghaus et al., 2021). In the blood, the effective dose of allicin is reduced by its reaction with GSH (Block, 2010; Borlinghaus et al., 2021). In our experiments, the viral load was only reduced by 60–70% after allicin treatment in Calu-3 cells, which is below 1-log scale and would not satisfy the desired antiviral effect required for therapeutics to enter pre-clinical trials. Future drug research should be directed to exploit the thiol-reactive activity of allicin derivatives with reduced toxicity, increased stability and higher antiviral activity as antiviral lead compounds.

DATA AVAILABILITY STATEMENT

The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Perez-Riverol et al., 2019; Deutsch et al., 2020) with the dataset identifier PXD024375.

AUTHOR CONTRIBUTIONS

KM performed the infection experiments and analyzed the data. VF and LA measured and analyzed the proteomic data. JB

constructed the Voronoi treemaps. MG and AS synthesized allicin. DN supervised the infection experiments and gave critical advise. VF and HA conducted the study and wrote the initial manuscript. All authors contributed to the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.746795/full#supplementary-material

REFERENCES

Al-Qattan, M. M. (2009). Garlic burns: case reports with an emphasis on associated and underlying pathology. *Burns* 35, 300–302. doi: 10.1016/j.burns.2008.01.004

Amin, S. A., Banerjee, S., Singh, S., Qureshi, I. A., Gayen, S., and Jha, T. (2021). First structure-activity relationship analysis of SARS-CoV-2 virus main protease (Mpro) inhibitors: an endeavor on COVID-19 drug discovery. *Mol. Divers* 25, 1827–1838. doi: 10.1007/s11030-020-10166-3

Anderson, T. W., and Darling, D. A. (1954). A test of goodness of fit. *J. Am. Statist. Assoc.* 49, 765–769. doi: 10.1080/01621459.1954.10501232

Arbach, M., Santana, T. M., Moxham, H., Tinson, R., Anwar, A., Groom, M., et al. (2019). Antimicrobial garlic-derived diallyl polysulfanes: interactions with biological thiols in Bacillus subtilis. *Biochim Biophys. Acta Gen. Subj.* 1863, 1050–1058. doi: 10.1016/j.bbagen.2019.03.012

Arreola, R., Quintero-Fabian, S., Lopez-Roa, R. I., Flores-Gutierrez, E. O., Reyes-Grajeda, J. P., Carrera-Quintanar, L., et al. (2015). Immunomodulation and anti-inflammatory effects of garlic compounds. *J. Immunol. Res.* 2015:401630. doi: 10.1155/2015/401630

Bastikar, V. A., Bastikar, A. V., and Chhajed, S. S. (2020). Understanding the role of natural medicinal compounds such as curcumin and allicin against SARS-CoV-2 proteins as potential treatment against COVID-19: an in silico approach. *Proteom. Bioinform* 13:7.

Bautista, D. M., Movahed, P., Hinman, A., Axelsson, H. E., Sterner, O., Hogestatt, E. D., et al. (2005). Pungent products from garlic activate the sensory ion channel TRPA1. *Proc. Natl. Acad. Sci. U S A.* 102, 12248–12252. doi: 10.1073/pnas.0505356102

Bizzotto, J., Sanchis, P., Abbate, M., Lage-Vickers, S., Lavignolle, R., Toro, A., et al. (2020). SARS-CoV-2 infection boosts MX1 antiviral effector in COVID-19 patients. *iScience* 23:101585. doi: 10.1016/j.isci.2020.101585

Block, E. (2010). *Garlic and Other Alliums-the Lore and the Science*. Cambridge: RSC Publishing.

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., et al. (2020). Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* 583, 469–472. doi: 10.1038/s41586-020-2332-7

Borlinghaus, J., Albrecht, F., Gruhlke, M. C., Nwachukwu, I. D., and Slusarenko, A. J. (2014). Allicin: chemistry and biological properties. *Molecules* 19, 12591–12618. doi: 10.3390/molecules190812591

Borlinghaus, J., Foerster, J., Kappler, U., Antelmann, H., Noll, U., Gruhlke, M., et al. (2021). Allicin, the odor of freshly crushed garlic: a review of recent progress in understanding allicin's effects on cells. *Molecules* 26:1505. doi: 10.3390/molecules26061505

Bouhaddou, M., Memon, D., Meyer, B., White, K. M., Rezelj, V. V., Correa Marrero, M., et al. (2020). The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* 182, 685–712.e19. doi: 10.1016/j.cell.2020.06.034

Cheng, B., and Li, T. (2020). Discovery of alliin as a putative inhibitor of the main protease of SARS-CoV-2 by molecular docking. *Biotechniques* 69, 108–112. doi: 10.2144/btn-2020-0038

Chi, B. K., Huyen, N. T. T., Loi, V. V., Gruhlke, M. C. H., Schaffer, M., Mäder, U., et al. (2019). The disulfide stress response and protein S-thioallylation caused by allicin and diallyl polysulfanes in Bacillus subtilis as revealed by transcriptomics and proteomics. *Antioxidants (Basel)* 8:605. doi: 10.3390/antiox8120605

Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., et al. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro. Surveill.* 25:2000045. doi: 10.2807/1560-7917.ES.2020.25.3.2000045

De Maio, F., Lo Cascio, E., Babini, G., Sali, M., Della Longa, S., Tilocca, B., et al. (2020). Improved binding of SARS-CoV-2 Envelope protein to tight junction-associated PALS1 could play a key role in COVID-19 pathogenesis. *Microbes Infect.* 22, 592–597. doi: 10.1016/j.micinf.2020.08.006

Deutsch, E. W., Bandeira, N., Sharma, V., Perez-Riverol, Y., Carver, J. J., Kundu, D. J., et al. (2020). The ProteomeXchange consortium in 2020: enabling 'big data' approaches in proteomics. *Nucleic Acids Res.* 48, D1145–D1152. doi: 10.1093/nar/gkz984

Durfee, L. A., Lyon, N., Seo, K., and Huibregtse, J. M. (2010). The ISG15 conjugation system broadly targets newly synthesized proteins: implications for the antiviral function of ISG15. *Mol. Cell* 38, 722–732. doi: 10.1016/j.molcel.2010.05.002

Fara, A., Mitrev, Z., Rosalia, R. A., and Assas, B. M. (2020). Cytokine storm and COVID-19: a chronicle of pro-inflammatory cytokines. *Open Biol.* 10:200160. doi: 10.1098/rsob.200160

Gruhlke, M. C., Hemmis, B., Noll, U., Wagner, R., Luhring, H., and Slusarenko, A. J. (2015). The defense substance allicin from garlic permeabilizes membranes of *Beta vulgaris*, *Rhoeo discolor*, *Chara corallina* and artificial lipid bilayers. *Biochim. Biophys. Acta* 1850, 602–611. doi: 10.1016/j.bbagen.2014.11.020

Gruhlke, M. C., Nicco, C., Batteux, F., and Slusarenko, A. J. (2016). The effects of allicin, a reactive sulfur species from garlic, on a selection of mammalian cell lines. *Antioxidants (Basel)* 6:1. doi: 10.3390/antiox6010001

Gruhlke, M. C., Portz, D., Stitz, M., Anwar, A., Schneider, T., Jacob, C., et al. (2010). Allicin disrupts the cell's electrochemical potential and induces apoptosis in yeast. *Free Radic. Biol. Med.* 49, 1916–1924. doi: 10.1016/j.freeradbiomed.2010.09.019

Gruhlke, M. C. H., Antelmann, H., Bernhardt, J., Kloubert, V., Rink, L., and Slusarenko, A. J. (2019). The human allicin-proteome: S-thioallylation of proteins by the garlic defence substance allicin and its biological effects. *Free Radic. Biol. Med.* 131, 144–153. doi: 10.1016/j.freeradbiomed.2018.11.022

Hitl, M., Kladar, N., Gavaric, N., Srdenovic Conic, B., and Bozin, B. (2021). Garlic burn injuries- a systematic review of reported cases. *Am. J. Emerg. Med.* 44, 5–10. doi: 10.1016/j.ajem.2021.01.039

Jarque, C. M., and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Econ. Lett.* 6, 255–259. doi: 10.1016/0165-1765(80)90024-5

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., et al. (2020). Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582, 289–293. doi: 10.1038/s41586-020-2223-y

Koch, H. P., and Lawson, L. D. (1996). *Garlic: The Science and Therapeutic Application of Allium sativum L. and Related Species*. Baltimore, MD: Williams & Wilkins.

Kublik, A., Deobald, D., Hartwig, S., Schiffmann, C. L., Andrades, A., Von Bergen, M., et al. (2016). Identification of a multi-protein reductive dehalogenase complex in Dehalococcoides mccartyi strain CBDB1 suggests a protein-dependent respiratory electron transport chain obviating quinone involvement. *Environ. Microbiol.* 18, 3044–3056. doi: 10.1111/1462-2920. 13200

Loi, V. V., Huyen, N. T. T., Busche, T., Tung, Q. N., Gruhlke, M. C. H., Kalinowski, J., et al. (2019). Staphylococcus aureus responds to allicin by global S-thioallylation - role of the Brx/BSH/YpdA pathway and the disulfide reductase MerA to overcome allicin stress. *Free Radic. Biol. Med.* 139, 55–69. doi: 10.1016/j.freeradbiomed.2019.05.018

Masucci, M. G. (2020). Viral ubiquitin and ubiquitin-like deconjugases-swiss army knives for infection. *Biomolecules* 10:1137. doi: 10.3390/biom10 081137

Matza, D., Badou, A., Jha, M. K., Willinger, T., Antov, A., Sanjabi, S., et al. (2009). Requirement for AHNAK1-mediated calcium signaling during T lymphocyte cytolysis. *Proc. Natl. Acad. Sci. U S A.* 106, 9785–9790. doi: 10.1073/pnas. 0902844106

McGauran, G., Dorris, E., Borza, R., Morgan, N., Shields, D. C., Matallanas, D., et al. (2020). Resolving the interactome of the human macrophage immunometabolism regulator (MACIR) with enhanced membrane protein preparation and affinity proteomics. *Proteomics* 20:e2000062. doi: 10.1002/ pmic.202000062

Mehlan, H., Schmidt, F., Weiss, S., Schuler, J., Fuchs, S., Riedel, K., et al. (2013). Data visualization in environmental proteomics. *Proteomics* 13, 2805–2821. doi: 10.1002/pmic.201300167

Miron, T., Listowsky, I., and Wilchek, M. (2010). Reaction mechanisms of allicin and allyl-mixed disulfides with proteins and small thiol molecules. *Eur. J. Med. Chem.* 45, 1912–1918. doi: 10.1016/j.ejmech.2010.01.031

Miron, T., Rabinkov, A., Mirelman, D., Wilchek, M., and Weiner, L. (2000). The mode of action of allicin: its ready permeability through phospholipid membranes may contribute to its biological activity. *Biochim. Biophys. Acta* 1463, 20–30. doi: 10.1016/S0005-2736(99)00174-1

Mosoian, A., Teixeira, A., High, A. A., Christian, R. E., Hunt, D. F., Shabanowitz, J., et al. (2006). Novel function of prothymosin alpha as a potent inhibitor of human immunodeficiency virus type 1 gene expression in primary macrophages. *J. Virol.* 80, 9200–9206. doi: 10.1128/JVI.00589-06

Müller, A., Eller, J., Albrecht, F., Prochnow, P., Kuhlmann, K., Bandow, J. E., et al. (2016). Allicin induces thiol stress in bacteria through S-allylmercapto modification of protein cysteines. *J. Biol. Chem.* 291, 11477–11490. doi: 10. 1074/jbc.M115.702308

Münchberg, U., Anwar, A., Mecklenburg, S., and Jacob, C. (2007). Polysulfides as biologically active ingredients of garlic. *Org. Biomol. Chem.* 5, 1505–1518. doi: 10.1039/B703832A

Muniz, I. A. F., Campos, D. E. S., Shinkai, R. S. A., Trindade, T. G. D., and Cosme-Trindade, D. C. (2021). Case report of oral mucosa garlic burn during COVID-19 pandemic outbreak and role of teledentistry to manage oral health in an older adult woman. *Spec Care Dentist* 41, 639–643. doi: 10.1111/scd.12605

Ou, X., Liu, Y., Lei, X., Li, P., Mi, D., Ren, L., et al. (2020). Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* 11:1620. doi: 10.1038/s41467-020-15562-9

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47, D442–D450. doi: 10.1093/nar/gky1106

Polak, Y., and Speth, R. C. (2021). Metabolism of angiotensin peptides by angiotensin converting enzyme 2 (ACE2) and analysis of the effect of excess zinc on ACE2 enzymatic activity. *Peptides* 137:170477. doi: 10.1016/j.peptides. 2020.170477

Qin, C., Zhou, L., Hu, Z., Zhang, S., Yang, S., Tao, Y., et al. (2020). Dysregulation of immune response in patients with Coronavirus 2019 (COVID-19) in Wuhan. China. *Clin. Infect. Dis.* 71, 762–768. doi: 10.1093/cid/ciaa248

Rabinkov, A., Miron, T., Konstantinovski, L., Wilchek, M., Mirelman, D., and Weiner, L. (1998). The mode of action of allicin: trapping of radicals and

interaction with thiol containing proteins. *Biochim. Biophys. Acta* 1379, 233– 244. doi: 10.1016/S0304-4165(97)00104-9

Reiter, J., Levina, N., Van Der Linden, M., Gruhlke, M., Martin, C., and Slusarenko, A. J. (2017). Diallylthiosulfinate (Allicin), a volatile antimicrobial from garlic (*Allium sativum*), kills human lung pathogenic bacteria including MDR strains as a vapor. *Molecules* 22:1711. doi: 10.3390/molecules22101711

Rossius, M., Hochgräfe, F., and Antelmann, H. (2018). Thiol-redox proteomics to study reversible protein thiol oxidations in bacteria. *Methods Mol. Biol.* 1841, 261–275. doi: 10.1007/978-1-4939-8695-8_18

Rouf, R., Uddin, S. J., Sarker, D. K., Islam, M. T., Ali, E. S., Shilpi, J. A., et al. (2020). Antiviral potential of garlic (*Allium sativum*) and its organosulfur compounds: a systematic update of pre-clinical and clinical data. *Trends Food Sci. Technol.* 104, 219–234. doi: 10.1016/j.tifs.2020.08.006

Schäfer, G., and Kaschula, C. H. (2014). The immunomodulation and anti-inflammatory effects of garlic organosulfur compounds in cancer chemoprevention. *Anticancer Agents Med. Chem.* 14, 233–240. doi: 10.2174/ 18715206113136660370

Schneider, W. M., Chevillotte, M. D., and Rice, C. M. (2014). Interferon-stimulated genes: a complex web of host defenses. *Annu. Rev. Immunol.* 32, 513–545. doi: 10.1146/annurev-immunol-032713-120231

Schulte-Schrepping, J., Reusch, N., Paclik, D., Bassler, K., Schlickeiser, S., Zhang, B., et al. (2020). Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* 182, 1419–1440.23.

Seidel, K., Kühnert, J., and Adrian, L. (2018). The complexome of *Dehalococcoides mccartyi* reveals its organohalide respiration-complex is modular. *Front. Microbiol.* 9:1130. doi: 10.3389/fmicb.2018.01130

Shekh, S., Reddy, K. K. A., and Gowd, K. H. (2020). In silico allicin induced S-thioallylation of SARS-CoV-2 main protease. *J. Sulphur Chem.* 42, 1–12. doi: 10.1080/17415993.2020.1817457

te Velthuis, A. J., Van Den Worm, S. H., Sims, A. C., Baric, R. S., Snijder, E. J., and Van Hemert, M. J. (2010). Zn(2+) inhibits coronavirus and arterivirus RNA polymerase activity in vitro and zinc ionophores block the replication of these viruses in cell culture. *PLoS Pathog* 6:e1001176. doi: 10.1371/journal.ppat. 1001176

Teixeira, A., Yen, B., Gusella, G. L., Thomas, A. G., Mullen, M. P., Aberg, J., et al. (2015). Prothymosin alpha variants isolated from CD8+ T cells and cervicovaginal fluid suppress HIV-1 replication through type I interferon induction. *J. Infect. Dis.* 211, 1467–1475. doi: 10.1093/infdis/jiu643

Thuy, B. T. P., My, T. T. A., Hai, N. T. T., Hieu, L. T., Hoa, T. T., Thi Phuong, et al. (2020). Investigation into SARS-CoV-2 resistance of compounds in garlic essential oil. *ACS Omega* 5, 8312–8320. doi: 10.1021/acsomega.0c 00772

Tian, W., Zhang, N., Jin, R., Feng, Y., Wang, S., Gao, S., et al. (2020). Immune suppression in the early stage of COVID-19 disease. *Nat. Commun.* 11: 5859.

Vargo, R. J., Warner, B. M., Potluri, A., and Prasad, J. L. (2017). Garlic burn of the oral mucosa: a case report and review of self-treatment chemical burns. *J. Am. Dent. Assoc.* 148, 767–771. doi: 10.1016/j.adaj.2017.02.053

Wang, H., Yang, P., Liu, K., Guo, F., Zhang, Y., Zhang, G., et al. (2008). SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Res.* 18, 290–301. doi: 10.1038/cr. 2008.15

Weber, N. D., Andersen, D. O., North, J. A., Murray, B. K., Lawson, L. D., and Hughes, B. G. (1992). In vitro virucidal effects of *Allium sativum* (garlic) extract and compounds. *Planta Med.* 58, 417–423. doi: 10.1055/s-2006-961504

Wei, L. L., Wang, W. J., Chen, D. X., and Xu, B. (2020). Dysregulation of the immune response affects the outcome of critical COVID-19 patients. *J. Med. Virol.* 92, 2768–2776. doi: 10.1002/jmv.26181

Wheeler, A. P., Wells, C. M., Smith, S. D., Vega, F. M., Henderson, R. B., Tybulewicz, V. L., et al. (2006). Rac1 and Rac2 regulate macrophage morphology but are not essential for migration. *J. Cell Sci.* 119, 2749–2757. doi: 10.1242/jcs.03024

Zecha, J., Lee, C. Y., Bayer, F. P., Meng, C., Grass, V., Zerweck, J., et al. (2020). Data, reagents, assays and merits of proteomics for SARS-CoV-2 research and testing. *Mol. Cell. Proteom.* 19, 1503–1522. doi: 10.1074/mcp.RA120.002164

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., et al. (2020). Crystal structure of SARS-CoV-2 main protease provides a basis for design

of improved alpha-ketoamide inhibitors. *Science* 368, 409–412. doi: 10.1126/science.abb3405

Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273. doi: 10.1038/s41586-020-2012-7

Author Disclaimer: Crushed garlic acts as strong irritant and activates the transient receptor potential channel (TRPA1) in pain-sensing neurons, resulting in pain, inflammation and neurotoxic effects (Bautista et al., 2005). Garlic can be highly toxic to human cells and caused severe garlic burns by its direct exposure to the skin or mucous membranes of the respiratory tract (Bautista et al., 2005; Al-Qattan, 2009; Vargo et al., 2017; Hitl et al., 2021; Muniz et al., 2021). Consequently, any ingestion, inhalation or introduction of fresh garlic can clearly be harmful and hazardous for patients, causing even more severe damage of the lung or the skin. This work does in no way suggest the use of garlic for self-medication of respiratory tract infections, including COVID-19. The authors explicitly warn against any form of self-medication with garlic or garlic compounds.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Lactobacillus Biofilms Influence Anti-*Candida* Activity

Carola Parolin[1]*, Vanessa Croatti[1], Luca Laghi[2], Barbara Giordani[1],
Maria Rosaria Tondi[1,3], Priscilla Romina De Gregorio[4], Claudio Foschi[5] and
Beatrice Vitali[1]

[1] Department of Pharmacy and Biotechnology, Alma Mater Studiorum – University of Bologna, Bologna, Italy, [2] Department of Agricultural and Food Sciences, Centre of Foodomics, Alma Mater Studiorum – University of Bologna, Bologna, Italy, [3] National Institute of Geophysics and Volcanology, Bologna, Italy, [4] Centro de Referencia para Lactobacilos (CERELA)-CONICET, San Miguel de Tucumán, Argentina, [5] Department of Diagnostic and Specialty Medicine, Section of Microbiology, Alma Mater Studiorum – University of Bologna, Bologna, Italy

Lactobacilli are the dominant members of the healthy human vaginal microbiota and represent the first defense line from pathogen infection, including vulvovaginal candidiasis. Biofilm is the predominant microbial growth form in nature, and the formation of biofilms inside the human body has important implications in health and disease. In particular, the formation of biofilm by members of the human resident microbiota is desirable, as it can improve microbial persistence and influence functionality. In the present study, we investigated the capability of 16 vaginal *Lactobacillus* strains (belonging to *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus vaginalis*, and *Lactobacillus plantarum* species) to form biofilms, and we correlated their mode of growth to anti-*Candida* activity. *L. plantarum* strains were the best biofilm producers, and high variability was registered in the level of biofilm formation among *L. crispatus* and *L. gasseri* strains. Culture supernatants derived from *Lactobacillus* biofilm and planktonic growth were tested toward a panel of *Candida* clinical isolates (*Candida albicans*, *Candida glabrata*, *Candida lusitaniae*, *Candida tropicalis*, *Candida krusei*, and *Candida parapsilosis*) and their metabolome assessed by ^{1}H-NMR. *L. crispatus* and *L. plantarum* strains exhibited the best fungistatic profile, and biofilms enhanced their anti-*Candida* activity; on the contrary, *L. gasseri* strains were more effective when grown in a planktonic mode. Biofilm/planktonic mode of growth also affects *Lactobacillus* metabolism, mainly influencing nitrogen and amino acid pathways, and anti-*Candida* activity is instead strictly related to carbohydrate metabolism. The present study underlined the strict interdependence between microbial mode of growth, metabolism, and functional properties. Biofilm formation by members of the healthy human microbiota represents a crucial issue in the field of microbial physiology and host–microbiota interactions, beyond supporting the development of new antimycotic strategies based on probiotics grown in adherence.

Keywords: *Lactobacillus*, *Candida*, biofilm, metabolome, vaginal microbiota, health benefits

INTRODUCTION

Members of the *Lactobacillus* genus are abundant and predominant in the vaginal niche of healthy women of reproductive age, reaching a concentration of 10^7 cfu/ml of vaginal sample, and 80% of the whole microbial content (Rönnqvist et al., 2006; Ceccarani et al., 2019). Most healthy women show the dominance of one species of *Lactobacillus*, and *Lactobacillus crispatus*, *Lactobacillus iners*, *Lactobacillus jensenii*, and *Lactobacillus gasseri* are the most frequent (Ravel et al., 2011). It is widely demonstrated that vaginal lactobacilli are involved in maintaining the state of vaginal eubiosis, preventing the colonization of genital, and sexually transmitted pathogens. Indeed, it has been reported that various vaginal *Lactobacillus* strains can interfere with *Candida* spp. overgrowth and reduce vulvovaginal candidiasis occurrence (Parolin et al., 2015; De Gregorio et al., 2019; Namarta et al., 2020). Anti-*Candida* activity of lactobacilli can be mediated through multiple mechanisms, including pH lowering, secretion of effective compounds, and impairment of fungal adhesion (Parolin et al., 2015, 2021; Calonghi et al., 2017; Abruzzo et al., 2018, 2021; De Gregorio et al., 2019).

Microorganisms' growth can occur in two different forms: planktonic and biofilm, and the latter is the predominant form in nature (Rieu et al., 2014). The biofilm mode of growth is characterized by high cell density and secretion of a polymeric matrix, in which microorganisms are immersed (Flemming and Wingender, 2010). The formation of microbial biofilms inside the human body has important implications in health and disease: for example, pathogens' biofilms generally show increased resistance to antibiotics, making their elimination challenging. On the other hand, the formation of biofilm by members of the human resident microbiota is desirable, as it can improve microbial persistence and favor functionality (Li et al., 2018). The planktonic growth or the alternative establishment of a sessile biofilm by a microbial species implies the remodeling of physiological and metabolic pathways, influencing in turn functional properties of probiotics. For some *Lactobacillus* species, it has been reported that biofilm mode of life can improve antagonistic effects toward bacterial pathogens and anti-inflammatory and immunomodulatory properties (Rieu et al., 2014; Aoudia et al., 2016; Liu et al., 2021), although few studies have investigated this issue till now.

In the present manuscript, 16 vaginal *Lactobacillus* strains, belonging to *L. crispatus*, *L. gasseri, Lactobacillus vaginalis*, and *Lactobacillus plantarum* species, were tested for their *in vitro* capacity to form biofilms. Culture supernatants derived from planktonic and biofilm growth were tested against a panel of *Candida* clinical isolates (*Candida albicans, Candida glabrata, Candida lusitaniae, Candida tropicalis*, *Candida krusei*, and *Candida parapsilosis*). Anti-*Candida* activity was then correlated with lactobacilli culture supernatants metabolome, assessed by ^{1}H-NMR.

MATERIALS AND METHODS

Microorganisms and Culture Conditions

All the 16 *Lactobacillus* strains included in this study were previously isolated from vaginal swabs of healthy premenopausal Caucasian women, following the protocol approved by the Ethics Committee of the University of Bologna, Bologna, Italy (52/2014/U/Tess) (Parolin et al., 2015). *Lactobacillus* strains are listed in **Table 1**. Lactobacilli were routinely cultured in de Man, Rogosa, and Sharpe (MRS) broth (Difco, Detroit, MI, United States) supplemented with 0.05% L-cysteine (Merck, Milan, Italy), at 37°C and in anaerobiosis. Anaerobic conditions were achieved by using jars containing GasPak™ (GasPak™ EZ Anaerobe Container System, Becton, Dickinson and Co., Sparks, MD, United States). Before the experiments, each strain was transferred from the frozen stock to MRS broth and allowed to grow for 24 h and then subcultured in the same medium and conditions for an additional 24 h.

The eight *Candida* strains used in the present study belong to a broad collection of yeasts isolated from vaginal swabs of premenopausal, VVC-affected women during routine diagnostic procedures at the "Microbiology Laboratory" in Sant'Orsola-Malpighi University Hospital of Bologna, Bologna, Italy. All the clinical isolates were coded to assure full anonymousness and are listed in **Table 1** (Oliver et al., 2020; Parolin et al., 2021). *Candida* strains were grown aerobically in Sabouraud dextrose (SD) medium (Difco) at 35°C in aerobiosis.

Lactobacillus Biofilm Formation Assay

To assess *Lactobacillus* strains' capability to form biofilms, bacterial pellets from the second subculture were washed in sterile saline and resuspended in MRS broth to a final concentration of 10^7 colony-forming units (cfu) per milliliter. The bacterial suspension was inoculated in polystyrene, U-bottom 96-well plates (0.2 ml per well) and allowed to grow for 72 h at 37°C, in anaerobiosis. At the end of the incubation, biofilm formation was evaluated by crystal violet staining, following the protocol described by Leccese Terraf et al. (2014). Briefly, culture supernatant was discarded and wells were gently washed with phosphate buffer, pH 7.4. Adherent biofilm was stained with 0.1% (w/v) crystal violet (CV) aqueous solution for 30 min and then washed with distilled water. Plates were dried overnight at room temperature, and the CV bound to the adherent biofilm was solubilized in 30% (v/v) acetic acid. CV absorbance was measured at 595 nm with an Enspire multiplate reader (Perkin Elmer, United States). Each strain was assayed in at least three independent experiments, each with four replicates. Additionally, a sterile culture medium was always included as negative control.

Preparation of Cell-Free Supernatants From *Lactobacillus* Planktonic and Biofilm Cultures

Lactobacillus strains were grown in planktonic and biofilm forms and culture supernatants collected. 10^7 cfu/ml bacterial

TABLE 1 | List of microorganisms used in the present manuscript and origin (VVC: vulvovaginal candidiasis).

Species	Strain denomination	Origin
***Lactobacillus* strains**		
L. crispatus	BC1	Healthy human vagina
L. crispatus	BC3	Healthy human vagina
L. crispatus	BC4	Healthy human vagina
L. crispatus	BC5	Healthy human vagina
L. crispatus	BC6	Healthy human vagina
L. crispatus	BC7	Healthy human vagina
L. gasseri	BC9	Healthy human vagina
L. gasseri	BC10	Healthy human vagina
L. gasseri	BC11	Healthy human vagina
L. gasseri	BC12	Healthy human vagina
L. gasseri	BC13	Healthy human vagina
L. gasseri	BC14	Healthy human vagina
L. vaginalis	BC16	Healthy human vagina
L. vaginalis	BC17	Healthy human vagina
L. plantarum	BC18	Healthy human vagina
L. plantarum	BC19	Healthy human vagina
***Candida* strains**		
C. albicans	SO1	VVC human vagina
C. albicans	SO2	VVC human vagina
C. glabrata	SO17	VVC human vagina
C. glabrata	SO18	VVC human vagina
C. lusitaniae	SO22	VVC human vagina
C. tropicalis	SO24	VVC human vagina
C. krusei	SO26	VVC human vagina
C. parapsilosis	SO27	VVC human vagina

suspensions were prepared as described above and inoculated in sterile glass tubes (10 ml, planktonic form) or polystyrene six-well plates (4 ml per well, biofilm form). Planktonic (pk) and biofilm (bf) samples were incubated for 72 h at 37°C, in anaerobiosis. Afterward, culture supernatants were recovered, centrifuged (2,750 × *g*, 10 min) and filtered through a 0.22-µm membrane filter to obtain cell-free supernatants (CFS). CFS obtained from planktonic (pk-CFS) and biofilm (bf-CFS) cultures were stored at −20°C until their use.

Metabolomics of *Lactobacillus* Cell-Free Supernatants

After centrifugation at 18,630 × *g* for 10 min at 4°C, 0.7 ml of pk-CFS and bf-CFS samples were added to 0.1 ml of a D2O solution of 3-(trimethylsilyl)-propionic-2,2,3,3-d4 acid sodium salt (TSP) 10 mM, set to pH 7.0 by means of a 1 M phosphate buffer. The solution contained also NaN$_3$ 2 mM, to avoid microorganisms' proliferation. ^{1}H-NMR spectra were recorded with an AVANCE III spectrometer (Bruker, Milan, IT, United States) at 298 K operating at a frequency of 600.13 MHz. To reduce broad signals caused by slowly tumbling molecules, a CPMG filter of 400 echoes, separated by an echo time of 400 µs, was applied. Water residual signal was reduced by presaturation. The

signals were assigned by comparing their chemical shift and multiplicity with Chenomx software data bank (Chenomx Inc., Canada, ver 11.05) (Laghi et al., 2014). Metabolite concentrations were calculated and reported as differences with respect to MRS medium.

Anti-*Candida* Activity of *Lactobacillus* Supernatants

Planktonic and biofilm supernatants anti-*Candida* activity was tested by microdilution assay, following EUCAST guidelines (EUCAST Committee, 2020). Briefly, stock *Candida* suspensions prepared in sterile water at an absorbance (measured at 600 nm) of 0.008–0.137 were diluted 1:10 in RPMI 1640 medium buffered to pH 7.0 with 0.165 M MOPS (morpholinepropanesulfonic acid buffer) and added with 2% glucose. Final yeast suspensions, corresponding to 1–5 × 10^5 cfu/ml, were inoculated in flat-bottomed 96-well plates (0.1 ml per well) and added with the same volume of each *Lactobacillus* pk-CFS and bf-CFS. Positive growth control wells contained 0.1 ml of *Candida* suspension added with the same volume of sterile MRS medium. The plates were incubated at 35°C for 24 h; afterward, *Candida* growth was evaluated by reading the absorbance at 530 nm with an Enspire multiplate reader. *Candida* growth inhibition was calculated relative to the absorbance of the corresponding positive controls. A subset of *Lactobacillus* pk-CFS and bf-CFS was tested for anti-*Candida* activity in simulated vaginal fluid (SVF), pH 4.2, by microdilution assay, following the method described above. In brief, *Candida* suspensions were diluted 1:10 in SVF, composed of 3.51 g/L NaCl, 1.40 g/L KOH, 0.222 g/L Ca(OH)$_2$, 0.018 g/L bovine serum albumin, 2 g/L lactic acid, 1 g/L acetic acid, 0.16 g/L glycerol, 0.4 g/L urea, and 5 g/L glucose. The pH was adjusted to 4.2 with HCl 0.1 N and the solution was filtered through a 0.22-µm membrane filter. CFS obtained from *L. crispatus* BC1, *L. crispatus* BC6, *L. gasseri* BC11, *L. gasseri* BC12, *L. vaginalis* BC17, and *L. plantarum* BC19 were tested.

Cell-free supernatants fungicidal activity was assessed by spotting 20 µl of active samples onto SD agar plates, then incubated at 35°C for 24/48 h. *Candida* viability was determined by the ability of forming colonies.

Data Analysis and Statistics

Data were analyzed by MatLab software (R2020b version 9.9.0.1524771, The MathWorks Inc., Natick, MA, United States) with the Statistics and Machine Learning Toolbox. Cluster analysis was performed by means of t-Distributed Stochastic Neighbor Embedding method (Van der Maaten and Hinton, 2008). Differences in biofilm formation were assessed by Kruskal–Wallis test followed by Tukey-Kramer test, and differences in metabolite concentrations were assessed by Wilcoxon rank test. Enrichment analysis was performed on significantly different metabolites, by using MetaboAnalyst web resource[1]. Differences were deemed significant for $p < 0.05$.

---

[1]https://www.metaboanalyst.ca/

RESULTS

Ability of Vaginal Lactobacilli to Form Biofilm

Sixteen *Lactobacillus* strains isolated from vaginal swabs were tested for their capability to form biofilm on an abiotic polystyrene surface (**Figure 1**). In general, all vaginal *Lactobacillus* used in this study can form biofilm, although at different amounts ($p < 0.05$). The strongest biofilm producers were *L. plantarum* BC18 and BC19 strains ($p < 0.05$), while *L. vaginalis* BC16 and BC17 strains were the weakest ($p < 0.05$). Within *L. crispatus* and *L. gasseri* species, very high variability was observed. In particular, among *L. crispatus*, BC3 and BC7 exhibited the highest biofilm production; for *L. gasseri*, BC12 was the best producer.

Anti-*Candida* Activity of Lactobacilli in Planktonic/Biofilm Cultures

The fungistatic activity of bf-CFS and pk-CFS of vaginal lactobacilli was evaluated toward eight *Candida* clinical isolates, belonging to *C. albicans* and non-*albicans* species (**Table 1**). The anti-*Candida* activity was firstly assessed following EUCAST guidelines for antifungal susceptibility testing (EUCAST Committee, 2020), which imply RPMI supplemented with glucose and MOPS, pH 7. Overall, the most sensitive *Candida* isolates were *C. albicans* SO2, *C. lusitaniae* SO22, and *C. parapsilosis* SO27, whereas *C. krusei* SO26 was the most resistant to *Lactobacillus* CFS (**Figure 2**). bf-CFS from *L. crispatus* and *L. plantarum* strains showed the highest fungistatic activity toward all *Candida* isolates, exhibiting the best anti-*Candida* profile. In detail, *L. crispatus* and *L. plantarum* bf-CFS showed higher values of *Candida* growth inhibition than the respective pk-CFS: bf-CFS almost completely inhibited (over 95% growth

inhibition) *C. albicans* SO2, *C. glabrata* SO17, *C. glabrata* SO18, *C. lusitaniae* SO22, and *C. parapsilosis* SO27 growth, and strongly inhibited (over 65% inhibition) *C. albicans* SO1 and *C. tropicalis* SO24. The corresponding pk-CFS showed inhibition rates ranging from 50 to 90% toward the same *Candida* strains. On the contrary, most *L. gasseri* strains exerted higher anti-*Candida* activity when they were cultured in planktonic mode rather than in biofilm, and great variability among *Candida* isolates' inhibition was registered. Overall, *L. gasseri* pk-CFS were mostly active toward *C. lusitaniae* SO22 and *C. parapsilosis* SO27, and to a lesser extent on *C. albicans* SO2. Intermediate anti-*Candida* activity was shown on *C. glabrata* SO17, *C. glabrata* SO18, and *C. tropicalis* SO24. *L. vaginalis* BC16 bf-CFS and pk-CFS showed the lowest fungistatic activity against all *Candida* isolates, exhibiting the worst anti-*Candida* profile. Overall, CFS were not fungicidal, with the only exception of *L. crispatus* bf-CFS toward *C. parapsilosis* SO27 and *L. plantarum* bf-CFS toward *C. lusitaniae* SO22.

In order to mimic the physiological condition typically found in the vaginal environment, the fungistatic activity of CFS was also tested in SVF, pH 4.2. A subset of active CFS was selected (i.e., CFS obtained from *L. crispatus* BC1, *L. crispatus* BC6, *L. gasseri* BC11, *L. gasseri* BC12, *L. vaginalis* BC17, and *L. plantarum* BC19) and tested toward the panel of *Candida* isolates (**Figure 3**). Also in these experimental conditions, CFS showed a remarkable anti-*Candida* activity; interestingly, *Lactobacillus* CFS were highly active also toward *C. glabrata* SO17, *C. glabrata* SO18, and *C. krusei* SO26 (over 60% inhibition). Overall, acidic pH of SVF enhanced *Candida* growth inhibition by CFS, especially for pk-CFS.

To better understand the analogies of CFS activities toward *Candida* isolates, a cluster analysis was performed using the t-SNE method on data obtained following EUCAST guidelines (**Figure 4**). Considering the position of samples in the 3D map,
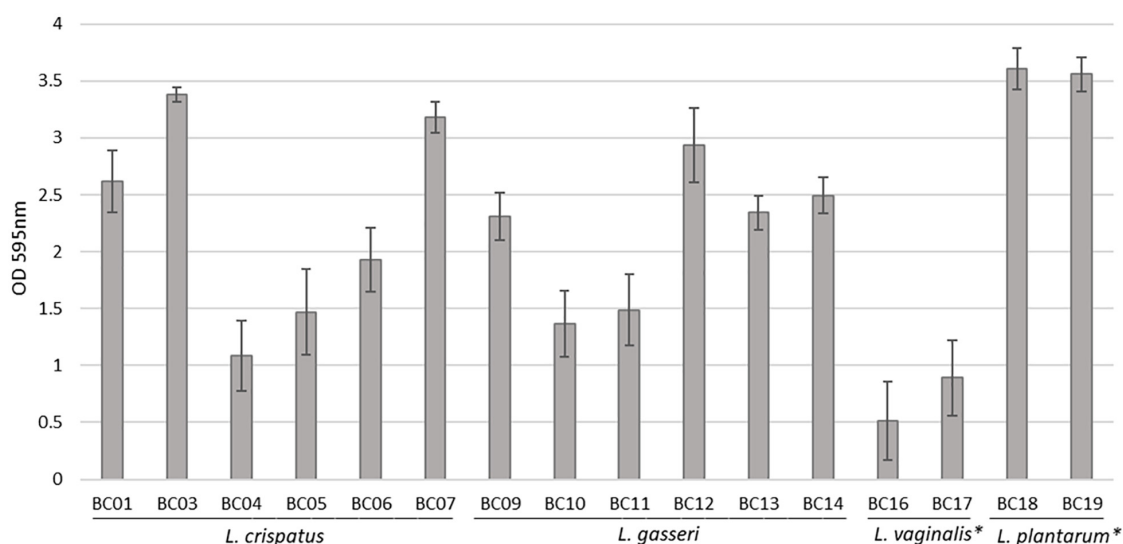
FIGURE 1 | Biofilm formation by *Lactobacillus* strains, evaluated by crystal violet staining (average ± st. dev.). *$p < 0.05$ selected species vs. other *Lactobacillus* species.

FIGURE 2 | Fungistatic activity of *Lactobacillus* bf-CFS and pk-CFS toward clinical *Candida* isolates, tested following EUCAST guidelines (pH 7). OD 530 nm values were normalized for positive controls and data were expressed as inhibition (%) of *Candida* growth. Fungistatic activity color scale was reported (10% activity variation).

FIGURE 3 | Fungistatic activity of *L. crispatus* BC1, *L. crispatus* BC6, *L. gasseri* BC11, *L. gasseri* BC12, *L. vaginalis* BC17, and *L. plantarum* BC19 bf-CFS and pk-CFS toward clinical *Candida* isolates, tested in simulated vaginal fluid, pH 4.2. OD 530 nm values were normalized for positive controls and data were expressed as inhibition (%) of *Candida* growth. Fungistatic activity color scale was reported.

FIGURE 4 | t-SNE cluster analysis on CFS fungistatic activity dataset. Three groups were identified, corresponding to low, medium, and high activity. Samples derived from biofilm and planktonic cultures were signed in black and red, respectively.

FIGURE 5 | Box plot of *Lactobacillus* CFS metabolite concentrations. Concentrations were calculated as differences from MRS broth (mmol/L). Lines within the boxes indicate the median values of the samples. Each box represents the interquartile range (25–75th percentile). The extremes of the bars indicate the 10th and 90th percentiles, respectively. Outlier values are indicated (+). Glucose median and acetate interquartile range are not represented because they were out of the range considered.

we identified three groups of CFS defined as low, medium, and high activity. All *L. crispatus* and *L. plantarum* bf-CFS are grouped in the high-activity cluster, while pk-CFS from the same species and *L. gasseri* pk-CFS are grouped in the medium one. *L. gasseri* bf-CFS, except for *L. gasseri* BC11 and BC14 bf-CFS, are grouped with *L. vaginalis* BC16 CFS and demonstrated a low fungistatic activity.

Lactobacillus Metabolome Analysis

In order to better characterize vaginal *Lactobacillus* CFS and bf-CFS and pk-CFS peculiarities, their metabolomic profiles were obtained by ^{1}H-NMR. Fifty-one molecules belonging to the families of amino acids, organic acids, monosaccharides/disaccharides, ketones, and alcohols were identified (**Figure 5**), and differences were calculated with respect to MRS medium. All metabolites showed similar concentration median values, except for acetate and glucose. The same molecules also showed high variability among samples. We thus investigated the metabolites that mainly differentiate bf-CFS and pk-CFS samples, and we found out that the quantities of 17 molecules were significantly different between the two groups (**Table 2**). Among those, 10 amino acids (or derivatives) were identified, suggesting a deep involvement of this molecule category. Data were then subjected to overrepresentation analysis (**Supplementary Figure 1**), which highlighted that the growth form (bf vs. pk) mainly affects the nitrogen metabolism, and specifically several amino acidic pathways (serine, alanine, and lysine).

TABLE 2 | List of metabolites significantly different between *Lactobacillus* bf-CFS and pk-CFS ($p < 0.05$) and concentrations, calculated as differences with respect to MRS medium [median value ± MAD (mmol/L)].

Metabolites	bf-CFS	pk-CFS
Amino acids and derivatives		
Alanine	0.7532 ± 0.1226	0.4625 ± 0.0935
Asparagine	0.0153 ± 0.0377	0.0558 ± 0.0415
Glutamate	−0.5813 ± 0.0343	−0.5487 ± 0.0210
Glutamine	0.2304 ± 0.0972	0.1152 ± 0.1133
2-Furoylglycine	0.0044 ± 0.0013	0.0003 ± 0.0037
Leucine	0.5522 ± 0.1429	0.4502 ± 0.0953
Lysine	0.1953 ± 0.0708	0.0982 ± 0.0380
Methionine	−0.1272 ± 0.2105	0.1092 ± 0.0117
Threonine	0.5409 ± 0.0835	0.6537 ± 0.1103
Sarcosine	0.2565 ± 0.2150	0.0209 ± 0.0204
Organic acids		
Citrate	0.0145 ± 0.0024	0.0051 ± 0.0020
Ribose	0.0277 ± 0.0072	0.0092 ± 0.0075
Alcohols		
Glycerol	0.1559 ± 0.0674	0.2926 ± 0.0900
Others		
Acetone	−0.0489 ± 0.0013	0.0250 ± 0.0081
1,3-Dihydroxyacetone	0.2195 ± 0.0592	0.3872 ± 0.1100
Acetoin	0.0034 ± 0.0062	0.0214 ± 0.0119
Allantoin	0.0080 ± 0.0011	0.0060 ± 0.0013

Correlation of Anti-*Candida* Activity With Supernatant Metabolomes

A principal component analysis (PCA) was performed for *Lactobacillus* bf-CFS and pk-CFS considering metabolomic and activity against *Candida* (**Figure 6A**). In the biplot, PC1 and PC2 represented the 44.3% of the total variance of the investigated samples (PC1 26.2%; PC2 18.1%). *Lactobacillus* CFS with different anti-*Candida* activity separated on PC1. In particular, *L. crispatus* and *L. plantarum* bf-CFS, which belong to the high anti-*Candida* activity group, were characterized by low values of PC1, while the low activity CFS samples showed high values of PC1 (**Figure 6B**). In order to identify the metabolites that mainly contribute to the anti-*Candida* activity of *Lactobacillus* CFS, we compared the metabolome of the active CFS to that of the less active ones. We found out that 18 molecule amounts were significantly different in the two groups (**Table 3**); most of them are represented by carbohydrates and amino acids. Enrichment analysis suggested a significant difference between the two CFS groups in the metabolism of sugars, mainly involving lactose and galactose pathways.

DISCUSSION

The two different forms of microbial growth, i.e., planktonic and biofilm, differ for physiological and metabolic features; indeed, the establishment of a sessile culture by a microbial strain requires the remodeling of several molecular pathways involved in adhesion, secretion, and other cellular functions. Such modifications of the global metabolism can, in turn, influence the functional properties of probiotics, beyond favoring the permanence of probiotics inside the human body (Donlan and Costerton, 2002; Ventolini, 2015).

In this context, we first assessed the ability of *Lactobacillus* strains of vaginal origin, belonging to *L. crispatus*, *L. gasseri*, *L. vaginalis*, and *L. plantarum* species, to form biofilm. We found out that *L. plantarum* species was the best biofilm producer, probably because of its adaptive features to niches (van den Nieuwboer et al., 2016). Indeed, *L. plantarum* is a versatile species able to adapt to different environmental niches and its genome, considered as one of the largest genome in lactic acid bacteria, encodes a huge repertoire of surface sugar, extracellular proteins, and regulatory peptide involved in the regulation of adherence (Kleerebezem et al., 2003; Sturme et al., 2005). Other authors previously reported that *L. plantarum* was able to produce a robust biofilm *in vitro*, resistant to environmental stresses (Kubota et al., 2008). Considering the most peculiar species of the vaginal microbiota, high variability was registered in the level of biofilm formation among *L. crispatus* and *L. gasseri* strains. Notably, *L. crispatus* BC3 biofilm was comparable to *L. plantarum* strains. Biofilm formation by *L. crispatus* and *L. gasseri* has been poorly investigated in the literature, and it has been reported that some *L. crispatus* strains of vaginal origin were not able to form biofilm and weakly autoaggregated (van der Veer et al., 2018), while Stivala et al. (2021) reported about vaginal *L. crispatus* and *L. gasseri* (and *L. rhamnosus*) strains as weak to moderate biofilm producers, with great variability.
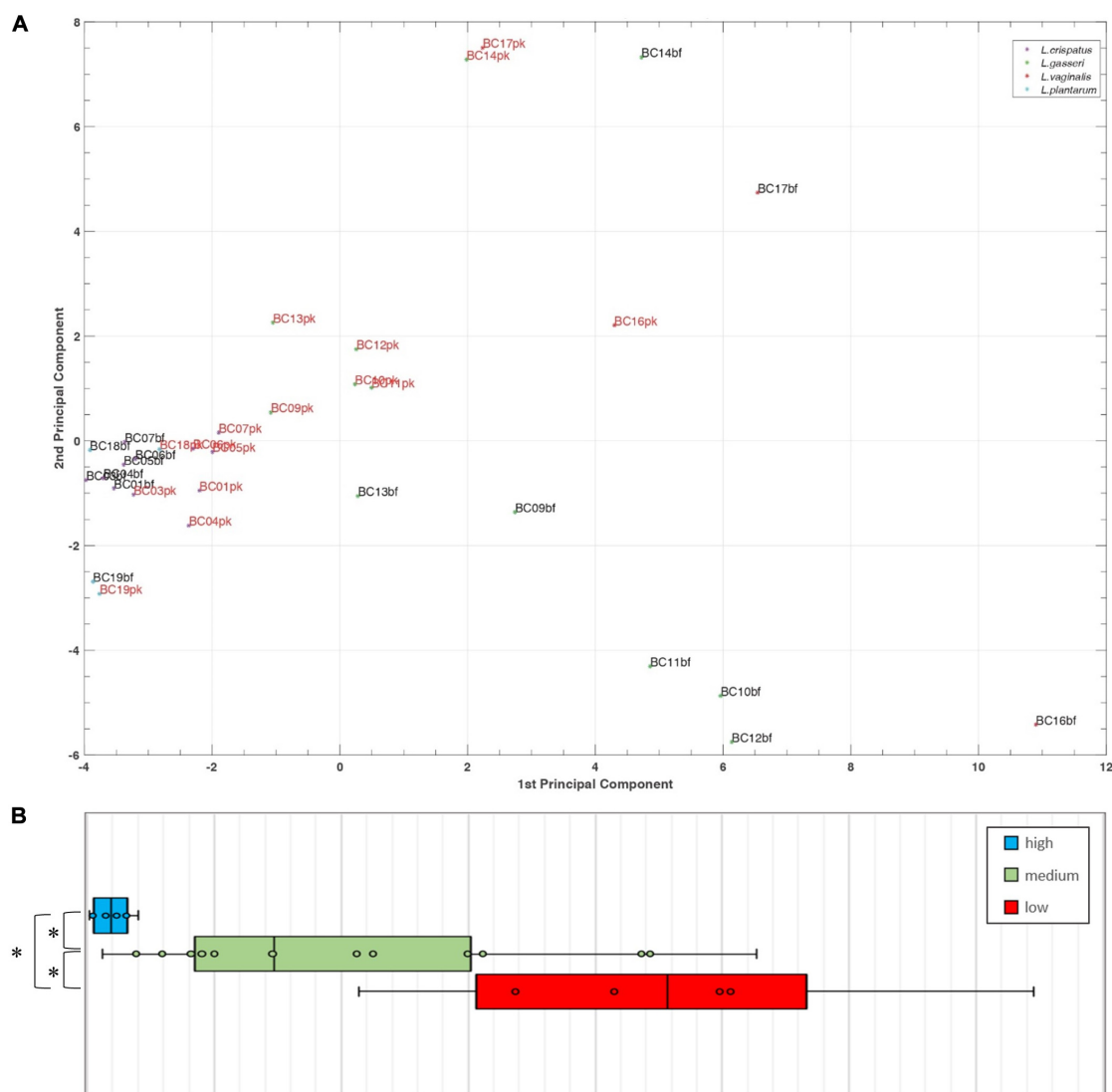
FIGURE 6 | (A) Biplot of a PCA performed on *Lactobacillus* bf-CFS and pk-CFS metabolome and anti-*Candida* activity data. **(B)** Boxplot of CFS distribution on PC1 based on anti-*Candida* activity score. Lines within the boxes indicate the median values of the samples. Each box represents the interquartile range (25–75th percentile). The extremes of the bars indicate the 10th and 90th percentiles, respectively.

From our results, the shift from a *Lactobacillus* planktonic culture to an adherent form implied major differences in the metabolism of nitrogen and amino acids, suggesting a key role of this molecule category in the formation of an adherent biomass, as already proposed by Liu et al. for *Lactobacillus paraplantarum* (Liu et al., 2021). In particular, alanine and lysine were produced/released at higher amounts by lactobacilli in the biofilm form compared to the planktonic one, whereas methionine was highly consumed by *Lactobacillus* in biofilm growth. These data support the finding that the methionine pathway is involved in the biofilm formation process, as proposed by Liu et al. (2018) in *L. paraplantarum*.

Although the anti-*Candida* activity of planktonic-derived *Lactobacillus* CFS has been reported and explored by several authors (Parolin et al., 2015; Wang et al., 2017;

De Gregorio et al., 2019; Jang et al., 2019; Tortelli et al., 2020), no data are available on bf-CFS anti-*Candida* potential. In the present manuscript, both planktonic- and biofilm-derived CFS were tested against a broad spectrum of *Candida* clinical isolates, including *C. albicans* and *C. non-albicans* species. Overall, CFS did not exert fungistatic activity, with few exceptions. Fungistatic activity was tested in standard conditions following EUCAST protocol and validated in SVF. In accordance with results reported by Parolin et al. (2015), *C. krusei* resulted in being the most resistant isolate, although acidic conditions enhanced CFS fungistatic effect. *C. glabrata* isolates were susceptible to *Lactobacillus* CFS and notably their growth was almost abrogated by *L. crispatus* bf-CFS, although this species is often reported as resistant to antifungal substances (Rodrigues et al., 2014). Unlike previous results (Parolin et al., 2015) that were

TABLE 3 | List of metabolites significantly different between *Lactobacillus* CFS belonging to high anti-*Candida* activity group and low anti-*Candida* activity group ($p < 0.05$) and concentrations, calculated as differences with respect to MRS medium [median value ± MAD (mmol/L)].

Metabolites	High-activity CFS	Low-activity CFS
Amino acids and derivatives		
Glycine	−0.3130 ± 0.0406	−0.1280 ± 0.0770
Histidine	0.2398 ± 0.0091	0.0886 ± 0.0748
Isoleucine	0.1051 ± 0.0188	0.2457 ± 0.0427
Methionine	0.0747 ± 0.0113	−0.3244 ± 0.1486
Proline	0.0878 ± 0.0080	0.1261 ± 0.0143
Valine	0.3559 ± 0.0342	0.5096 ± 0.0647
Sarcosine	0.0570 ± 0.0155	0.5450 ± 0.1427
Organic acids		
Acetate	4.6614 ± 0.5951	−12.2848 ± 3.5317
Lactate	−0.0499 ± 0.0293	0.1147 ± 0.0749
Pyruvate	0.1285 ± 0.0199	0.0311 ± 0.0244
Sugars		
Fructose	−3.0765 ± 0.0353	−0.5228 ± 0.9609
Galactose	0.0436 ± 0.0171	0.0169 ± 0.0102
Glucose	−83.9781 ± 0.0512	−58.5522 ± 13.3528
Ribose	0.0266 ± 0.0017	0.0609 ± 0.0340
myo-Inositol	−0.0588 ± 0.0197	−0.1496 ± 0.0204
Others		
sn-Glycero-3-phosphocholine	0.0284 ± 0.0051	0.0083 ± 0.0072
Orotate	0.0059 ± 0.0015	−0.0005 ± 0.0008
Uracil	−0.0167 ± 0.0064	0.0527 ± 0.0253

obtained with CFS from lactobacilli in exponential growth phase, *Lactobacillus* CFS from the stationary phase were active toward *C. parapsilosis*, pointing out the importance of growth phase and resulting metabolites on the anti-*Candida* activity. In addition, bf-CFS were more effective against *C. parapsilosis* than the respective pk-CFS.

Lactobacillus CFS that demonstrated high anti-*Candida* activity all derived from biofilm mode of growth and were produced by *L. crispatus* and *L. plantarum* strains; on the contrary, low-activity CFS were derived from *L. gasseri* (BC11–BC14) biofilm cultures and *L. vaginalis* BC16. This observation suggested that the anti-*Candida* activity of *L. crispatus* strains was enhanced when they were able to establish an adherent biofilm, while *L. gasseri* strains behaved in the opposite manner. This result, in turn, can offer new hints in the perspective of using *L. crispatus* or *L. gasseri* as probiotics to prevent vulvovaginal candidiasis or to restore the vaginal eubiosis status, also in the light of developing fourth-generation probiotics, based on biofilm-forming strains and encapsulation techniques (Salas-Jara et al., 2016). Peculiarly, *L. vaginalis* BC16, especially when grown in the biofilm mode, slightly stimulated *Candida* proliferation. This behavior correlated with *L. vaginalis* BC16 weak attitude to form an adherent biofilm, suggesting that its preferred mode of growth (i.e., planktonic) assured the best conditions to prevent from *Candida* infection, sustaining in turn the healthy status of the vaginal ecological niche.

In the present manuscript, we also sought for peculiarities in the metabolome of *Lactobacillus* CFS that demonstrated highly effective toward *Candida* clinical isolates. Highly active CFS differentiated from low active ones mostly for carbohydrates and amino acid quantification. Among carbohydrates, high anti-*Candida* activity was accompanied by increased consumption of glucose and fructose (corresponding to lower amounts), two well-known glycolytic fuel molecules. A marked consumption of glucose has already been correlated with *Lactobacillus* antimicrobial effect (Nardini et al., 2016), highlighting the mechanism of competition for nutrients between human beneficial microbes and pathogens as crucial. On the other hand, galactose concentration was higher in active CFS compared to non-active ones and lactose/galactose metabolism has been identified by overrepresentation analysis as a differential metabolic pathway between active and non-active CFS, denoting the preference of active *Lactobacillus* strains to utilize other carbon sources than galactose to sustain their growth.

Another recognized mechanism supporting the antimicrobial activity of lactobacilli is lactic acid production by fermentation and acidification (Boskey et al., 2001; O'Hanlon et al., 2013). In this regard, lower amount of lactate was quantified in active CFS compared to non-active ones, suggesting that the observed anti-*Candida* effect was not directly attributable to this organic acid. CFS endowed with high anti-*Candida* activity also showed higher amounts of pyruvate with respect to low-activity CFS, suggesting that the antifungal activity did not depend on a burst of the classical lactic fermentation pathway. In addition, active CFS contained high amounts of acetate, although *L. crispatus, L. gasseri*, and *L. plantarum* are recognized as homofermentative lactic acid bacteria. *Lactobacillus* CFS fungistatic activity was assessed in standard and acidic conditions, suggesting that the anti-*Candida* effect is not merely dependent on the acidic pH. Nevertheless, when CFS acted in an acidic environment, the anti-*Candida* activity was enhanced, pointing up the existence of a synergic activity among different components of the *Lactobacillus* supernatants and suggesting that the vaginal physiological environment could improve CFS anti-*Candida* potential.

To the best of our knowledge, the present study reports the anti-*Candida* activity of vaginal *Lactobacillus* grown in a sessile mode for the first time. Biofilm formation by lactobacilli represents a crucial aspect in the vaginal ecology, as this phenomenon can possibly influence endogenous microbiota persistence and functionality. Here, we demonstrated that *Lactobacillus* adherent growth strongly affects the anti-*Candida* potential and could in turn modulate *Candida* colonization of the vaginal niche. In particular, biofilm formation enhanced *L. crispatus* and *L. plantarum* fungistatic activity, while *L. gasseri* became more active when grown in a free form. Lactobacilli metabolism turned different between planktonic and biofilm forms, though the correlation between metabolism and functional properties demands more investigation. Further studies on lactobacilli biofilm formation in the vaginal niche will be approached to better understand its physiological role, also in the perspective to support the development of new antimycotic strategies based on probiotics grown in adherence.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

CP and BV contributed to conception and design of the study. CP supervised the experiments. VC, LL, BG, and PD performed the experiments. CP, VC, LL, and MT analyzed the data. VC and MT produced the figures. CP and VC wrote the first draft of the manuscript. CF provided *Candida* strains. BV supervised the study. All authors contributed to manuscript revision, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.750368/full#supplementary-material

REFERENCES

Abruzzo, A., Giordani, B., Parolin, C., De Gregorio, P. R., Foschi, C., Cerchiara, T., et al. (2021). Lactobacillus crispatus BC1 biosurfactant delivered by hyalurosomes: an advanced strategy to counteract candida biofilm. *Antibiotics* 10:33. doi: 10.3390/antibiotics10010033

Abruzzo, A., Giordani, B., Parolin, C., Vitali, B., Protti, M., Mercolini, L., et al. (2018). Novel mixed vesicles containing lactobacilli biosurfactant for vaginal delivery of an anti- Candida agent. *Eur. J. Pharm. Sci.* 112, 95–101. doi: 10.1016/j.ejps.2017.11.012

Aoudia, N., Rieu, A., Briandet, R., Deschamps, J., Chluba, J., Jego, G., et al. (2016). Biofilms of Lactobacillus plantarum and Lactobacillus fermentum: effect on stress responses, antagonistic effects on pathogen growth and immunomodulatory properties. *Food Microbiol.* 53, 51–59. doi: 10.1016/j.fm.2015.04.009

Boskey, E. R., Cone, R. A., Whaley, K. J., and Moench, T. R. (2001). Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source. *Hum. Reprod.* 16, 1809–1813. doi: 10.1093/humrep/16.9.1809

Calonghi, N., Parolin, C., Sartor, G., Verardi, L., Giordani, B., Frisco, G., et al. (2017). Interaction of vaginal Lactobacillus strains with HeLa cells plasma membrane. *Benef. Microbes* 8, 625–633. doi: 10.3920/BM2016.0212

Ceccarani, C., Foschi, C., Parolin, C., D'Antuono, A., Gaspari, V., Consolandi, C., et al. (2019). Diversity of vaginal microbiome and metabolome during genital infections. *Sci. Rep.* 9:14095. doi: 10.1038/s41598-019-50410-x

EUCAST Committee (2020). Method for the determination of broth dilution minimum inhibitory concentrations of antifungal agents for yeasts. *E. Def.* 7.3.2, 1–21.

De Gregorio, P., Silva, J., Marchesi, A., and Nader-Macías, M. (2019). Anti-Candida activity of beneficial vaginal lactobacilli in in vitro assays and in a murine experimental model. *FEMS Yeast Res.* 19:foz008.

Donlan, R. M., and Costerton, J. W. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin. Microbiol. Rev.* 15, 167–193. doi: 10.1128/CMR.15.2.167-193.2002

Flemming, H. C., and Wingender, J. (2010). The biofilm matrix. *Nat. Rev. Microbiol.* 8, 623–633. doi: 10.1038/nrmicro2415

Jang, S. J., Lee, K., Kwon, B., You, H. J., and Ko, G. P. (2019). Vaginal lactobacilli inhibit growth and hyphae formation of Candida albicans. *Sci. Rep.* 9:8121. doi: 10.1038/s41598-019-44579-4

Kleerebezem, M., Boekhorst, J., Van Kranenburg, R., Molenaar, D., Kuipers, O. P., Leer, R., et al. (2003). Complete genome sequence of Lactobacillus plantarum WCFS1. *Proc. Natl. Acad. Sci. U. S. A.* 100, 1990–1995. doi: 10.1073/pnas.0337704100

Kubota, H., Senda, S., Nomura, N., Tokuda, H., and Uchiyama, H. (2008). Biofilm Formation by Lactic Acid Bacteria and Resistance to Environmental Stress. *J. Biosci. Bioeng.* 106, 381–386. doi: 10.1263/jbb.106.381

Laghi, L., Picone, G., Cruciani, F., Brigidi, P., Calanni, F., Donders, G., et al. (2014). Rifaximin modulates the vaginal microbiome and metabolome in women affected by bacterial vaginosis. *Antimicrob. Agents Chemother.* 58, 3411–3420. doi: 10.1128/AAC.02469-14

Leccese Terraf, M. C., Mendoza, L. M., Juárez Tomás, M. S., Silva, C., and Nader-Macías, M. E. F. (2014). Phenotypic surface properties (aggregation, adhesion and biofilm formation) and presence of related genes in beneficial vaginal lactobacilli. *J. Appl. Microbiol.* 117, 1761–1772. doi: 10.1111/jam.12642

Li, Z., Behrens, A. M., Ginat, N., Tzeng, S. Y., Lu, X., Sivan, S., et al. (2018). Biofilm-inspired encapsulation of probiotics for the treatment of complex infections. *Adv. Mater.* 30:e1803925. doi: 10.1002/adma.201803925

Liu, L., Guo, S., Chen, X., Yang, S., Deng, X., Tu, M., et al. (2021). Metabolic profiles of Lactobacillus paraplantarum in biofilm and planktonic states and investigation of its intestinal modulation and immunoregulation in dogs. *Food Funct.* 12, 5317–5332. doi: 10.1039/d1fo00905b

Liu, L., Wu, R., Zhang, J., and Li, P. (2018). Overexpression of luxSPromotes stress resistance and biofilm Formation of lactobacillus paraplantarumL-ZS9 by regulating the expression of multiple genes. *Front. Microbiol.* 9:2628. doi: 10.3389/fmicb.2018.02628

Namarta, K., Jatinder, S., and Manpreet, K. (2020). Microbiota in vaginal health and pathogenesis of recurrent vulvovaginal infections: a critical review. *Ann. Clin. Microbiol. Antimicrob.* 19:5. doi: 10.1186/s12941-020-0347-4

Nardini, P., Ñahui Palomino, R. A., Parolin, C., Laghi, L., Foschi, C., Cevenini, R., et al. (2016). Lactobacillus crispatus inhibits the infectivity of Chlamydia trachomatis elementary bodies, in vitro study. *Sci. Rep.* 6:29024. doi: 10.1038/srep29024

O'Hanlon, D. E., Moench, T. R., and Cone, R. A. (2013). Vaginal pH and microbicidal lactic acid when lactobacilli dominate the microbiota. *PLoS One* 8:e80074. doi: 10.1371/journal.pone.0080074

Oliver, J. C., Laghi, L., Parolin, C., Foschi, C., Marangoni, A., Liberatore, A., et al. (2020). Metabolic profiling of Candida clinical isolates of different species and infection sources. *Sci. Rep.* 10:16716. doi: 10.1038/s41598-020-73889-1

Parolin, C., Abruzzo, A., Giordani, B., Oliver, J. C., Marangoni, A., Luppi, B., et al. (2021). Anti-candida activity of hyaluronic acid combined with lactobacillus crispatus lyophilised supernatant: a new antifungal strategy. *Antibiotics* 10:628. doi: 10.3390/antibiotics10060628

Parolin, C., Marangoni, A., Laghi, L., Foschi, C., Palomino, R. A. Ñ, Calonghi, N., et al. (2015). Isolation of vaginal lactobacilli and characterization of anti-candida activity. *PLoS One* 10:e0131220. doi: 10.1371/journal.pone.0131220

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S. K., McCulle, S. L., et al. (2011). Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. U. S. A.* 108, 4680–4687. doi: 10.1073/pnas.1002611107

Rieu, A., Aoudia, N., Jego, G., Chluba, J., Yousfi, N., Briandet, R., et al. (2014). The biofilm mode of life boosts the anti-inflammatory properties of Lactobacillus. *Cell. Microbiol.* 16, 1836–1853. doi: 10.1111/cmi.12331

Rodrigues, C. F., Silva, S., and Henriques, M. (2014). Candida glabrata: a review of its features and resistance. *Eur. J. Clin. Microbiol. Infect. Dis.* 33, 673–688. doi: 10.1007/s10096-013-2009-3

Rönnqvist, D., Forsgren-Brusk, B., and EGrahn-Håkansson, E. (2006). Lactobacilli in the female genital tract in relation to other genital microbes and vaginal

pH. *Acta Obstet. Gynecol. Scand.* 85, 726–735. doi: 10.1080/000163406005
78357

Salas-Jara, M., Ilabaca, A., Vega, M., and García, A. (2016). Biofilm Forming
Lactobacillus: new Challenges for the Development of Probiotics.
Microorganisms 4:35. doi: 10.3390/microorganisms4030035

Stivala, A., Carota, G., Fuochi, V., and Furneri, P. M. (2021). Lactobacillus
rhamnosus AD3 as a promising alternative for probiotic products. *Biomolecules*
11:94. doi: 10.3390/biom11010094

Sturme, M. H. J., Nakayama, J., Molenaar, D., Murakami, Y., Kunugi, R., Fujii, T.,
et al. (2005). An agr-like two-component regulatory system in Lactobacillus
plantarum is involved in production of a novel cyclic peptide and regulation
of adherence. *J. Bacteriol.* 187, 5224–5235. doi: 10.1128/JB.187.15.5224-52
35.2005

Tortelli, B., Lewis, W., Allsworth, J., Member-Meneh, N., Foster, L., Reno, H., et al.
(2020). Associations between the vaginal microbiome and Candida colonization
in women of reproductive age. *Physiol. Behav.* 176, 139–148. doi: 10.1016/j.ajog.
2019.10.008.Associations

van den Nieuwboer, M., van Hemert, S., Claassen, E., and de Vos, W. M. (2016).
Lactobacillus plantarum WCFS1 and its host interaction: a dozen years after the
genome. *Microb. Biotechnol.* 9, 452–465. doi: 10.1111/1751-7915.12368

Van der Maaten, L., and Hinton, G. (2008). Multiobjective evolutionary algorithms
to identify highly autocorrelated areas: the case of spatial distribution in
financially compromised farms. *Ann. Oper. Res.* 219, 187–202. doi: 10.1007/
s10479-011-0841-3

van der Veer, C., Hertzberger, R. Y., Bruisten, S. M., Tytgat, H. L. P., Swanenburg, J.,
Angelino-Bart, A., et al. (2018). Comparative genomics of human Lactobacillus
crispatus isolates reveals genes for glycosylation and glycogen degradation:

implications for in vivo dominance of the vaginal microbiota. *bioRxiv*
[Preprint]. doi: 10.1101/441972

Ventolini, G. (2015). Vaginal lactobacillus: biofilm formation in vivo – clinical
implications. *Int. J. Womens. Health* 7, 243–247. doi: 10.2147/IJWH.S77956

Wang, S., Wang, Q., Yang, E., Yan, L., Li, T., and Zhuang, H. (2017). Antimicrobial
compounds produced by vaginal Lactobacillus crispatus are able to strongly
inhibit Candida albicans growth, hyphal formation and regulate virulence-
related gene expressions. *Front. Microbiol.* 8:564. doi: 10.3389/fmicb.2017.
00564

Check for
updates

The Polygonal Cell Shape and Surface Protein Layer of Anaerobic Methane-Oxidizing *Methylomirabilis lanthanidiphila* Bacteria

Lavinia Gambelli[1,2,3], Rob Mesman[1], Wouter Versantvoort[1], Christoph A. Diebolder[4†], Andreas Engel[5], Wiel Evers[5,6], Mike S. M. Jetten[1], Martin Pabst[7], Bertram Daum[2,8] and Laura van Niftrik[1]*

[1] Department of Microbiology, Faculty of Science, Radboud University, Nijmegen, Netherlands, [2] Living Systems Institute, University of Exeter, Exeter, United Kingdom, [3] College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, United Kingdom, [4] Netherlands Centre for Electron Nanoscopy (NeCEN), Leiden University, Leiden, Netherlands, [5] Department of Bionanoscience, Delft University of Technology, Delft, Netherlands, [6] Department of Chemical Engineering, Delft University of Technology, Delft, Netherlands, [7] Department of Biotechnology, Delft University of Technology, Delft, Netherlands, [8] College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom

Methylomirabilis bacteria perform anaerobic methane oxidation coupled to nitrite reduction via an intra-aerobic pathway, producing carbon dioxide and dinitrogen gas. These diderm bacteria possess an unusual polygonal cell shape with sharp ridges that run along the cell body. Previously, a putative surface protein layer (S-layer) was observed as the outermost cell layer of these bacteria. We hypothesized that this S-layer is the determining factor for their polygonal cell shape. Therefore, we enriched the S-layer from *M. lanthanidiphila* cells and through LC-MS/MS identified a 31 kDa candidate S-layer protein, mela_00855, which had no homology to any other known protein. Antibodies were generated against a synthesized peptide derived from the mela_00855 protein sequence and used in immunogold localization to verify its identity and location. Both on thin sections of *M. lanthanidiphila* cells and in negative-stained enriched S-layer patches, the immunogold localization identified mela_00855 as the S-layer protein. Using electron cryo-tomography and sub-tomogram averaging of S-layer patches, we observed that the S-layer has a hexagonal symmetry. Cryo-tomography of whole cells showed that the S-layer and the outer membrane, but not the peptidoglycan layer and the cytoplasmic membrane, exhibited the polygonal shape. Moreover, the S-layer consisted of multiple rigid sheets that partially overlapped, most likely giving rise to the unique polygonal cell shape. These characteristics make the S-layer of *M. lanthanidiphila* a distinctive and intriguing case to study.

Keywords: *Methylomirabilis*, NC10 phylum, anaerobic methane oxidation, S-layer, cell shape, cryo-tomography, sub-tomogram averaging

INTRODUCTION

"Candidatus Methylomirabilis" bacteria are anaerobic methanotrophs that oxidize methane to carbon dioxide coupled to the reduction of nitrite to dinitrogen gas (Ettwig et al., 2010). For a long time, this process was thought to be impossible due to the high activation energy needed to break the C-H bond of methane. However, in 2006 microbe-mediated nitrite-dependent anaerobic oxidation of methane (AOM) was discovered (Raghoebarsing et al., 2006). Nitrite-dependent AOM was

attributed to "*Candidatus* Methylomirabilis oxyfera," a bacterium belonging to the NC10 phylum (Raghoebarsing et al., 2006). Since their discovery, NC10 bacteria have been found in a variety of natural [freshwater (Raghoebarsing et al., 2006; Deutzmann and Schink, 2011; Chen et al., 2015; Yan et al., 2015; Graf et al., 2018), marine (He et al., 2015; Padilla et al., 2016), brackish (Zhang et al., 2018)] and man-made (wastewater treatment plants; Luesken et al., 2011; Ho et al., 2013) ecosystems, indicating that these bacteria might contribute significantly to the global carbon and nitrogen cycles.

Methylomirabilis bacteria are cultivated as flocculent enrichment cultures in bioreactor systems. Even though these bacteria grow anaerobically, the postulated pathway for nitrite-dependent methane oxidation still relies on oxygen. Based on genomic analysis and isotope labeling experiments, it was proposed that nitrite is converted to nitric oxide which is then dismutated by a putative NO dismutase (NOD) enzyme (Ettwig et al., 2010, 2012) to nitrogen and oxygen. The oxygen would then be used for methane activation by the particulate methane monooxygenase (pMMO) enzyme. In the end, methane is oxidized to carbon dioxide as end product (Ettwig et al., 2010).

An ultrastructural study of the first enriched and described *Methylomirabilis* species, *M. oxyfera* (Wu et al., 2012), showed that these microorganisms are also unique with respect to their cell morphology. *M. oxyfera* are ca. 300 nm wide and 1.5 μm long diderm bacteria with a polygonal cell shape. The outermost layer of the cell envelope contains multiple longitudinal ridges that end in a cap-like structure at the cell poles. This outermost layer of *M. oxyfera* was proposed to be a surface protein layer (S-layer), about 8 nm thick, of oblique or square symmetry (Wu et al., 2012). Since S-layers, have been described to promote and maintain the cell shape of bacteria, it was postulated that the S-layer of *M. oxyfera* is the responsible factor for its polygonal shape (Wu et al., 2012).

Besides the canonical shapes of cocci, bacilli, filaments and spirals, several unusual cell shapes have been described for bacteria and archaea. Some examples are the flat and square cells of the archaeon *Haloquadratum walsbyi* (Walsby, 1980), the star-shaped bacteria of the genus *Stella* (Vasilyeva, 1985) and the square or triangular archaeal cells of *Haloarcula* species (Nishiyama et al., 1992; Oren et al., 1999). It is often difficult to assess what the evolutionary advantage behind these unusual morphologies is. The most popular hypothesis is that microorganisms need to increase their surface area to volume ratio (SA:V) to balance nutrient uptake through diffusion and disposal of by-products. Any shape more complex than a sphere increases the SA:V compared to a sphere of the same volume (Purcell, 1977; Koch, 1996). However, maintaining a rod or filamentous shape (or any shape other than a sphere) requires the cell to counterbalance the surface tension and the cell osmotic pressure, which would force the cell into a sphere (Young, 2004). For this reason, the evolution of a robust cell envelope and a cytoskeleton are paramount (van Teeseling et al., 2017).

S-layers are an almost universal feature in the cell envelope of Archaea and widespread in Bacteria (Baumeister et al., 1988; Sleytr et al., 1999; Klingl, 2014). In most cases, S-layers consist of a single (glyco)protein that self-assembles into a 2-dimensional, paracrystalline array covering the entire cell surface. S-layer

lattices exist in oblique (p1, p2), square (p4), and hexagonal (p3, p6) symmetries, with a center-to-center spacing of the S-layer unit cell between 4 and 35 nm. A high percentage of the S-layer lattice (30–70%) is occupied by pores of varying diameters (Sleytr and Sára, 1997; Sleytr et al., 2014). S-layer proteins have been documented in a wide range of masses, generally between 40 and 200 kDa (Sleytr and Sára, 1997; Sleytr et al., 2014). These proteins are mostly weakly acidic and contain 40–60% hydrophobic amino acids (Sleytr and Sára, 1997). The bonds between the S-layer proteins are usually non-covalent and the attachment to the underlying cell envelope layer is a combination of weak bonds, such as hydrophobic, hydrogen and ionic bonds (Sleytr et al., 2014). However, in the case of diderm Bacteria and Archaea, the underlying cell envelope layer is the lipid bilayer of the outer membrane and of the cytoplasmic membrane, respectively. In monoderm bacteria, however, the S-layer is anchored to the cell wall peptidoglycan or to secondary cell wall polymers (Sleytr et al., 2014; Engelhardt, 2016). S-layer proteins undergo a number of post-translational modifications, among which glycosylation is the most frequent form (Sleytr et al., 2014). Many S-layer proteins are either N- or O- glycosylated, even though in a few cases both types of glycosylation are found on one protein (Schäffer et al., 2001). S-layers have been studied since 1953 (Houwink, 1953) and to date several S-layer proteins have been isolated and characterized. It has become apparent that S-layer proteins show a low degree of conservation on the level of the primary sequence (Sleytr et al., 2014).

S-layers fulfill several functions, most of which are still only hypothetical. Restricting the field to non-pathogenic microorganisms, S-layers can serve as protective coat, molecular sieve, molecule and ion trap; mediate cell adhesion and surface recognition and are involved in the determination of the cell shape (Baumeister et al., 1988; Engelhardt, 2007a,b; Sleytr et al., 2014).

In this study we characterized the cell shape and S-layer (composition and structure) of a recently enriched *Methylomirabilis* bacteria species, *M. lanthanidiphila* (Versantvoort et al., 2018), which similar to *M. oxyfera* exhibits a polygonal cell shape. To study the role of the S-layer in the determining the cell shape of *M. lanthanidiphila*, we used an array of biochemical and cryo-electron microscopy (cryo-EM) methods to enrich, identify, localize, and characterize the S-layer protein and lattice. We find that the S-layer has a hexagonal symmetry and is composed of a single repeating protein unit identified as mela_00855. As other S-layers, the *M. lanthanidiphila* one consisted of individual S-layer sheets. However, unlike other S-layers observed to date, the S-layer sheets of *M. lanthanidiphila* protrude from the outer membrane, detaching from the cell body, and giving rise to the ridges that likely determine the polygonal shape of the cell.

MATERIALS AND METHODS

Enrichment Conditions

M. lanthanidiphila (Versantvoort et al., 2018) was cultivated as a flocculent enrichment culture (∼70% *M. lanthanidiphila*) in a 10 L continuous sequencing batch reactor (Applicon

Biotechnology BV, Delft, the Netherlands), originally inoculated with sediment samples from a ditch in the Ooijpolder (the Netherlands) (Ettwig et al., 2009). The culture was kept anoxic by a continuous supply of a gas mixture composed of methane and carbon dioxide (95:5, v/v) and the medium was continuously flushed with a mixture of argon and carbon dioxide (95:5, v/v). The temperature was kept stable at 30°C and the system was stirred at 100 RPM. The bioreactor volume was kept at 10 L by a level sensor-controlled pump in sequential cycles of feeding (22.5 h) and resting (30 min to settle, 60 min to pump out excess medium). The HRT (hydraulic retention time) was 10 days and the composition of the medium was 30–40 mM $NaNO_2$ (depending on culture performance), 0.78 mM $MgSO_4$, 1.96 mM $CaCl_2$, and 0.73 mM KH_2PO_4 and the following trace elements: 5.4 μM $FeSO_4$, 0.26 μM $ZnSO_4$, 0.15 μM $CoCl_2 \cdot 6H_2O$, 2.82 μM $CuSO_4$, 0.24 μM $NiCl_2$, 0.07 μM H_3BO_3, 0.3 μM $MnCl_2$, 0.05 μM Na_2WO_4, 0.12 μM Na_2MoO_4, 0.14 μM SeO_2, and 0.12 μM $CeCl_2$, which was adapted from a previously published study (Ettwig et al., 2010).

The presence of *M. lanthanidiphila* was monitored during enrichment by fluorescence *in situ* hybridization (FISH) using a newly developed probe specific for *M. lanthanidiphila* (MLanth181, [TCCCATGAGATCCTCACAGG]), which targets bases 181–200 in the 16 s rRNA gene sequence (NCBI sequence ID CABIKM010000010.1:2618–4176), a region just upstream of the target of the DAMOBACT-0193 probe, which targets all known *Methylomirabilis* species. The MLanth181 probe has 6 mismatches with the *M. oxyfera* sequence (NCBI sequence ID FP565575: 1586103–1587651) as shown using pairwise alignment of the 16 S rRNA gene sequences of *M. oxyfera* and *M. lanthanidiphila*. No additional targets of the MLanth181 probe (Raghoebarsing et al., 2006) could be identified using the SILVA database (Quast et al., 2013).

Enrichment of the *M. lanthanidiphila* S-Layer Patches

The enrichment of S-layer patches was performed as described previously (van Teeseling et al., 2015). Approximately 60 ml of biomass were harvested from the *M. lanthanidiphila* enrichment culture and gently pottered on ice to disrupt the bacterial aggregates. Cells were concentrated by centrifugation (Allegra X-15R, Beckman Coulter) at 10,000 g for 20 min at 4°C. The supernatant was discarded, and the pellet was resuspended in 4 ml of medium. The sample was boiled in 4% SDS for 1 h. Subsequently, the sample was allowed to cool down and washed three times by ultracentrifugation (Beckman-Coulter Optima-XE-90 ultracentrifuge) at 138,550 g for 10 min at 25°C. Each time the pellet was thoroughly resuspended in 8 ml of MilliQ. To digest the peptidoglycan sacculi, a lysozyme treatment (final concentration 10 mg/ml) was performed overnight at 37°C with vigorous shaking. The lysozyme and peptidoglycan residues were removed from the sample by an additional ultracentrifugation step.

To isolate S-layer suitable for cryo-EM the isolation procedure was further optimized. A sample of 3 ml biomass from the *M. lanthanidiphila* enrichment culture was passed 10 times through a ball-bearing homogenizer (Isobiotech), using a ball

with 8 μm clearance, to obtain a homogeneous sample. This sample was boiled for 1.5 h in 4% SDS to liberate the S-layer. S-layers were pelleted and washed 3 times by ultracentrifugation as described above. The resulting pellet was resuspended in 200 μl MilliQ and dialyzed (50 kDA cutoff) against 300 ml MilliQ containing activated biobeads (Biorad) for 2 days at 4°C to adsorb the remaining SDS. Sample quality was tested by transmission electron microscopy using negative stain (2% uranyl acetate) on grids containing a continuous carbon film (200 mesh, copper).

LC-MS/MS for Identification of the *M. lanthanidiphila* S-Layer Protein and Antibody Generation

Three dilutions of the enriched S-layer patches (undiluted, diluted 1:10 and 1:100 with MilliQ) were processed by in-solution digestion to obtain peptides for LC-MS/MS. Each sample had a total volume of 100 μl. Samples were diluted 1:1 with 8 M urea in 10 mM Tris-HCl pH 8. 1 μl 50 mM DTT was added to the sample and incubated for 20 min at RT. Subsequently, 1 μl of 50 mM iodoacetamide was added and the sample was incubated for an additional 20 min in the dark. LysC enzyme (1 μl, 0.5 μg/μl stock) was added to the sample and incubated for 3 h at RT. Samples were then diluted threefold with 10 mM ammonium bicarbonate and trypsin (1 μl, 0.5 μg/μl stock) was added. Samples were incubated overnight at 37°C and stored at −80°C. Tryptic peptides were desalted and concentrated using C18 solid phase extraction (Omix tips, Agilent Technologies). Peptides were analyzed using nanoflow ultra-high pressure C18 reversed phase liquid chromatography coupled online to a quadrupole orthogonal time-of-flight mass spectrometer (maXis Plus, Bruker Daltonics) via a vacuum assisted axial desolvation nanoflow electrospray ionization source (Captivesprayer, Bruker Daltonics). Peptides were separated using a linear gradient of 5–45% acetonitrile in 0.1% acetic acid for 60 min at a flow rate of 500 nl/min. The mass spectrometer was programmed to acquire 1 MS spectrum at 3 Hz with subsequent data dependent MS/MS spectra at precursor intensity scaled acquisition speeds (3 Hz at 10,000 intensity up to 16 Hz at 150,000 intensity). Total MS + MS/MS duty cycle was 3 s. Dynamic exclusion was enabled to prevent re-analysis of the same precursor ions. Acquired data files were processed in DataAnalysis 4.2 (Bruker Daltonics) to extract MS/MS data for subsequent database searches. Proteins were identified using the MASCOT search tool (Matrix Science, London, United Kingdom) and an in-house protein sequence database of the predicted *M. lanthanidiphila* proteome. MASCOT search parameters included tryptic specificity, maximum of two missed cleavages, Carbamidomethyl (C) as fixed modification, and Oxidation (M) and Deamidation (NQ) as variable modifications. Precursor mass tolerance was set to 10 ppm and fragment ion mass tolerance to 0.05 Da. A single protein was identified by the MS/MS database search which was mela_00855 (NCBI sequence ID: VUZ84482.1). The full protein sequence was handed over to the company GenScript. A protein fragment (33–260 aa) of mela_00855 was synthesized and heterologously expressed in *E. coli*. The protein fragment was then used to

raise antibodies in rabbit. The pre-immune serum and the final bleed (crude antiserum) were used for immunoblotting and immunogold localization.

Cell-Free Extract, PAGE and Immunoblotting

A sample of 500 ml was harvested from the *M. lanthanidiphila* bioreactor and centrifuged at 20,000 g for 10 min at 4°C in a Sorvall centrifuge (Sorvall Lynx 4,000). The pellet was resuspended in 60 ml of 20 mM potassium-phosphate (KPi) buffer at pH 7. The cells were centrifuged again at 20,000 g for 10 min at 4°C. The pellet was resuspended in 60 ml of 20 mM KPi with 2.5 mM EDTA at pH 7 and six protease inhibitor tablets (Pierson Protease Inhibitor, Mini tablets, EDTA-free, Thermo Scientific). Cells were sonicated on ice for 10 min, 5 s on and 5 s off. To remove debris and unbroken cells, the sonicated sample was centrifuged at 10,000 g for 10 min at 4°C. The supernatant of lysed cells was used for further PAGE and immunoblotting.

The cell-free extract was boiled for 7 min in SDS sample buffer (50 mM Tris-HCl buffer pH 6.8 containing 5% β-mercaptoethanol, 2% SDS, 10% glycerol and 0.005% bromophenol blue), and 20 μg protein per lane was loaded onto 4–15% Criterion TGX precast gels (Bio-Rad) for PAGE according to manufacturer's instructions. After PAGE separation, the proteins were transferred from the gel onto a Trans-blot Turbo, Midi format 0.2 μm nitrocellulose transfer membrane (Bio-Rad) with the Turbo blotter system (Bio-Rad) according to manufacturer's instructions. The blotting was performed at 2.5 A, 25 V for 7 min. Dried blots were stored at 4°C.

Prior to starting the immunoblot protocol, blots were kept at RT for 30 min and then incubated in MilliQ water for an additional 30 min. Blocking was performed for 1 h in 1% BSA in 10 mM TBS (10 mM Tris-HCl, 137 mM NaCl, 2.7 mM KCl, pH 7.4). The blots were then incubated for 60 min in α-mela_00855 antiserum diluted 125-, 500- or 1,000-fold in blocking buffer. The negative controls were incubated in the pre-immune serum diluted 125-fold or incubated in blocking buffer without primary antibody. The blots were then washed three times for 10 min in TBS containing 0.05% Tween and incubated for 60 min in monoclonal mouse anti-rabbit IgG alkaline phosphatase conjugate (Sigma-Aldrich) diluted 150,000-fold in blocking buffer. The blots were washed two times for 10 min in TBS containing 0.05% Tween and two times for 10 min in TBS. Finally, blots were incubated with a 5-bromo-4-chloro-3-indolylphosphate (BCIP)/nitroblue tetrazolium (NBT) liquid substrate system (Sigma-Aldrich) for 5 min and rinsed for 10 min in MilliQ water.

Immunogold Localization of the S-Layer Protein in *M. lanthanidiphila* Ultrathin Sections

M. lanthanidiphila cells were harvested from the bioreactor and cryo-fixed by high-pressure freezing (Leica HPM 100, Leica Microsystems, Vienna, Austria). Samples were placed into a 100-μm cavity of a type-A platelet (3 mm diameter; 0.1–0.2-mm depth, Leica Microsystems) and closed with the flat side of a lecithin-coated type B platelet (3-mm diameter; 0.3-mm depth). The frozen samples were stored in liquid nitrogen.

For Lowicryl HM20 embedding, frozen samples were freeze-substituted in 0.2% uranyl acetate in acetone. The substitution started at −90°C for 48 h; brought to −70°C at 2°C per hour and kept at −70°C for 12 h; brought to −50°C at 2°C per h and kept at −50°C for 12 h in a freeze-substitution unit (AFS2; Leica Microsystems, Vienna, Austria). To remove unbound uranyl acetate, the samples were washed twice with 100% acetone for 30 min at −50°C. Keeping the temperature stable at −50°C, the sample was infiltrated with a graduated series of Lowicryl HM20 (10, 25, 50, and 75%) in acetone. Each step was 1 h long. Three final infiltration phases were performed with 100% Lowicryl HM20: first for 1.5 h, then overnight and finally 2 h. Polymerization of the resin was obtained by irradiating the sample with UV light for 120 h at −50°C after which the temperature was brought to 0°C in 24 h at 2.1°C per hour. UV light was switched off and the temperature was increased to 20°C in 5 h at 4°C per hour. Ultrathin sections of 55 nm were cut using a Leica UCT ultramicrotome (Leica Microsystems, Vienna, Austria) and collected on carbon-Formvar-coated 100 mesh copper grids (Agar Scientific).

Grids containing ultrathin sections were rinsed for 10 min in 0.1 M PHEM (1X PHEM contains: 60 mM PIPES, 25 mM HEPES, 10 mM EGTA, and 4 mM $MgSO_4 \cdot 7H_2O$) buffer pH 7 and blocked for 20 min in 1% BSAc (Aurion) in PHEM buffer. Grids were incubated for 60 min at RT with the primary antiserum targeting the S-layer protein diluted 1:400 in blocking buffer. Negative controls were incubated for 60 min in blocking buffer without primary antiserum and with the pre-immune serum diluted 1:400. After this incubation, the grids were washed for 10 min in 0.1% BSAc in 0.1 M PHEM buffer and incubated for 30 min with protein A coupled to 10 nm gold particles (PAG 10, CMC UMC Utrecht), diluted 1:60 in blocking buffer. The grids were then washed again first in 0.1% BSAc in 0.1 M PHEM buffer for 5 min and then in 0.1 M PHEM buffer for 5 min. To fix the labeling, the grids were incubated for 5 min in 1% glutaraldehyde in 0.1 M PHEM buffer and subsequently washed for 10 min in MilliQ. Post-staining was performed by incubating for 20 min on drops of 2% uranyl acetate, after which the grids were quickly washed in 4 drops of MilliQ and allowed to air dry. The grids containing labeled ultrathin sections were investigated at 60 kV in a JEOL (Tokyo, Japan) JEM-1010 TEM.

Immunogold Localization of the *M. lanthanidiphila* S-Layer Patches

A sample of 3.5 μl of enriched S-layer patches (obtained as described in the section "Enrichment of the *M. lanthanidiphila* S-layer patches") was spotted on carbon-Formvar-coated 200 mesh copper grids (Agar Scientific) and incubated for 15 min at RT under moisture-controlled conditions. Subsequently the grids were blotted to remove excess sample and were allowed to air dry overnight. Prior to starting the immunogold localization protocol, the grids were heated at 120°C for 20 min. The grids were rinsed for 6 min on drops of 0.1 M PHEM buffer pH 7 and blocked with 1% BSAc (Aurion) in 0.1 M PHEM buffer pH 7
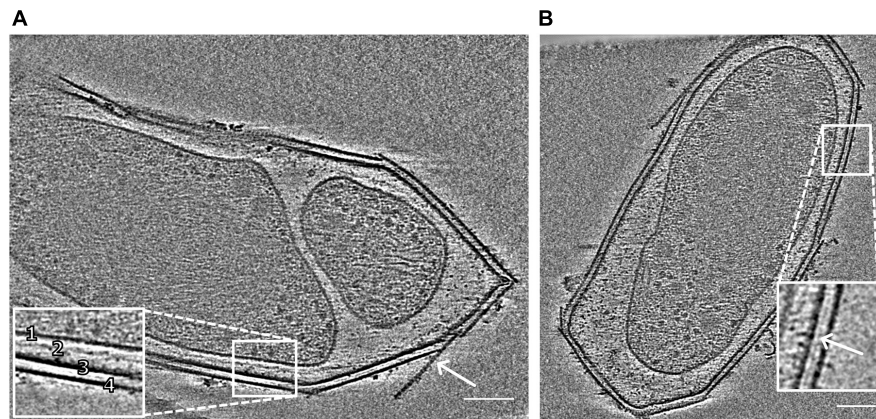
FIGURE 1 | Snapshots of electron cryo-tomograms of two *M. lanthanidiphila* cells. The inset in **(A)** shows the cell envelope (1, cytoplasmic membrane; 2, peptidoglycan layer; 3, outer membrane; 4, S-layer). The peptidoglycan layer is not always visible. The arrow shows an S-layer sheet protruding from the cell. The inset in **(B)** shows a layer (white arrowhead) between the outer membrane and the S-layer. Scale bars 100 nm.

for 5 min. The incubation with primary antiserum targeting the S-layer protein was performed with a 1:40 dilution in blocking buffer for 25 min. Negative controls were incubated for 25 min in blocking buffer without primary antiserum and with the pre-immune serum diluted 1:40. Grids were rinsed for 6 min in 0.1 M PHEM containing 0.1% BSAc and then incubated for 30 min with the protein A coupled to 10 nm gold particles (PAG 10, CMC UMC Utrecht), diluted 1:60 in blocking buffer. To remove BSAc, the grids were washed in 0.1 M PHEM and fixed with 1% glutaraldehyde in 0.1 M PHEM buffer. The glutaraldehyde was removed with 10 washes with MilliQ. The labeled S-layer patches were stained with 2% uranyl acetate in 0.1% acetic acid for 1 min. The excess staining solution was blotted off and grids were allowed to air dry. The grids containing labeled negative-stained S-layer patches were investigated at 60 kV in a JEOL (Tokyo, Japan) JEM-1010 TEM.

Cryo-Tomography
Whole Cell Cryo-ET

For the whole cell cryo-ET (**Figures 1**, **2** and **Supplementary Movies 1, 2**), a sample of 15 ml was collected from the *M. lanthanidiphila* enrichment culture and centrifuged (Beckman Coulter) for 2 min at 5,000 g. The pellet was resuspended in 1 ml of bioreactor medium. Cells were bead-beaten using glass beads (0.5 mm, BioSpec) for 1 min at 50/s oscillation in a tissuelyser (Tissuelyser LT, Quiagen) to break the cell aggregates. To collect the single cells, the sample was pulse-centrifuged in a tabletop microcentrifuge. The supernatant was diluted 1:2 with reactor medium until an OD600 of 0.45.

Prior to sub-tomogram averaging of S-layers *in situ* (**Figure 3**), the samples were reproduced, under slightly different conditions as before. 3 ml biomass from the *M. lanthanidiphila* bioreactor was passed 3 times through a ball-bearing homogenizer (Isobiotech) using a ball with 8 μm clearance, to break the cell aggregates.

For both datasets, ∼2 μl containing *M. lanthanidiphila* single cells and 0.5 μl 10 nm protein A gold (CMC, UMC Utrecht,

the Netherlands) were applied on a glow-discharged Quantifoil type R2/2 copper 200-mesh grids. The cells were frozen in liquid ethane using a Vitrobot Mk4 (FEI, Thermo Fisher Scientific) with controlled humidity (100%) using ablotforce of 1 and a blot time of 3 s. Tilt series were recorded at the Netherlands Center for Electron Nanoscopy (NeCEN, Leiden, the Netherlands) on a Titan Krios (FEI, Thermo Fisher Scientific) microscope at 300 kV equipped with a Schottky type field emission gun and a K3 camera (Gatan) operated in counting mode.

Tilt series were acquired using UCSF Tomo (Zheng et al., 2007). Initially, a linear tilt scheme from −60° to + 60° with a 2° increment at −6 μm defocus and a magnification of 26,000× (pixel size of 4.6 Å) was employed. To improve data quality for sub-tomogram averaging, data were collected at 33,000× magnification (pixel size of 2.64 Å) and dose-symmetric tilt scheme (Hagen et al., 2017) from −60 to + 60 with a 2° increment at −4 μm defocus was employed using SerialEM (Mastronarde, 2005). The total dose on the record area did not exceed 100 e-/Å2 and varied between 93 and 96.5 e^{-}/Å^{2}. Tomograms were reconstructed and CTF corrected in IMOD (Kremer et al., 1996) and binned threefold (final pixel size of 9.267 Å for whole cell cryo-ET; pixel size of 7.92 Å, for sub-tomogram averaging). Tomograms were either reconstructed using SIRT (simultaneous iterative reconstruction technique) to optimize contrast or weighted back projection for sub-tomogram averaging. The whole cell movies (**Supplementary Movies 1, 2**) were made in IMOD.

Cryo-ET of Isolated S-Layers and Sub-Tomogram Averaging

For cryo-ET of isolated S-layers and sub-tomogram averaging (**Figure 4**), isolated S-layers were obtained as described above (section "Enrichment of the *M. lanthanidiphila* S-layer patches"). Before plunge freezing, potential aggregates in the sample were pelleted and 2 μl of the supernatant was mixed with 0.5 μl 10 nm ProteinA gold solution (CMC, UMC Utrecht) on glow discharged Quantifoil type R2/2 copper 200-mesh grids. Grids were plunge

FIGURE 2 | Snapshots of a model of a *M. lanthanidiphila* cell (based on the electron cryo-tomogram in **Figure 1B**). The S-layer patches are depicted in shades of blue, the outer membrane in magenta, and the cytoplasmic membrane in gold. The complete cell envelope (S-layer, outer membrane, and cytoplasmic membrane) is shown in **(A1)**, the S-layer and the outer membrane are shown in **(B1)**, and the S-layer only (with arrows pointing at different patches) is shown in **(C1)**. **(A2,B2,C2)** are different views of **(A1,B1,C1)**, respectively.

frozen in liquid ethane using a Vitrobot Mk4 (FEI/Thermo Fisher Scientific) with controlled humidity (100%) using a blot force of 2, and a blot time of 2.5–3 s. Tilt series were recorded at the Netherlands Centre for Electron Nanoscopy (NeCEN, Leiden, Netherlands) using a Titan Krios (FEI, Thermo Fisher Scientific) microscope at 300 kV equipped with a Schottky type field emission gun and a K3 camera (Gatan) operated in counting mode. The tilt series were acquired in low-dose using SerialEM (Mastronarde, 2005), employing a dose-symmetric tilt-scheme (Hagen et al., 2017) from −60° to + 60° with a 2° increment at −4 µm defocus. The total dose on the record area did not exceed 100 e-/Å2. Recorded tilt series were reconstructed using IMOD

(Kremer et al., 1996), and tomograms were generated using the weighted back-projection algorithm.

Sub-Tomogram Averaging

To obtain the structure of the S-layer *in situ* and in its isolated form, sub-volumes were initially selected by applying a random grid of points over the S-layer patches in 3Dmod (IMOD software package). Using PEET (Nicastro et al., 2006; Heumann et al., 2011), sub-volumes were extracted, aligned, and averaged. A total of 8,938 sub-volumes of $120 \times 120 \times 60$ pixels in dimension were used for averaging the isolated S-layer patches and 1,330 sub-volumes for obtaining the average from the S-layer *in situ*. P6 symmetry was applied to the final average obtained from unbinned and unfiltered tomograms. The final S-layer average was visualized and segmented using UCSF Chimera (Pettersen et al., 2004). A resolution of 21 Å nm of the average was estimated based on the reflections in its power spectrum calculated by IMOD (Kremer et al., 1996). **Supplementary Movie 3** was made using UCSF Chimera.

S-Layer Sequence Coverage and Protein Glycosylation Analysis

In-Solution Digestion of Isolated S-Layer Protein Enrichment

A volume of 50 µl of S-layer enrichment sample (approx. 0.1 mg/ml protein content) was diluted 1:1 with 8 M Urea in 10 mM Tris-HCl buffer pH 8.0. To this solution, 1 µl of freshly prepared 50 mM DTT solution was added and incubated at 37°C for 1 h. Next, 1 µl of a freshly prepared 50 mM IAA (iodoacetamide) solution was added and incubated in the dark for 30 min. The solution was diluted to below 1 M Urea using 200 mM bicarbonate buffer and one aliquot protein each were digested using sequencing grade Trypsin or Chymotrypsin, respectively (Promega) at 37°C overnight (protease/protein of approx. 1:50). Finally, the protein digests were desalted using an Oasis HLB 96 well plate (Waters) according to the manufacturer protocols. The purified peptide fraction was speed-vac dried and stored at −20°C until further processed.

Shotgun Proteomic Analysis

The speed-vac dried peptide fraction was resuspended in 3% acetonitrile and 0.1% formic acid. An aliquot corresponding to approx. 250 ng protein digest was analyzed using a one-dimensional shotgun proteomics approach. Thereby, 1 µl of the protein digest was analyzed using a linear gradient from 5 to 30% B over 40 min, and further to 60% B over 15 min, maintaining a flow rate of 350 nL/min. MS1 level scans were performed over the mass range from 385 to 1250 m/z at 70 K resolution with an AGC target of 3 e6 and a max IT of 100 ms. Top10 DDA fragmentation spectra were acquired at 17.5 K resolution, with an AGC target of 2e5, a max IT of 54 ms and by using a NCE of 28. Unassigned, singly charged as well as 6, 8, and > 8 charged mass peaks were excluded.

Database Search and Data Processing

Raw data were analyzed using PEAKS Studio X (Bioinformatics Solutions Inc., Canada) allowing 20 ppm parent ion and
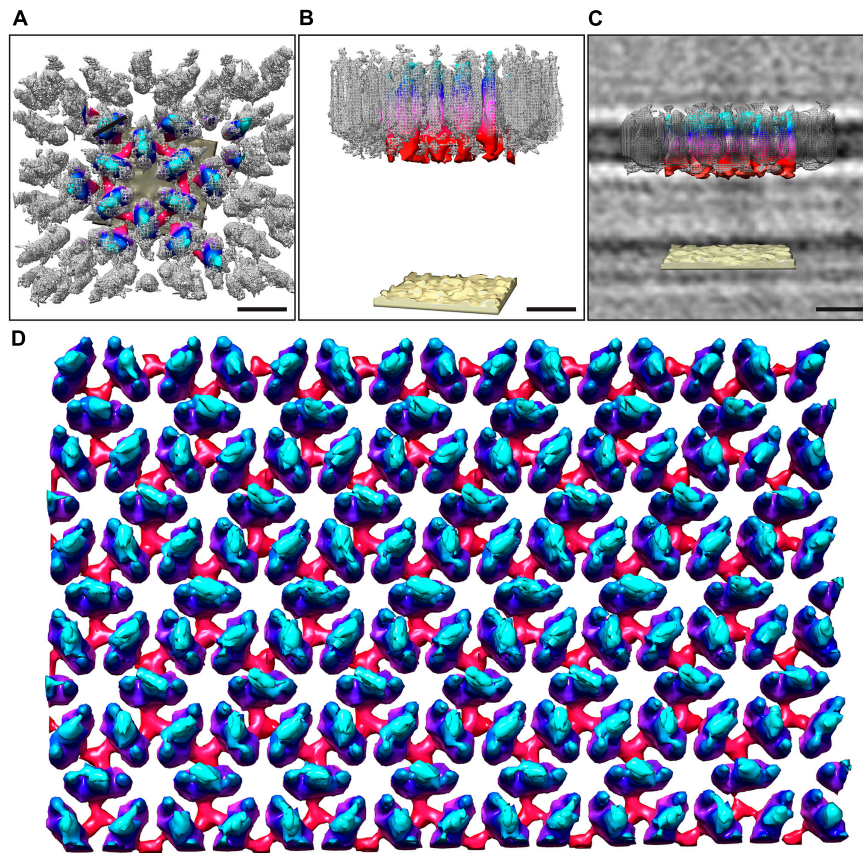
FIGURE 3 | Overlay of the sub-tomogram average obtained from isolated S-layer patches and whole-cell electron cryo-tomography. The map obtained from the whole cell cryo-tomogram is colored by proximity to the cell outer membrane (cyan, membrane-distal; red, membrane-proximal; sand, outer membrane), whereas the sub-tomogram average obtained from isolated S-layer patches is in gray mesh. Both the outward-facing **(A)** and the side views **(B)** show a good correlation between the two models. **(C)** Shows the side view with the median filtered S-layer model after scaling the height of both models to fit the height of the S-layer as observed in the cryo-tomograms. Visual rendering **(D)** of an S-layer sheet after median filtering (6 iterations). The height of the S-layer has been scaled to 7 nm as observed from perpendicular cross-sections in the cryo-tomograms. Scale bars 10 nm.

0.02 Da fragment mass error tolerance. Search conditions further included considering 4 missed cleavages, carbamidomethylation as fixed and methionine oxidation and N/Q deamidation as variable modifications. Semi-specific cleavages were allowed. Data were matched against a *M. lanthanidiphila* in-house annotated protein database (Genbank, VUZ84482.1). Database search included the GPM crap contaminant database[1] and a decoy fusion for determining false discovery rates. Peptide spectrum matches were filtered against 1% false discovery rate (FDR), or a −10lgP peptide score of −41.3, respectively.

Sequence Analyses

The sequence alignment was performed with the Praline server (Heringa, 1999).[2]

The sequence analysis for glycosylation sites was performed with the GlycoPP v1.0 server (Chauhan et al., 2012).[3]

---

[1]https://www.thegpm.org/crap/

[2]https://www.ibi.vu.nl/programs/pralinewww/

[3]http://crdd.osdd.net/raghava/glycopp/

Data Deposition

The sub-tomogram averaging electron density maps of the whole cells and isolated S-layer have been deposited in EMDB with accession code EMD-13672 and EMD-13670, respectively. Raw cryo-tomography data have been deposited in EMPIAR with accession code EMPIAR-10822 for the isolated S-layer patches and EMPIAR-10829 for the whole cells. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium[4] via the PRIDE partner repository with the dataset identifier PXD029319.

RESULTS

Cell Structure of *M. lanthanidiphila* Bacteria

The cell shape and cell plan of *M. lanthanidiphila* were investigated by cryo-ET on whole cells. Our data revealed the

---

[4]http://www.proteomexchange.org/
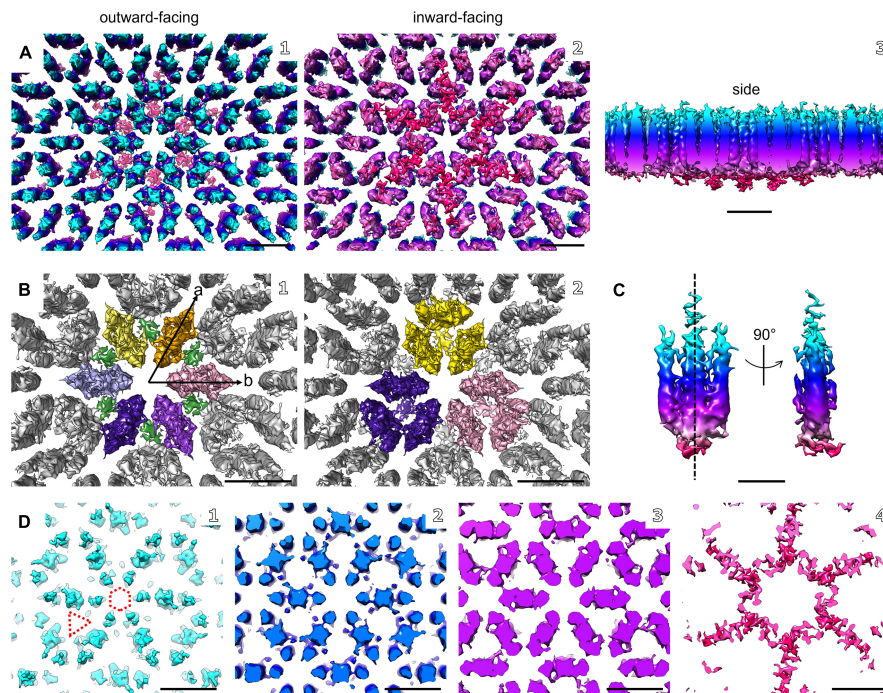
FIGURE 4 | Electron cryo-tomography and sub-tomogram averaging of *M. lanthanidiphila* S-layer. **(A)** Sub-tomogram average of isolated *M. lanthanidiphila* S-layer patches. Maps are colored by proximity to the cell outer membrane (cyan, membrane-distal; red, membrane-proximal), as in **(C,D)**. **(B)** in **(B1)** putative dimers of mela_00855 as the S-layer building block are highlighted in different colors, and the densities projecting toward the outer membrane are in green. The unit cell dimensions are 11.9 × 11.6 nm (a and b arrows, respectively) including an angle of 60°. **(B2)** shows the staggered conformation of the repeating unit-cell within the S-layer. **(C)** shows the isolated putative dimer of mela_00855. The structure is symmetrical along the dashed line. **(D)** slices through the sub-tomogram average from membrane-distal **(D1)** to membrane-proximal **(D4)**. The dashed hexagon and triangle in **(D1)** highlight the hexagonal and triangular pores. Scale bar in **(A,B, and D)** is 10 nm; scale bar in **(C)** is 5 nm.

overall polygonal shape of the *M. lanthanidiphila* cell (**Figure 1**; **Supplementary Movies 1, 2**). The S-layer was clearly discernable as a distinct layer ∼ 11 nm above the outer membrane. Instead of adhering to the cell body as usually seen in other microorganisms (van Teeseling et al., 2014; Bharat et al., 2017; Gambelli et al., 2019; Gaisin et al., 2020; von Kügelgen et al., 2020; Oatley et al., 2020. For examples of reviews see Albers and Meyer (2011); Pavkov-Keller et al. (2011), Rodrigues-Oliveira et al. (2017), and Pum et al. (2021), the S-layer of *M. lanthanidiphila* formed several distinct patches. These patches had the appearance of planar sheets and intersected at sharp ridges, thereby likely defining the polygonal shape of the cell. The distance between the S-layer and the outer membrane remained constant, indicating that the S-layer proteins are anchored to the outer membrane, but not to the cytoplasmic membrane that did not have a polygonal shape. The peptidoglycan layer also did not have a polygonal shape. In intact cells (**Figure 1A** and **Supplementary Figure 1**) the peptidoglycan appeared to slightly protrude toward the edges of the polygon, without however strictly assuming a polygonal shape. Furthermore, in isolated peptidoglycan sacculi (**Figure 5A**) the peptidoglycan was round. An additional electron dense layer was present between the outer membrane and the S-layer in all observed *M. lanthanidiphila* cells (**Figure 1B**, inset). At the ridges, the S-layer sheets protrude, overlapping with each other, and partially detaching from the

outer membrane (**Figure 1A**, arrow). The inset in **Figure 1A** shows a snapshot of the cell envelope elements: cytoplasmic membrane, peptidoglycan layer, outer membrane, and S-layer. The segmentation of the *M. lanthanidiphila* cell in **Figure 1B** showed that the S-layer was arranged around the cell in multiple sheets of different sizes that partially overlapped with each other (**Figure 2**).

Identification of the *M. lanthanidiphila* S-Layer Protein

Several established protocols were extensively tested and optimized to enrich or isolate the *M. lanthanidiphila* S-layer protein. Specifically, we changed the pH of the solution (low pH, acid treatment), added detergent (Triton X-100), denaturing agents (urea, guanidine hydrochloride, lithium chloride), or chelating agents (EDTA). In almost all cases, the S-layer was released from the cell, but could not be disassembled and did not show any protein bands on SDS-PAGE. Therefore, *M. lanthanidiphila* cells were boiled in 4% SDS and subsequently ultracentrifuged to enrich S-layer patches. The patches were digested in-solution and analyzed for protein content by LC-MS/MS. Only one protein was retrieved that matched to the predicted proteome of *M. lanthanidiphila*. This protein was mela_00855 (gene length 950 bp, protein length 316 aa, MW
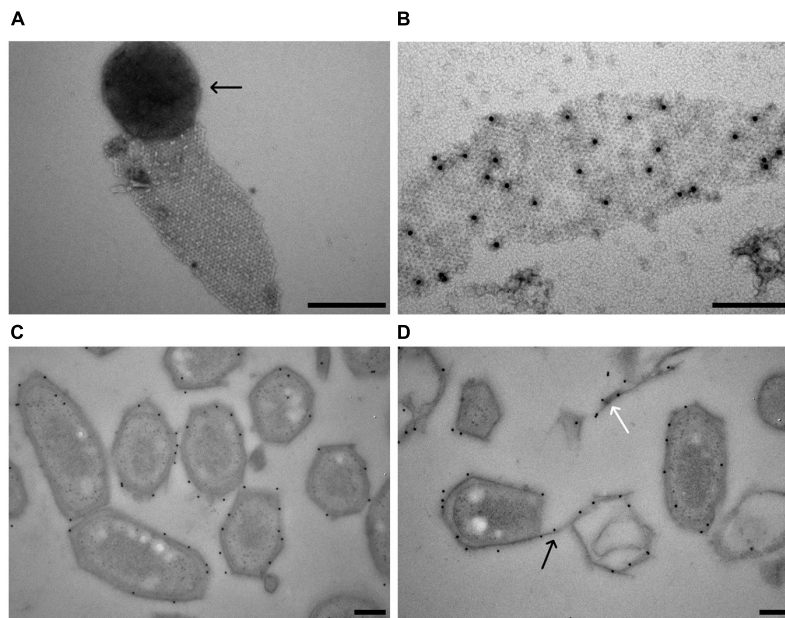
FIGURE 5 | Localization of the *M. lanthanidiphila* S-layer protein. **(A)** Transmission electron micrograph of a negative-stained *M. lanthanidiphila* S-layer patch. A round peptidoglycan sacculus (arrow) is still attached to the S-layer patch. **(B)** Transmission electron micrograph of negative-stained *M. lanthanidiphila* S-layer patch with immunogold labeling (black dots) of the S-layer protein mela_00855. **(C,D)** Transmission electron micrographs of high-pressure frozen, freeze-substituted and Lowicryl-embedded *M. lanthanidiphila* cells showing immunogold localization of the S-layer protein mela_00855. In **(D)** the labeling is present also on the S-layer edges protruding from the cell (black arrow) and on shed S-layer patches (white arrow). Scale bars 200 nm.

31.6 kDa, pI 8.42, **Supplementary Table 1**). This protein was highly transcribed; 1,539.33 RPKM value, which was in the range of other important proteins such as 1,328.26 for PmoB beta subunit and 6,254.36 for the putative NO dismutase (Nod2, mela_02434) (unpublished data). The protein mela_00855 was annotated as hypothetical protein based on Blastp analysis. The only protein that matched mela_0855 in Blastp analysis was an *M. oxyfera* protein annotated as "exported protein of unknown function" (NCBI sequence ID: CBE67388). The two proteins shared 40.88% protein sequence identity and 97% coverage (**Supplementary Figure 2**). The amino acid sequence of mela_00855 (**Supplementary Table 1**) was analyzed for the presence of conserved regions consulting InterProScan (Jones et al., 2014). The protein sequence had two transmembrane regions, one being a predicted N-terminal signal peptide (1–34 aa), and the other a predicted C-terminal IPTL-CTERM protein sorting domain (285–312 aa).

To confirm the identity of the *M. lanthanidiphila* S-layer protein, antibodies were raised against a synthesized protein fragment (from 33 to 260 aa of mela_00855). The affinity and specificity of the crude antiserum was tested and confirmed by immunoblot analysis using *M. lanthanidiphila* cell extracts (**Supplementary Figure 3**). The S-layer sheets were isolated from *M. lanthanidiphila* cells (**Figure 5A**) and the mela_00855 antiserum was used to localize the S-layer on negative-stained enriched S-layer patches and thin-sections of resin-embedded *M. lanthanidiphila* cells. The immunogold localization on negative-stained enriched S-layer patches showed abundant and specific binding of the antiserum to the patches (**Figure 5B**).

The immunogold localization on thin sections resulted in the specific labeling of the *M. lanthanidiphila* cell envelope and it was particularly abundant in the area occupied by the outer membrane and the S-layer (**Figures 5C,D**). Labeling was also present on edges protruding from the bacterial cell wall and on shed S-layer sheets (**Figure 5D**).

Structure of the *M. lanthanidiphila* S-Layer

The structure of the *M. lanthanidiphila* S-layer was investigated by cryo-ET and sub-tomogram averaging of enriched S-layer patches (**Figure 4**, **Supplementary Figure 4**, and **Supplementary Movie 3**). The final map was achieved by averaging 8,938 sub-volumes (**Figure 4A**) and had a resolution of 21 Å based on the Fast Fourier Transform (FFT) power spectrum (see **Supplementary Figure 5** for both FFT and FSC). Further, analysis of the FFT power spectrum of the averaged S-layer volumes showed that the S-layer had a hexagonal, p6 symmetry, and the unit cell dimensions were ∼11.6 × ∼11.9 nm including an angle of 60° (**Figure 4B1**). A single S-layer building block is shown in **Figure 4C**. This density showed a fork-like appearance, with three extracellular protrusions that join into a single stalk at the inward-facing side of the lattice. When observed in cross sections (**Figure 4D**) the S-layer building block appeared to have two-fold symmetry with respect to its long axis, which may suggest that the subunit is a dimer. Three S-layer putative dimers constitute the repeating unit cell of the S-layer (**Figure 4B2**). The S-layer lattice contained two main pore

types: a ∼12 nm^2 hexagonal pore surrounded by six units, and a ∼3.9 nm^2 triangular pore surrounded by three units (**Figure 4D1**). On the inward-facing side of the S-layer, the three protein densities surrounding a triangular pore converged and connected at the center of it. As each of the triangular pores has six triangular neighbors and each dimer-like unit is shared between two neighboring triangular pores, the entire S-layer is interconnected. The S-layer patches were roughly perpendicular to the electron beam during data acquisition and, therefore, the map was stretched in Z due to the missing wedge. From the cryo-tomograms of whole *M. lanthanidiphila* cells it could be determined that the S-layer has a thickness of ∼7 nm. For visualization purposes a median filter (6 iterations) was applied to the obtained average and the Z-scaling adjusted to match the thickness obtained from the cryo-tomograms. Multiple copies of the average were aligned to simulate the structure of the S-layer sheet (**Figure 3D**).

To determine the orientation of the obtained S-layer model with respect to the cell envelope, cryo-ET and subsequent sub-tomogram averaging was performed on whole *M. lanthanidiphila* cells. The *in situ* average of the S-layer confirmed that the fork-like protrusions of the S-layer units face the extracellular environment, while the densities that connect the units with each other face the outer membrane (**Figures 3A–C**).

S-layer proteins are usually highly glycosylated. To investigate, if this is also the case for mela_00855, we analyzed its primary sequence for N- or O-linked glycosylation sequons using GlycoPP v1.0 (Chauhan et al., 2012). The server predicted 46 putative O-linked and 4 putative N-linked glycosylated sites (**Supplementary Table 1**). To confirm this prediction, we performed an in-depth characterization of the purified S-layer by high-resolution shotgun proteomics (**Supplementary Figure 6**). Apart from the C-terminal tail, the complete amino acid sequence could be covered with peptide fragments. However, interestingly no indications for N or O-linked glycans could be found.

DISCUSSION

We analyzed the cell shape and S-layer structure of *M. lanthanidiphila* to understand the origin of its polygonal cell shape. To this end, we enriched, identified, localized, and characterized the S-layer protein and lattice. The building block of the S-layer was identified as the protein mela_00855. This protein assembled in planar sheets of hexagonal (p6) symmetry. The S-layer sheets partially overlapped and entirely covered the outer membrane. The polygonal shape was clearly observed in both the S-layer and the outer membrane, but not in the peptidoglycan and cytoplasmic membrane, leading us to conclude that the S-layer is very likely responsible for the polygonal shape of this bacterium. This S-layer is highly stable and resilient. In fact, we were unable to break the S-layer patches down to its monomers. Extensive attempts using different protocols were made to disrupt the S-layer protein-protein interactions, but always resulted in the enrichment of S-layer patches and not of single subunits visible on SDS-PAGE. Also in other instances, S-layers have been shown to be impervious

to disassembly, for example archaea of the *Thermoproteus* (Messner et al., 1986), *Pyrobaculum* (Phipps et al., 1990, 1991) and *Staphylothermus* (Peters et al., 1995, 1996) genera. However, this resilience is unusual for bacterial S-layers and indicates that this lattice is highly stable, suggesting that the protein-protein interaction occurring between these S-layer proteins can be very strong and possibly of kinds not usually found in interprotein interactions in S-layer lattices, such as covalent bonds.

The S-layer is constituted by several mostly straight patches, distributed with different orientations in order to cover the cell entirely. Moreover, the sheets protrude from the outer membrane toward the extracellular environment still maintaining their orientation. From these observations we concluded that these S-layer sheets are very rigid. A dense network of protein bridges links the S-layer morphological units. These proteinaceous connections are present over every triangular pore throughout the S-layer lattice. We hypothesize that the dense net of protein bridges, together with perhaps stronger protein-protein interactions, establishes the observed rigidity of the S-layer. S-layers are typically highly interconnected structures, and protein connections between the morphological units have been observed in other lattices (Phipps et al., 1990; Pum et al., 1991; Lupas et al., 1994; Arbing et al., 2012; Bharat et al., 2017; Gambelli et al., 2019). Future experiments to probe the rigidity and elasticity of this S-layer could include Atomic Force Microscopy (AFM) to analyze the binding forces between S-layer proteins both on the cell and on enriched S-layer patches (Toca-Herrera et al., 2004; Martín-Molina et al., 2006; Scheuring and Dufrene, 2010; Viljoen et al., 2020).

As opposed to the cytoplasmic membrane and peptidoglycan, the outer membrane followed the polygonal cell shape and cryo-ET revealed an additional layer between the S-layer and outer membrane. A similarly weak electron density between the S-layer and the outer membrane has been observed in *C. crescentus*, another Gram-negative bacterium (Bharat et al., 2017). In this bacterium, the S-layer protein RsaA is attached to the outer membrane via the N-terminus region of the S-layer protein that interacts with the LPS (Ford et al., 2007; von Kügelgen et al., 2020). We thus hypothesize that a similar mechanism could mediate the attachment of the *M. lanthanidiphila* S-layer protein to the outer membrane, via the N-terminus region of mela_00855.

Based on the observations that the outer membrane and the S-layer are the only polygonal components in the cell wall of *M. lanthanidiphila*, we concluded that the S-layer is likely to be responsible for the unusual cell shape of these bacteria. However, additional elements could contribute to the cell shape determination. Particularly, protein filament systems have been shown to be widespread in prokaryotic cells and play a major role in cell shape determination (Wagstaff and Löwe, 2018). Protein filaments such as crescentin and bactofilins are localized in the cytoplasm and are membrane-bound (Ausmees et al., 2003; Deng et al., 2020). Because the inner membrane of *M. lanthanidiphila* does not follow the sharp angles of the outer membrane and S-layer, we exclude that cytoplasmic elements have a structural role in the cell shape maintenance of these bacteria. However, other cytoskeletal elements of still

unknown nature could function as a scaffold for the S-layer.S-layer proteins in different microorganisms share little to no similarity (Engelhardt and Peters, 1998; Bharat et al., 2021), sometimes even in the case of closely related microorganisms (Avall-Jääskeläinen and Palva, 2005; Fagan and Fairweather, 2014). The *M. lanthanidiphila* mela_00855 had no homologs (after BLASTp) besides a 40.88% protein sequence identity with a protein from *M. oxyfera* (NCBI sequence ID: CBE67388). The power spectrum of the surface of freeze-etched *M. oxyfera* cells showed its S-layer to have either an oblique (p2) or square (p4) symmetry, and a center-to-center spacing of the S-layer morphological unit of ∼7 nm (Wu et al., 2012). These symmetries are different from what we observed for *M. lanthanidiphila*. Nevertheless, both *M. lanthanidiphila* and *M. oyfera*, have a similar polygonal cell shape and the S-layer occurs as straight patches that partially overlap and protrude from the edges of the cell. It would be very interesting to investigate and compare the S-layers of these two *Methylomirabilis* species in the future.

The mela_00855 protein sequence showed 46 putative O-linked and 4 putative N-linked glycosylated sites. Moreover, the genome organization around mela_00855 shows genes involved in glycosylation (glycosyl transferase, mela_00856, mela_00851; UDP-N-acetyl-D-glucosamine dehydrogenase, mela_00857). Therefore, an in-depth characterization of the enriched S-layer by high-resolution shotgun proteomics was performed aiming for high amino acid sequence coverage. Except for the C-terminal tail, the complete amino acid sequence could be covered with peptide fragments. Thereby, no indications for N- or O-linked glycans were found. Nevertheless, because peptides from the C-terminal tail could not be detected, we cannot exclude the presence of glycosylation in this area (even though preliminary gels and staining attempts seem to exclude this scenario, data not shown). Nevertheless, the C-terminal part includes a very hydrophilic sequence region, and therefore is difficult to capture by the applied analysis approach. Glycosylation is a frequent post-translational modification in S-layer proteins (Ristl et al., 2010; Sleytr et al., 2014), albeit there have been reports of S-layer proteins with limited sugar content (Peters et al., 1987) as well as its absence (Masuda and Kawata, 1983; Li et al., 2018). Glycans have been hypothesized to favor protein stability (Mengele and Sumper, 1992; Engelhardt and Peters, 1998; Li et al., 2020). Interestingly, however, despite the apparent lack of glycosylation, the S-layer of *M. lanthanidiphila* shows remarkable resilience to disassembly. In the case of *H. volcanii* it was shown that environmental conditions (particularly salinity) modulate the glycan structure of the S-layer protein as well as the glycosylation sites (Guan et al., 2011). Therefore, it could be speculated that under different growth conditions glycosylation might be present on the *M. lanthanidiphila* S-layer.

We identified mela_00855 as the only component of the *M. lanthanidiphila* S-layer. Nonetheless, we cannot exclude the possibility that other proteins may be associated to the lattice. Infact, due to the impervious nature of disassembly, the S-layer patches were boiled in SDS prior to LC-MS/MS, likely causing the loss of less resilient protein components (such as membrane-associated proteins). Additionally, no other

S-layers have been characterized from members of the NC10 phylum, thereby limiting our understanding of this system. Multi-component S-layers have been described for bacteria and archaea, in which the lattice is constituted by more than one structural and functional component (Mayr et al., 1996; Veith et al., 2009; Bradshaw et al., 2017; Gambelli et al., 2019). Food for thought is how the rigid S-layer sheets are assembled on the outside of the *M. lanthanidiphila* cell. Assembly could be comparable to other S-layers that entirely adhere to the cell body. In the case of *C. crescentus*, S-layer proteins are secreted and diffuse on the lipopolysaccharide until they are incorporated at the edges of growing 2D crystals, preferentially at the cell poles and at the division site (Comerci et al., 2019). *C. difficile* and *H. volcanii* also assemble new S-layer proteins at the mid-cell during cell growth (Abdul-Halim et al., 2020; Oatley et al., 2020). Given that the *M. lanthanidiphila* S-layer sheets partially overlap each other, they could be assembled continuously, leading to individual sheets being pushed further out – perhaps until they snap off. Performing time-lapse microscopy of a growing and dividing *M. lanthanidiphila* cell might answer this question but it is currently technically impossible considering that these bacteria grow as aggregates in an enrichment culture, are anaerobic, have a slow generation time and are not genetically tractable.

Another open question regards the evolutionary advantage of the polygonal shape of *M. lanthanidiphila* (and the *Methylomirabilis* genus in general). We calculated that a *Methylomirabilis* cell with a polygonal shape has a SA:V ratio 0.2 μm-1 higher than a common rod of the same size (**Supplementary Table 2**). Therefore, the selective advantage of this shape for nutrient access seems limited. In laboratory enrichment cultures *Methylomirabilis* bacteria prefer to grow in aggregates with other microorganisms rather than planktonically (Gambelli et al., 2018). S-layers have been implicated to promote formation and maintenance of the aggregates (Zu et al., 2020), the flat surfaces of the cell walls of *Methylomirabilis* bacteria might aid in the attachment of one cell to the other. Moreover, flat surfaces and aggregates could help against shearing forces that these microorganisms might be exposed to in their natural habitats (Mader et al., 1999). Alternatively, predation may also be a selective factor (Beveridge et al., 1997). Indeed, the cell wall is the first element with which predators come into contact. A rigid and resilient cell wall might be harder to digest for predators, making *Methylomirabilis* bacteria a less desirable target. Finally, the possibility that the polygonal shape is not a selective trait that directly confers a survival advantage, but rather a secondary or superfluous trait, as a by-product of other selective features, cannot be excluded. Studies with knock-out mutants could unravel the molecular mechanisms and function of the polygonal cell shape. However, these techniques are currently unfeasible in the case of *Methylomirabilis* bacteria because no pure culture and genetic system are available to date. These characteristics make *Methylomirabilis* bacteria hard to manipulate and a challenging case study for cell shape investigation. Without a genetic system, a possible alternative could be to heterologously express the S-layer protein and observe if the host cell would acquire a polygonal shape.

S-layers have been a subject of investigation through electron microscopy for nearly seven decades (see Sleytr et al., 2014 for an overview). However, the recent developments in cryo-EM have sparked a renewed interest in these fascinating structures due to advancements in cryo-EM allowing us to reach (near) atomic resolution (Bharat et al., 2017; Fioravanti et al., 2019; Gambelli et al., 2019; Herrmann et al., 2020; von Kügelgen et al., 2020). Therefore, we expect additional studies that will broaden our understanding of the role of S-layers in the microbial lifestyle, and how we can exploit them as nanotechnologies.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: ProteomeXchange, PXD029319; EMDB, EMD-13672 and EMD-13670; and EMPIAR, EMPIAR-10822 and EMPIAR-10829.

AUTHOR CONTRIBUTIONS

LN and MJ designed the project. LG, RM, WV, CD, AE, WE, BD, and LN designed the experiments. WV maintained the enrichment culture. WV, LG, and RM performed the enrichment of the S-layer patches. LG performed the SDS-PAGE and immunoblotting. LG and RM performed the immunogold localization. WE and CD collected the cryo-tomography data. LG, CD, and RM analyzed the cryo-tomography data with input from BD and AE. MP performed the glycosylation analysis. LN and LG wrote the manuscript with input from all authors.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.766527/full#supplementary-material

REFERENCES

Abdul-Halim, M. F., Schulze, S., DiLucido, A., Pfeiffer, F., Bisson Filho, A. W., and Pohlschroder, M. (2020). Lipid anchoring of archaeosortase substrates and midcell growth in haloarchaea. *Mbio* 11, e00349–20. doi: 10.1128/mBio.00349-20

Albers, S.-V., and Meyer, B. H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol.* 9, 414–426. doi: 10.1038/nrmicro2576

Arbing, M. A., Chan, S., Shin, A., Phan, T., Ahn, C. J., Rohlin, L., et al. (2012). Structure of the surface layer of the methanogenic archaean Methanosarcina acetivorans. *PNAS* 109, 11812–11817. doi: 10.1073/pnas.1120595109

Ausmees, N., Kuhn, J. R., and Jacobs-Wagners, C. (2003). The bacterial cytoskeleton: an intermediate filament-like function in cell shape. *Cell* 115, 705–713. doi: 10.1016/S0092-8674(03)00935-8

Avall-Jääskeläinen, S., and Palva, A. (2005). Lactobacillus surface layers and their applications. *FEMS Microbiol. Rev.* 29, 511–529. doi: 10.1016/j.fmrre.2005.04.003

Baumeister, W., Wildhaber, I., and Engelhardt, H. (1988). Bacterial surface proteins - Some structural, functional and evolutionary aspects. *Biophys. Chem.* 29:10. doi: 10.1016/0301-4622(88)87023-6

Beveridge, T. J., Pouwels, P. H., Sára, M., Kotiranta, A., Lounatmaa, K., Kari, K., et al. (1997). V. Functions of S-layers. *FEMS Microbiol. Rev.* 20, 99–149. doi: 10.1111/j.1574-6976.1997.tb00305.x

Bharat, T. A. M., Kureisaite-Ciziene, D., Hardy, G. G., Yu, E. W., Devant, J. M., Hagen, W. J. H., et al. (2017). Structure of the hexagonal surface layer on *Caulobacter crescentus* cells. *Nat. Microbiol.* 2:17059. doi: 10.1038/nmicrobiol.2017.59

Bharat, T. A. M., von Kügelgen, A., and Alva, V. (2021). Molecular logic of prokaryotic surface layer structures. *Trends Microbiol.* 29, 405–415. doi: 10.1016/j.tim.2020.09.009

Bradshaw, W. J., Roberts, A. K., Shone, C. C., and Acharya, K. R. (2017). The structure of the S-layer of Clostridium difficile. *J. Cell Commun. Signal.* 12, 319–331. doi: 10.1007/s12079-017-0429-z

Chauhan, J. S., Bhat, A. H., Raghava, G. P., and Rao, A. (2012). GlycoPP: a webserver for prediction of N- and O-glycosites in prokaryotic protein sequences. *PLoS One* 7:e40155. doi: 10.1371/journal.pone.0040155

Chen, J., Zhou, Z., and Gu, J.-D. (2015). Complex community of nitrite-dependent anaerobic methane oxidation bacteria in coastal sediments of the Mai Po wetland by PCR amplification of both 16S rRNA and pmoA genes. *Appl. Microbiol. Biotechnol.* 99:10. doi: 10.1007/s00253-014-6051-6

Comerci, C. J., Herrmann, J., Yoon, J., Jabbarpour, F., Zhou, X., Nomellini, J. F., et al. (2019). Topologically-guided continuous protein crystallization controls bacterial surface layer self-assembly. *Nat. Commun.* 10:2731. doi: 10.1038/s41467-019-10650-x

Deng, X., Llamazares, A. G., Wagstaff, J., Hale, V. L., Cannone, G., McLaughlin, S. H., et al. (2020). The structure of bactofilin filaments reveals their mode of membrane binding and lack of polarity. *Nat. Microbiol.* 4, 2357–2368. doi: 10.1038/s41564-019-0544-0

Deutzmann, J. S., and Schink, B. (2011). Anaerobic oxidation of methane in sediments of Lake Constance, an oligotrophic freshwater lake. *Appl. Environ. Microbiol.* 77, 4429–4436. doi: 10.1128/AEM.00340-11

Engelhardt, H. (2007a). Are S-layers exoskeletons? The basic function of protein surface layers revisited. *J. Struct. Biol.* 160, 115–124. doi: 10.1016/j.jsb.2007.08.003

Engelhardt, H. (2007b). Mechanism of osmoprotection by archaeal S-layers: a theoretical study. *J. Struct. Biol.* 160, 190–199. doi: 10.1016/j.jsb.2007.08.004

Engelhardt, H. (2016). *S-layers.* United States: Wiley. doi: 10.1002/9780470015902.a0021936

Engelhardt, H., and Peters, J. (1998). Structural research on surface layers: a focus on stability, surface layer homology domains, and surface layer-cell wall interactions. *J. Struct. Biol.* 124, 276–302. doi: 10.1006/jsbi.1998.4070

Ettwig, K. F., Butler, M. K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M. M., et al. (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* 464, 543–548. doi: 10.1038/nature08883

Ettwig, K. F., Speth, D. R., Reimann, J., Wu, M. L., Jetten, M. S. M., and Keltjens, J. T. (2012). Bacterial oxygen production in the dark. *Front. Microbiol.* 3:273. doi: 10.3389/fmicb.2012.00273

Ettwig, K. F., van Alen, T., van de Pas-Schoonen, K. T., Jetten, M. S. M., and Strous, M. (2009). Enrichment and molecular detection of denitrifying methanotrophic bacteria of the NC10 phylum. *Appl. Environ. Microbiol.* 75, 3656–3662. doi: 10.1128/AEM.00067-09

Fagan, R. P., and Fairweather, N. F. (2014). Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* 12, 211–222. doi: 10.1038/nrmicro3213

Fioravanti, A., Van Hauwermeiren, F., Van der Verren, S. E., Jonckheere, W., Goncalves, A., Pardon, E., et al. (2019). Structure of S-layer protein Sap reveals a mechanism for therapeutic intervention in anthrax. *Nat. Microbiol.* 4, 1805–1814. doi: 10.1038/s41564-019-0499-1

Ford, M. J., Nomellini, J. F., and Smit, J. (2007). S-layer anchoring and localization of an S-layer-associated protease in Caulobacter crescentus. *J. Bacteriol.* 189, 2226–2237. doi: 10.1128/JB.01690-06

Gaisin, V. A., Kooger, R., Grouzdev, D. S., Gorlenko, V. M., and Pilhofer, M. (2020). Cryo-electron tomography reveals the complex ultrastructural organization of multicellular filamentous chloroflexota (Chloroflexi) bacteria. *Front. Microbiol.* 11:1373. doi: 10.3389/fmicb.2020.01373

Gambelli, L., Guerrero-Cruz, S., Mesman, R. J., Cremers, G., Jetten, M. S. M., Op den Camp, H. J. M., et al. (2018). Community composition and ultrastructure of a nitrate-dependent anaerobic methane-oxidizing enrichment culture. *Appl. Environ. Microbiol.* 84, e02186–17. doi: 10.1128/AEM.02186-17

Gambelli, L., Meyer, B. H., McLaren, M., Sanders, K., Quax, T. E. F., Gold, V. A. M., et al. (2019). Architecture and modular assembly of Sulfolobus S-layers revealed by electron cryotomography. *Proc. Natl. Acad. Sci. U. S. A.* 116, 25278–25286. doi: 10.1073/pnas.1911262116

Graf, J. S., Mayr, M. J., Marchant, H. K., Tienken, D., Hach, P. F., Brand, A., et al. (2018). Bloom of a denitrifying methanotroph, 'Candidatus Methylomirabilis limnetica', in a deep stratified lake. *Environ. Microbiol.* 20, 2598–2614.

Guan, Z., Naparstek, S., Calo, D., and Eichler, J. (2011). Protein glycosylation as an adaptive response in Archaea: growth at different salt concentrations leads to alterations in Haloferax volcanii S-layer glycoprotein N-glycosylation. *Environ. Microbiol.* 14, 743–753. doi: 10.1111/j.1462-2920.2011.02625

Hagen, W. J. H., Wan, W., and Briggs, J. A. G. (2017). Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *J. Struct. Biol.* 197, 191–198. doi: 10.1016/j.jsb.2016.06.007

He, Z., Geng, S., Cai, C., Liu, S., Liu, Y., Pan, Y., et al. (2015). Anaerobic oxidation of methane coupled to nitrite reduction by halophilic marine NC10 bacteria. *Appl. Environ. Microbiol.* 81, 5538–5545. doi: 10.1128/AEM.00984-15

Heringa, J. (1999). Two strategies for sequence comparison: profile-preprocessed and secondary structure-induced multiple alignment. *Comput. Chem.* 23, 341–364. doi: 10.1016/s0097-8485(99)00012-1

Herrmann, J., Li, P. N., Jabbarpour, F., Chan, A. C. K., Rajkovic, I., Matsui, T., et al. (2020). A bacterial surface layer protein exploits multistep crystallization for rapid self-assembly. *Proc. Natl. Acad. Sci. U. S. A.* 117, 388–394. doi: 10.1073/pnas.1909798116

Heumann, J. M., Hoenger, A., and Mastronarde, D. N. (2011). Clustering and variance maps for cryo-electron tomography using wedge-masked differences. *J. Struct. Biol.* 175, 288–299. doi: 10.1016/j.jsb.2011.05.011

Ho, A., Vlaeminck, S. E., Ettwig, K. F., Schneider, B., Frenzel, P., and Boon, N. (2013). Revisiting methanotrophic communities in sewage treatment plants. *Appl. Environ. Microbiol.* 79:5. doi: 10.1128/AEM.03426-12

Houwink, A. L. (1953). A macromolecular mono-layer in the cell wall of Spirillum spec. *Biochim. Biophys. Acta* 10:6. doi: 10.1016/0006-3002(53)90266-2

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031

Klingl, A. (2014). S-layer and cytoplasmic membrane - exceptions from the typical archaeal cell wall with a focus on double membranes. *Front. Microbiol.* 5:624. doi: 10.3389/fmicb.2014.00624

Koch, A. L. (1996). What size should a bacterium be? A question of scale. *Annu. Rev. Microbiol.* 50:31. doi: 10.1146/annurev.micro.50.1.317

Kremer, J. R., Mastronarde, D. N., and McIntosh, R. (1996). Computer visualization of three-dimentional image data using IMOD. *J. Struct. Biol.* 116:5. doi: 10.1006/jsbi.1996.0013

Li, L., Ren, M., Xu, Y., Jin, C., Zhang, W., and Dong, X. (2020). Enhanced glycosylation of an S-layer protein enables a psychrophilic methanogenic archaeon to adapt to elevated temperatures in abundant substrates. *FEBS Lett.* 594, 665–677. doi: 10.1002/1873-3468.13650

Li, P. N., Herrmann, J., Tolar, B. B., Poitevin, F., Ramdasi, R., Bargar, J. R., et al. (2018). Nutrient transport suggests an evolutionary basis for charged archaeal surface layer proteins. *ISME J.* 12, 2389–2402. doi: 10.1038/s41396-018-0191-0

Luesken, F. A., van Alen, T., van der Biezen, E., Frijters, C., Toonen, G., Kampman, C., et al. (2011). Diversity and enrichment of nitrite-dependent anaerobic methane oxidizing bacteria from wastewater sludge. *Appl. Microbiol. Biotechnol.* 92:9. doi: 10.1007/s00253-011-3361-9

Lupas, A., Engelhardt, H., Peters, J., Santarius, U., Volker, S., and Baumeister, W. (1994). Domain structure of the Acetogenium kivui surface layer revealed by electron crystallography and sequence analysis. *J. Bacteriol.* 176, 1224–1233. doi: 10.1128/jb.176.5.1224-1233.1994

Mader, C., Küpcü, S., Sára, M., and Sleytr, U. B. (1999). Stabilizing effect of an S-layer on liposomes towards thermal or mechanical stress. *Biochim. Biophys. Acta* 1418, 106–116. doi: 10.1016/s0005-2736(99)00030-9

Martín-Molina, A., Moreno-Flores, S., Perez, E., Pum, D., Sleytr, U. B., and Toca-Herrera, J. L. (2006). Structure, surface interactions, and compressibility of bacterial S-layers through scanning force microscopy and the surface force apparatus. *Biophys. J.* 90, 1821–1829. doi: 10.1529/biophysj.105.067041

Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* 152, 36–51. doi: 10.1016/j.jsb.2005.07.007

Masuda, K., and Kawata, T. (1983). Distribution and chemical characterization of the regular arrays in the cell walls of strains of the genus Lactobacillus. *FEMS Microbiol. Lett.* 20, 245–150. doi: 10.1111/j.1574-6968.1983.tb00105

Mayr, J., Lupas, A., Kellermann, J., Eckerskorn, C., Baumeister, W., and Peters, J. (1996). A hyperthermostable protease of the subtilisin family bound to the surface layer of the Archaeon Staphylothermus marinus. *Curr. Biol.* 6, 739–749. doi: 10.1016/S0960-9822(09)00455-2

Mengele, R., and Sumper, M. (1992). Drastic differences in glycosylation of related S-layer glycoproteins from moderate and extreme halophiles. *J. Biol. Chem.* 267, 8182–8185. doi: 10.1016/S0021-9258(18)42424-6

Messner, P., Pum, D., Sára, M., Stetter, K. O., and Sleytr, U. B. (1986). Ultrastructure of the cell envelope of the archaebacteria Thermoproteus tenax and Thermoproteus neutrophilus. *J. Bacteriol.* 166, 1046–1054. doi: 10.1128/jb.166.3.1046-1054.1986

Nicastro, D., Schwartz, C., Pierson, J., Gaudette, R., Porter, M. E., and McIntosh, R. (2006). The molecular architecture of axonemes revealed by cryoelectron tomography. *Science* 313:944. doi: 10.1126/science.1128618

Nishiyama, Y., Takashina, T., Grant, W. D., and Horikoshi, K. (1992). Ultrastructure of the cell wall of the triangular halophilic archaebacterium Haloarcula japonica strain TR-1. *FEMS Microbiol. Lett.* 99:5. doi: 10.1111/j.1574-6968.1992.tb05539.x

Oatley, P., Kirk, J. A., Ma, S., Jones, S., and Fagan, R. P. (2020). Spatial organization of Clostridium difficile S-layer biogenesis. *Sci. Rep.* 10:14089. doi: 10.1038/s41598-020-71059-x

Oren, A., Ventosa, A., Gutiérrez, M. C., and Kamekura, M. (1999). Haloarcula quadrata sp. nov., a square, motile archaeon isolated from a brine pool in Sinai (Egypt). *Int. J. Syst. Bacteriol.* 49:6. doi: 10.1099/00207713-49-3-1149

Padilla, C. C., Bristow, L. A., Sarode, N., Garcia-Robledo, E., Gomez Ramirez, E., Benson, C. R., et al. (2016). NC10 bacteria in marine oxygen minimum zones. *ISME J.* 10, 2067–2071. doi: 10.1038/ismej.2015.262

Pavkov-Keller, T., Howorka, S., and Keller, W. (2011). The structure of bacterial S-layer proteins. *Prog. Mol. Biol. Transl. Sci.* 103, 73–130.

Peters, J., Baumeister, W., and Lupas, A. (1996). Hyperthermostable surface layer protein Tetrabrachion from the archaebacterium Staphylothermus marinus: evidence for the presence of a right-handed coiled coil derived from the primary structure. *J. Mol. Biol.* 257, 1031–1041. doi: 10.1006/jmbi.1996.0221

Peters, J., Nitsch, M., Kühlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., et al. (1995). Tetrabrachion: a filamentous archaebacterial surface protein assembly of unusual structure and extreme stability. *J. Mol. Biol.* 245, 385–401. doi: 10.1006/jmbi.1994.0032

Peters, J., Peters, M., Lottspeich, F., Schäfer, W., and Baumeister, W. (1987). Nucleotide sequence analysis of the gene encoding the Deinococcus radiodurans surface protein, derived amino acid sequence, and complementary protein chemical studies. *J. Bacteriol.* 169, 5216–5223. doi: 10.1128/JB.169.11. 5216-5223.1987

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084

Phipps, B. M., Engelhardt, H., Huber, R., and Baumeister, W. (1990). Three-dimensional structure of the crystalline protein envelope layer of the hyperthermophilic archaebacterium Pyrobaculum islandicum. *J. Struct. Biol.* 103, 152–163. doi: 10.1016/1047-8477(90)90019-9

Phipps, B. M., Huber, R., and Baumeister, W. (1991). The cell envelope of the hyperthermophilic archaebacterium Pyrobaculum organothrophum consists of two regularly arrayed protein layers: three-dimensional structure of the outer later. *Mol. Microbiol.* 5, 253–265. doi: 10.1111/j.1365-2958.1991.tb02106

Pum, D., Breitwieser, A., and Sleytr, U. B. (2021). Patterns in nature – S-layer lattices of bacterial and archaeal cells. *Crystals* 11:869. doi: 10.3390/cryst11080869

Pum, D., Messner, P., and Sleytr, U. B. (1991). Role of the S layer in morphogenesis and cell division of the archaebacterium Methanocorpusculum sinense. *J. Bacteriol.* 173, 6865–6873. doi: 10.1128/jb.173.21.6865-6873.1991

Purcell, E. M. (1977). Life at low Reynolds number. *Am. J. Phys.* 45, 3–11. doi: 10.1063/1.30370

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Raghoebarsing, A. A., Pol, A., van de Pas-Schoonen, K. T., Smolders, A. J., Ettwig, K. F., Rijpstra, W. I., et al. (2006). A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature* 440, 918–921. doi: 10. 1038/nature04617

Ristl, R., Steiner, K., Zarschler, K., Zayni, S., Messner, P., and Schäffer, C. (2010). The S-layer glycome—adding to the sugar coat of bacteria. *Int. J. Microbiol.* 2011:127870. doi: 10.1155/2011/127870

Rodrigues-Oliveira, T., Belmok, A., Vasconcellos, D., Schuster, B., and Kyaw, C. M. (2017). Archaeal S-layers: overview and current state of the art. *Front. Microbiol.* 8:2597. doi: 10.3389/fmicb.2017.02597

Schäffer, C., Graninger, M., and Messner, P. (2001). Prokaryotic glycosilation. *Protemics* 1, 248–61. doi: 10.1002/1615-9861(200102)1:2<248::AID-PROT 248<3.0.CO;2-K

Scheuring, S., and Dufrene, Y. F. (2010). Atomic force microscopy: probing the spatial organization, interactions and elasticity of microbial cell envelopes at molecular resolution. *Mol. Microbiol.* 75, 1327–1336. doi: 10.1111/j.1365-2958. 2010.07064

Sleytr, U. B., and Sára, M. (1997). Bacterial and archaeal S-layer proteins: structure-function relationship and their biotechnological applications. *Trends Biotechnol.* 15:7. doi: 10.1016/S0167-7799(96)10063-9

Sleytr, U. B., Messner, P., Pum, D., and Sára, M. (1999). Crystalline bacterial cell surface layers (S-layers): from supramolecular cell structure to biomimetics and nanotechnology. *Angewadte Chem.* 38:20.

Sleytr, U. B., Schuster, B., Egelseer, E. M., and Pum, D. (2014). S-layers: principles and applications. *FEMS Microbiol. Rev.* 38, 823–864. doi: 10.1111/1574-6976. 12063

Toca-Herrera, J. L., Moreno-Flores, S., Friedmann, J., Pum, D., and Sleytr, U. B. (2004). Chemical and thermal denaturation of crystalline bacterial S-layer proteins: an atomic force microscopy study. *Microsc. Res. Tech.* 65, 226–234. doi: 10.1002/jemt.20127

van Teeseling, M. C. F., de Almeida, N. M., Klingl, A., Speth, D. R., Op den Camp, H. J. M., Rachel, R., et al. (2014). A new addition to the cell plan of anammox bacteria: "Candidatus Kuenenia stuttgartiensis" has a protein surface layer as the outermost layer of the cell. 196, 80–89. doi: 10.1128/JB.00988-13

van Teeseling, M. C. F., de Pedro, M. A., and Cava, F. (2017). Determinants of bacterial morphology: from fundamentals to possibilities for antimicrobial targeting. *Front. Microbiol.* 8:1264. doi: 10.3389/fmicb.2017.01264

van Teeseling, M. C., Mesman, R. J., Kuru, E., Espaillat, A., Cava, F., Brun, Y. V., et al. (2015). Anammox Planctomycetes have a peptidoglycan cell wall. *Nat. Commun.* 6:6878.

Vasilyeva, L. V. (1985). Stella, a new genus of soil Prosthecobacteria, with proposals for Stella humosa sp. nov. and Stella acuolata sp. nov. *Int. J. Syst. Evol. Microbiol.* 35, 518–521. doi: 10.1099/00207713-35-4-518

Veith, A., Kling, A., Zolghard, B., Lauber, K., Mentele, R., Lottspeich, F., et al. (2009). Acidianus, Sulfolobus and Metallosphaera surface layers: structure, composition and gene expression. *Mol. Microbiol.* 73, 58–72. doi: 10.1111/j. 1365-2958.2009.06746

Versantvoort, W., Guerrero-Cruz, S., Speth, D. R., Frank, J., Gambelli, L., Cremers, G., et al. (2018). Comparative genomics of Candidatus Methylomirabilis Species and Description of Ca. Methylomirabilis lanthanidiphila. *Front. Microbiol.* 9:1672. doi: 10.3389/fmicb.2018.01672

Viljoen, A., Foster, S. J., Fantner, G. E., Hobbs, J. K., and Dufrêne, Y. F. (2020). Scratching the surface: bacterial cell envelopes at the nanoscale. *Mbio* 11, e03020–19. doi: 10.1128/mBio.03020-19

von Kügelgen, A., Tang, H., Hardy, G. G., Kureisaite-Ciziene, D., Brun, Y. V., Stansfeld, P. J., et al. (2020). *In situ* structure of an intact lipopolysaccharide-bound bacterial surface layer. *Cell* 180, 348–358.e15. doi: 10.1016/j.cell.2019.12. 006

Wagstaff, J., and Löwe, J. (2018). Prokaryotic cytoskeletons: protein filaments organizing small cells. *Nat. Rev. Microbiol.* 16, 187–201. doi: 10.1038/nrmicro. 2017.153

Walsby, A. E. (1980). A square bacterium. *Nature* 283:3. doi: 10.1038/283069a0

Wu, M. L., van Teeseling, M. C., Willems, M. J., van Donselaar, E. G., Klingl, A., Rachel, R., et al. (2012). Ultrastructure of the denitrifying methanotroph "Candidatus Methylomirabilis oxyfera," a novel polygon-shaped bacterium. *J. Bacteriol.* 194, 284–291. doi: 10.1128/JB.05816-11

Yan, P., Li, M., Wei, G., Li, H., and Gao, Z. (2015). Molecular fingerprint and dominant environmental factors of nitrite-dependent anaerobic methane-oxidizing bacteria in sediments from the Yellow River Estuary, China. *PLoS One* 10:e0137996. doi: 10.1371/journal.pone.0137996

Young, K. D. (2004). Bacterial shape. *Mol. Microbiol.* 49, 571–580. doi: 10.1046/j. 1365-2958.2003.03607.x

Zhang, M., Luo, Y., Lin, L., Lin, X., Hetharua, B., Zhao, W., et al. (2018). Molecular and stable isotopic evidence for the occurrence of nitrite-dependent anaerobic methane-oxidizing bacteria in the mangrove sediment of Zhangjiang Estuary, China. *Appl. Microbiol. Biotechnol.* 102, 2441–2454. doi: 10.1007/s00253-017-8718-2

Zheng, S. Q., Keszthelyi, B., Branlund, E., Lyle, J. M., Braunfeld, M. B., Sedat, J. W., et al. (2007). UCSF tomography: an integrated software suite for real-time electron microscopic tomographic data collection, alignment, and reconstruction. *J. Struct. Biol.* 157, 138–147. doi: 10.1016/j.jsb.2006.06. 005

Zu, Y., Hong, S., Xu, C., Li, W., Chen, S., and Li, J. (2020). Cell wall surface layer (S-layer) promotes colony formation in Microcystis: comparison of S-layer characteristics between colonial and unicellular forms of Microcystis and function conformation. *Environ. Sci. Pollut. Res. Int.* 27, 42254–42263. doi: 10.1007/s11356-020-08254-w

frontiers
in Microbiology

Check for
updates

Discovering the Influence of Microorganisms on Wine Color

Rosanna Tofalo*, Giovanna Suzzi and Giorgia Perpetuini

Faculty of Bioscience and Technology for Food, Agriculture and Environment, University of Teramo, Teramo, Italy

Flavor, composition and quality of wine are influenced by microorganisms present on the grapevine surface which are transferred to the must during vinification. The microbiota is highly variable with a prevalence of non-*Saccharomyces* yeasts, whereas *Saccharomyces cerevisiae* is present at low number. For wine production an essential step is the fermentation carried out by different starter cultures of *S. cerevisiae* alone or in mixed fermentation with non-*Saccharomyces* species that produce wines with significant differences in chemical composition. During vinification wine color can be influenced by yeasts interacting with anthocyanin. Yeasts can influence wine phenolic composition in different manners: direct interactions—cell wall adsorption or enzyme activities—and/or indirectly—production of primary and secondary metabolites and fermentation products. Some of these characteristics are heritable trait in yeast and/or can be strain dependent. For this reason, the stability, aroma, and color of wines depend on strain/strains used during must fermentation. *Saccharomyces cerevisiae* or non-*Saccharomyces* can produce metabolites reacting with anthocyanins and favor the formation of vitisin A and B type pyranoanthocyanins, contributing to color stability. In addition, yeasts affect the intensity and tonality of wine color by the action of β-glycosidase on anthocyanins or anthocyanidase enzymes or by the pigments adsorption on the yeast cell wall. These activities are strain dependent and are characterized by a great inter-species variability. Therefore, they should be considered a target for yeast strain selection and considered during the development of tailored mixed fermentations to improve wine production. In addition, some lactic acid bacteria seem to influence the color of red wines affecting anthocyanins' profile. In fact, the increase of the pH or the ability to degrade pyruvic acid and acetaldehyde, as well as anthocyanin adsorption by bacterial cells are responsible for color loss during malolactic fermentation. Lactic acid bacteria show different adsorption capacity probably because of the variable composition of the cell walls. The aim of this review is to offer a critical overview of the roles played by wine microorganisms in the definition of intensity and tonality of wines' color.

Keywords: wine, color, yeasts, lactic acid bacteria, fermentation, metabolism, polyphenols

INTRODUCTION

Wine market is facing several challenges due to consumer demands for high quality wines. The quality of a wine depends on several factors, including grape variety, soil management, winemaking techniques, alcoholic strength, residual sugar content, total and volatile acidity, aroma, flavor, astringency, bitterness, and color. In fact, color intensity and tonality are considered one of the

main parameters contributing to the quality of wine and a matter of concern to winemakers (Forino et al., 2020). In general, the color of young red wines mainly relies on the concentration of monomeric anthocyanins and related compounds, which are extracted from grape skins during the maceration process (Ribereau-Gayon et al., 2006). Polymeric pigments and anthocyanin-derived compounds such as visitins are more resistant to bisulfite bleaching and oxidation and are the main responsible of observed color in aged red wines (Morata et al., 2016). However, a reduction of their concentration occurs during aging and storage of red wine, because of the conversion of monomeric anthocyanins to polymeric pigments and the formation of anthocyanin derivatives (Morata et al., 2016). Wine color is influenced by several factors including the grapevine variety, agricultural practices, and fruit maturation, as well as oenological protocols, such as destemming and crushing conditions, yeast strains used for alcoholic fermentation, malolactic fermentation (MLF), maceration procedures, and wine aging (Ribereau-Gayon et al., 2006; Morata et al., 2016). Yeasts play a key role in the definition of wine color. In fact, they can reduce color intensity and modify wine tonality by deglycosylation of anthocyanins catalyzed by β-glycosidase or anthocyanidase enzymes (Manzanares et al., 2000), through the direct adsorption of pigments on yeasts' cell wall, and producing metabolites such as pyruvic acid and acetaldehyde that have been found to react with different phenolic compounds (Morata et al., 2003, 2006, 2016; Medina et al., 2005; Caridi et al., 2007). Yeast adsorption and its impact on wine color has been demonstrated in several studies and actually is considered an important target for yeast selection. Wine color can also be affected by the metabolic activity of lactic acid bacteria (LAB) (Devi and Ka, 2019). In fact, color loss is common in wines that have undergone MLF (Virdis et al., 2021). LAB can also liberate hydroxycinnamic acids from their tartaric esters and have the potential to break down anthocyanin glucosides, thus impacting wine color (Virdis et al., 2021). This review focuses on the role of wine microorganisms in the definition of wine color.

WINE FERMENTATION

Wine fermentations are characterized by a heterogeneous microbiota and yeasts play a major role in this process. This complex microbial array influences the characteristics of the final product thanks to the coexistence and succession of different species/strains along the fermentation process. From a microbiological point of view, winemaking involves two main steps, the alcoholic fermentation (AF) and MLF. Alcoholic fermentation, mainly driven by *Saccharomyces cerevisiae*, leads to the formation of metabolites of oenological interest (Álvarez-Pérez et al., 2012). However, recent studies demonstrated that hybrids with other species of the *Saccharomyces* complex (e.g., *S. bayanus*, *S. kudriavzevii*, and *S. mikatae*) showed similar fermentation power and vigor and sometimes are preferred in fermentation trials (Gonzalez et al., 2006; Bellon et al., 2013; Peter et al., 2018). However, despite the predominant status of

S. cerevisiae, many non-*Saccharomyces* (NS) yeasts participate to wine fermentation and can shape the sensory characteristics of the wines. In fact, these yeasts may influence the production of secondary and volatile compounds such as esters, higher alcohols, acids and monoterpenes increasing wine quality and complexity (for a review see Padilla et al., 2016). Their occurrence in wine environment has been known for more than a 100 years but they have been considered as spoilage microorganisms or irrelevant species. Thanks to the microbiological studies performed during the last decades enriched with the help of metataxonomic studies (Setati et al., 2012; Bokulich et al., 2013) their role in winemaking has been reconsidered. According to Jolly et al. (2014) NS yeasts can be divided into 3 groups: (i) aerobic yeasts such as *Candida* spp., *Cryptococcus* spp., *Debaryomyces* spp., *Pichia* spp., and *Rhodoturula* spp.; (ii) low fermentative yeasts including *Hanseniaspora uvarum* (*Kloeckera apiculata*), *Hanseniaspora guilliermondii* (*Kloeckera apis*), and *Hanseniaspora occidentalis* (*Kloeckera javanica*); (iii) fermentative yeasts e.g., *Kluyveromyces marxianus* (*Candida kefyr*), *Metschnikowia pulcherrima* (*Candida pulcherrima*), *Torulaspora delbrueckii* (*Candida colliculosa*), and *Zygosaccharomyces bailii*.

S. cerevisiae and NS yeast species do not simply passively coexist during wine fermentation, but a metabolic interplay occurs between them. For instance, mixed fermentations between *S. cerevisiae* and *T. delbrueckii* and *H. vineae*, seem to be a good strategy to enhance wine aroma diversity (Liu et al., 2019). Moreover, *Starm. bacillaris* (syn. *C. zemplinina*) if used in mixed fermentation with *S. cerevisiae*, improve the fermentation kinetic with low ethyl acetate and acetic acid production (Tofalo et al., 2016). Therefore, it is essential not only to select yeasts with suitable oenological properties, but also to consider other aspects including inoculation density, timing, and combination of strains in the organoleptic properties of wines (Englezos et al., 2018). Several efforts must be undertaken in order to establish a link between an inoculation protocol and the chemical composition as well as the chromatic characteristics of wines using the same couple of strains and fermentation conditions. **Table 1** reports the main activities of non-*Saccharomyces* yeasts during wine fermentation and the inoculation protocols applied.

Lactic acid bacteria are responsible of MLF which usually takes place after the AF. Malolactic fermentation is a process required for most red wines and some white wines; it consists of decarboxylation of the L-malic acid to L-lactic acid and induces pH increase, makes wines more palatable by reducing the sour taste associated to malic acid, and provides additional advantages, like microbial stability and improved aroma complexity (Virdis et al., 2021). In particular, LAB belonging to *Lactiplantibacillus, Pediococcus, Leuconostoc,* and *Oenococcus* genera drive the MLF. They are also involved in the definition of wine aroma releasing diacetyl, esters and volatile thiols. They also show pectinolytic activity, which could be useful to improve clarification and the ability to break down acetaldehyde (Virdis et al., 2021). Moreover, recent studies highlighted their role in the definition of wine color (Virdis et al., 2021).

Several studies highlighted that inoculation strategies and timing (i.e., simultaneous or sequential inoculation of LAB and

yeasts) lead to the production of different aroma compounds modifying wine profile (Virdis et al., 2021). Moreover, the development of tailored starter cultures of LAB and yeasts are useful to minimize the sulfur dose (Nardi, 2020). For instance, *T. delbrueckii* has been proposed as an alternative to the use of SO_2 if inoculated at the beginning of the white winemaking process (Simonin et al., 2018).

POLYPHENOLS ADSORPTION AND YEAST CELL WALL

The main role of the yeast cell wall is conferring protection and resistance to environmental conditions. The *S. cerevisiae* cell wall is 100–150 nm thick representing 10–25% of cell dry mass (Yin et al., 2007) and has a bi-layered structure. The outer layer of about 30–40 nm thick is mainly composed of mannoproteins covalently linked to the underlying glycans. The inner layer of about 70–100 nm consisting of a network of branched β-glucans (mainly β-1,3 glucans), serving as a scaffold for the entire cell wall and chitin molecules (Klis et al., 2006; **Figure 1**). Mannoproteins (MPs), β-1,3 glucans, β-1,6 glucans, and chitin, four polysaccharides that are covalently joined, constitute the structure of wall (Schiavone et al., 2017). Highly glycosylated mannoproteins constitute yeast mannan, a complex oligosaccharide comprising 10 to more than 50 mannose units linked in α-(1,2), α-(1,3), α-(1,5), and α-(1,6), which is attached to proteins by either Asn (large manno-oligosaccharides for N-glycosylation) or Ser/Thr residues (short manno- oligosaccharides to make the O-glycosylation) (**Figure 2**). Linear chains of about 1,500 glucose units linked in β-1,3 and β-1,6 compose the β-glucan, whereas 140–350 glucose units linked in β-1,6 form glucan. Chitin is a polymer 100–190 *N*-acetylglucosamine units linked by β-1,4 linkages (for reviews see Lesage and Bussey, 2006; Francois, 2016). The dry mass of cell wall is made of 50% of β-glucan, 40% of mannans, 3–5% of chitin

and then proteins (Schiavone et al., 2014). These components are assembled each other to form a supramolecular architecture, cross-linked in various ways to form higher-order complexes. A central role in this cross-linking is carried out by β-1,6 glucan, even if it is a minor cell wall component from a quantitative point of view (Kollár et al., 1997). The cross-linking cell wall protein (CWPs) to β-1,3 glucans is carried out by β-1,6 glucans in connection with the glycosylphosphatidyl inositol (GPI) anchor attached to these proteins. Proteins with internal repeats (PIR)-CWPs are cell wall proteins directly linked to β-1,3 glucans through γ-carboxylic group of glutamates (Cabib et al., 2012). Cell wall composition varies over the yeast species and strains (Nguyen et al., 1998). Yeast cell wall proteins contain several tandem repeats, which vary greatly in number. Mutations in such repeats are associated to a great functional diversity, which allow yeasts to adapt to different ecological niches or facilitating their exploration of new ones (Verstrepen and Fink, 2009). The number of genes that encode enzymes directly involved in biosynthesis or remodeling of the wall, or non-enzymatic wall proteins, is about 200 genes. During growth and development yeast wall composition and degree of cross-linking can vary (Francois, 2016). Four functions have been recognized for the cell wall, namely stabilization and internal osmotic conditions, protection against stresses, maintenance of the cell shape and integrity, and a scaffold for cell wall proteins (Klis et al., 2006). These functions can be influenced by different factors such as single-strain characteristic, fermentation processes, chemical and environmental stress, substrate composition, and others. The cell wall polysaccharides possess technological properties, relevant for different applications in food safety, biotechnology, and technology (reviewed in Chen and Seviour, 2007; Kogani et al., 2008; Braconi et al., 2011; Pfliegler et al., 2015). In winemaking the role of yeast cell wall components is of great interest for managing fermentations, wine stabilization and aging processes. Several studies recognized a key role of MPs in the determination of wine color. MPs are polysaccharides released by

TABLE 1 | Main roles of non-*Saccharomyces* yeasts in winemaking.

Species	Inoculation strategy	Role in winemaking	References
T. delbrueckii/ S. cerevisiae	Sequential	High production of terpenols, and 2-phenylethanol, higher concentrations of thiols	Bely et al., 2008; Renault et al., 2009, 2016; Comitini et al., 2011; Sadoudi et al., 2012; Van Breda et al., 2013; Belda et al., 2017; Benito, 2018
	Co-inoculation	Low volatile acidity	
Starm. bacillaris/ S. cerevisiae	Co-inoculation	Reduced amount of acetic acid	Tofalo et al., 2012; Lencioni et al., 2018
	Co-inoculation, sequential	High production of glycerol production and low ethanol yield, acetic acid decrease	
L. thermotolerans/ S. cerevisiae	Co-inoculation, sequential	Production of lactic acid, 2-phenylethanol, glycerol, and polysaccharides	Gobbi et al., 2013; Benito et al., 2016; Vilela, 2018
M. pulcherrima/ S. cerevisiae	Sequential	Production of volatile terpene and varietal thiols	Barbosa et al., 2018
	Co-inoculation	Acetic acid decrease	
	Co-inoculation, sequential	Ethyl ester increase	
Sch. pombe/ S. cerevisiae	Co-inoculation, sequential	Reduction of malic acid amount, production of pyruvic acid and polysaccharides	Benito et al., 2012, 2014; Domizio et al., 2017, 2018

Modified from Padilla et al. (2016).

FIGURE 1 | Yeast cell wall. It is made up of mannoproteins, β-1,3-glucans, β-1,6-glucans, and chitin. Cell wall may undergo several changes during alcoholic fermentation because they are exposed to several stresses (osmotic stress, low pH, high acidity, nitrogen depletion, elevated ethanol) which should increase anthocyanins adsorption. Macromolecules of *S. cerevisiae* cell wall (Klis et al., 2006; De Iseppi et al., 2020).

yeast cells during wine fermentation and during aging of wine on lees by endo-glucanases, exo-D-mannose, and α-D-mannosidase (Arévalo Villena et al., 2005; Belda et al., 2016; Balmaseda et al., 2021). These proteins are mainly composed of mannose and glucose with a protein content ranging between 1 and 10% with a molecular weight ranging from 50 to 500 kDa (Yue et al., 2021). MPs protect wine against protein precipitation and stabilize wine color intensity. In fact, yeast MPs can combine with anthocyanins and tannins increasing color stability (Escot et al., 2001). In fact, the addition of MPs before AF enhances the content of anthocyanins and phenolic acids improving the color stability (Yue et al., 2021) and could protect the degradation of phenolic acids during the fermentation process playing a protective role (Fernando et al., 2018; Rinaldi et al., 2019). However, Mekoue Nguela et al. (2019) reported that polyphenol adsorption on yeast cell outer surface can have negative consequences on the cell wall metabolic activity interfering with cell signaling functions and nutrient transport. Interactions between yeast and polyphenols have been observed in wine aging on lees, a practice applied after fermentation to maintain the wine in contact with dead yeast cells (lees) (Mazauric and Salmon, 2005, 2006). The MPs' influence on wine color is also dependent upon the strain of yeast used (Escot et al., 2001). Even if *S. cerevisiae* cell wall is considered the main source of MPs, also some NS yeasts such as *Schizosaccharomyces pombe, Pichia fermentans, M. pulcherrima, Saccharomycodes ludwigii, T. delbrueckii, Lachancea thermotolerans,* and *Wickerhamomyces anomalus,* demonstrated the ability to produce and release MPs into the wine during aging on lees (Morata et al., 2006, 2019; Belda et al., 2016; Ferrando et al., 2020; Balmaseda et al., 2021). MPs are released

continuously during the growth of several NS yeasts, reflecting a high production of these polysaccharides during the first phase of fermentation (Domizio et al., 2014). Some studies highlighted that NS yeasts showed a higher release of MPs in wine than *S. cerevisiae* (for a review see Vejarano, 2020) with *S'codes ludwigii* are found among the species with high potentials for releasing polysaccharides (Palomero et al., 2009; Domizio et al., 2017). According to Palomero et al. (2009) *S'codes ludwigii* released 110.51 mg/L of MPs against the 36.65 mg/L of *S. cerevisiae*. non-*Saccharomyces* MPs showed a different structure in terms of protein, mannose, glucose, and galactose content compared to those characterizing *S. cerevisiae*. For instance, the % of mannose residues is 88% in *S. cerevisiae*, while range from 55% in *Sch. pombe* to 93% in *S'codes ludwigii*. Moreover, α-galactomannose rather than mannose has been found as part of the structure of polysaccharides in *Sch. pombe*. The polysaccharides from these NS yeasts show a greater molecular size and may potentially impact the wine's palatability (Vejarano, 2020).

During aging on lees the color loss can occur since lees can interact with anthocyanins through pigments adsorption by lees and anthocyanins degradation by β-glucosidase enzymes (Darriet et al., 2012). Lees from different yeasts can have a different adsorption. For instance, lees from *M. pulcherrima, S'codes ludwigii*, or *Sch. pombe* have shown a low adsorption of anthocyanins with respect to the lees of *S. cerevisiae, T. delbrueckii*, or *L. thermotolerans* (Herderich et al., 2013). The less color loss could be also related to the ability of these NS yeasts, especially, *S. pombe*, to produce pyranoanthocyanins which are more stable than anthocyanins (Claus and Mojsov, 2018). In addition, *S. pombe*—besides its deacidification activity due to its

FIGURE 2 | Some structural characteristics of mannoproteins. Asn, asparagine; GlcNAc, N-acetylglucosamine; Man, mannose; P, phosphate; Ser, serine; Thr, threonine.
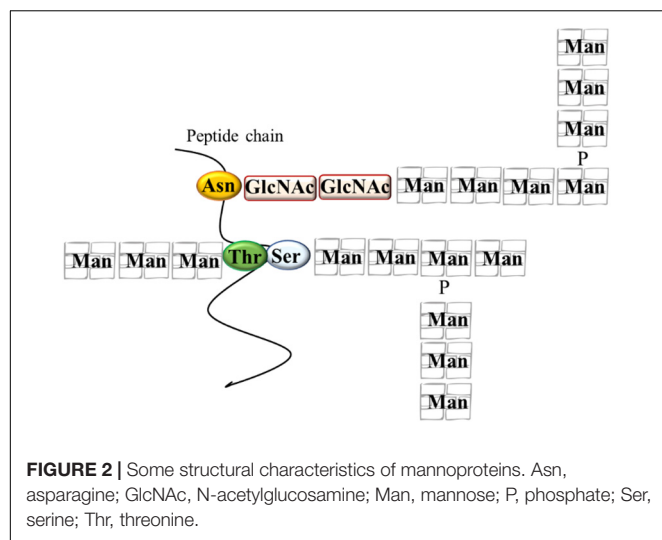
ability to convert malic acid into ethanol and carbon dioxide—is able to improve color stability of red wine producing vitisin A and anthocyanin-vinyl phenol derivatives thanks to its capacity to release pyruvic acid and hydroxycinnamate decarboxylase activity (Morata et al., 2012).

INFLUENCE OF YEASTS ON POLYPHENOLIC PROFILE OF WINES

Wines are characterized by a certain variety of phenolic compounds also known as polyphenols or biophenols. Grape polyphenols are secondary compounds extracted during the winemaking process which contribute to wine color and flavor especially in red wines (Goldner and Zamora, 2007). In fact, in red wine their concentration is approximately about six times higher than that in white one because red juice has longer contact time with the grape skins and seeds. In particular, the minimum and maximum levels of total phenolic contents reported (expressed in mg of gallic acid equivalents per liter) ranged from 1,531 to 3,192 and from 210 to 402 for red and white wines, respectively (Visioli et al., 2020). They are composed of one or more hydroxyl groups linked with one or more aromatic or benzene rings (Visioli et al., 2020). Moreover, these compounds can be conjugated to one or more sugar residues linked by β-glycosidic (O-glycosylated) bonds or by direct linkages of sugar to an aromatic ring carbon atom (C-glycosides) (Visioli et al., 2020). They can be classified into 2 groups: the flavonoids and the non-flavonoids. The first one includes anhocyanins, flavonols, and flavonoids. The second one encompasses hydroxybenzoic acids, hydroxycinnamates, and the stilbenoids (Jaganath and Crozier, 2010; **Table 2**). The concentration of polyphenols in wines is influenced by viticulture (grape variety and clone, light exposure, degree of ripeness), vinification process (destemming, crushing, pre-fermentation maceration, alcoholic fermentation, pressing), and yeast strains (Jagatić Korenika et al., 2021). For instance, must freezing, cryogenic maceration, extended maceration, and temperature
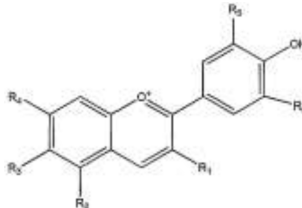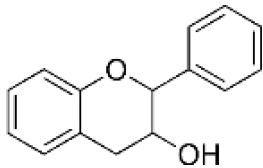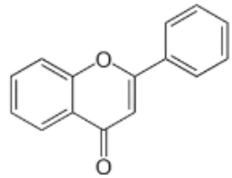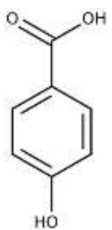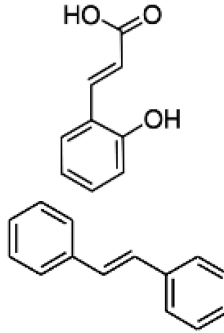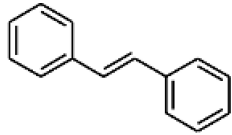
increase phenolics concentrations in wines, while mechanical harvesting could decrease their concentration through reactions with oxidative radicals (Olejar et al., 2015).

The first evidence of yeast influence on polyphenols content was reported by Caridi et al. (2004) who highlighted the correlations between yeast strain and chromatic properties, and phenolic profile of wines. Subsequently, Medina et al. (2005) reported the importance of yeast selection to shape anthocyanin concentration and Morata et al. (2016) demonstrated that yeast strains are involved in stable pigments formation and in the adsorption of color molecules on the cell wall. In fact, yeast metabolism may lead to different values of metabolic precursors during must fermentation for the formation of pyranoanthocyanins, oligomeric, and polymeric pigments (Escott et al., 2018).

Actually, 3 different mechanisms have been described to explain the interaction between the wine's polyphenols and yeasts (Giovinazzo et al., 2019). The first one is based on the adsorption of polyphenols on the yeast cell wall. This phenomenon is strain dependent and not yet completely understood. It probably depends on cell wall surface structure and composition being apolar anthocyanins better adsorbed than polar ones. According to Echeverrigaray et al. (2019), yeasts could be grouped as low, medium, and high anthocyanins adsorption strains. The same authors demonstrated that yeast anthocyanin adsorption occurs by pigment-binding molecules constitutively expressed in the inner part of the cell walls of all *Saccharomyces* strains, regardless of their assigned adsorption behavior in red wine fermentation (high, medium, and low pigment adsorption yeast strains) (Echeverrigaray et al., 2020). Recent studies showed that yeast cells primarily adsorb grape pigments by the end of fermentation (Echeverrigaray et al., 2019, 2020). The second mechanism is linked to yeast β-glucosidase activity, which breaks the polyphenols–sugar bond, anthocyanidase enzymes, or pectinolytic enzymes which favor the extraction of color from pomace (Manzanares et al., 2000). Glycosidase activities have been described in various NS yeasts (*Candida, Hanseniaspora, Pichia, Metschnikowia, Rhodotorula, Trichosporon, Wickerhamomyces*) (for a review see Claus and Mojsov, 2018). Hydrolysis of glucose usually results in a corresponding anthocyanidin, which is converted to the colorless pseudobase, which may affect color and stability (Mansfield et al., 2002). The last one is based on the release by yeast strains of polysaccharides, like MPs, able to entrap polyphenols during fermentation (see previous paragraph).

Of particular interest for the determination of wine color is the adsorption of anthocyanins. During fermentation and aging, anthocyanins are subjected to chemical modifications through their interaction with other compounds including pyruvic acid, flavan-3-ols, condensed tannins, etc. (Morata et al., 2016; **Figure 3**). These modifications together with structural or metabolic modifications of yeast increase their adsorbability by yeast cell wall. These changes have been found to occur during alcoholic fermentation as a response to stresses such as high osmotic pressure, low pH, high acidity, nitrogen depletion, elevated ethanol, in order to maintain metabolic activity and cell viability (Klis et al., 2002; Aguilar-Uscanga and François, 2003;

TABLE 2 | Characteristics of main polyphenols occurring in wine (modified by Visioli et al., 2020).

Polyphenols	Characteristics
Flavonoids	
Anthocyanins	They are water-soluble pigments and the main anthocyanins in wines are anthocyanidin—3-O-glucosides, peonidin-3-glc, cyanidin-3-glc, petunidin-3-glc, and delphinidin—3-glc. Their concentration can reach 400 mg/L in red wines, while are absent in white ones. Their effect on wine color is influenced by several factors and the main are pH and co-pigmentation. At low pH the red color is stable, while in presence of alkaline conditions they appear purple/blue. Anthocyanins can interact with other polyphenols (co-pigmentation) stabilizing wine color. In particular, during fermentation and aging, anthocyanins are chemically modified by their interaction with pyruvic acid, coumaric acid, ethanal, flavan-3-ols, condensed tannins, and other reactive molecules yielding pyroanthocyanins and polymerized pigments
Flavan-3-ol	They are yellow pigments responsible of wines' astringency, bitterness, and structure and can be found in monomeric form (catechin and epicatechin) and in their polymeric form (proanthocyanidins, also called condensed or non-hydrolysable tannins). The main ones detected in grapes and wine are myricetin, quercetin, laricitrin, kaempferol, isorhamnetin, and syringetin. They can be found in both white and red wines with values ranging from 15 to 25 mg/L and from 4 to 120 mg/L, respectively.
Flavonoids	They exist as glycosides in combination with monosaccharides such as glucose, rhamnose, galactose, xylose, and arabinose.
Non-flavonoids	
Hydroxybenzoic acids	The most abundant are p-hydroxybenzoic, gallic, vanillic, gentisic, syringic, salicylic, and protocatechuic acids. The gallic acid has been found in red and white wines with concentrations 70 and 10 mg/L, respectively.
Hydroxycinnamic acids	Hydroxycinnamic acids are the main group of polyphenols in must and white wine. They are generally conjugated with tartaric acid esters or diesters and responsible of wine browning processes since they can be oxidized.
Stilbenes	The main stilbenes found in wines are trans-piceid and trans-resveratrol, hopeaphenol, ampelosin A, isohopeaphenol, piceatannol, pallidol, e-viniferin, miyabenol C, r-viniferin, and r2-viniferin. In general, they occur in low concentrations, but if grapes are subjected to abiotic and biotic stresses the amount of resveratrol can reach values of 100 mg/L.

Duc et al., 2017). Yeast stress response depend on several genes and varies among yeasts strains winemaking conditions (Stanley et al., 2010; Duc et al., 2017). In general, cell viability and cell wall integrity are negatively correlated with pigment adsorption and induce the phenotypic differences among strains (Echeverrigaray et al., 2020; for a review see Zhang et al., 2021).

Morata et al. (2016) showed that adsorption mechanisms can be responsible for up to 6% reduction of anthocyanins that contribute to wine color intensity, significantly impacting the quality of the final product. Yeast adsorption and its impact on wine color has been corroborated by multiple studies (Morata et al., 2003, 2016; Medina et al., 2005; Caridi et al., 2007), and this trait has been identified as a target for yeast strain selection toward the improvement for wine production. The importance of yeast strain selection has been shown also by Carew et al. (2013) who demonstrated a significant influence of yeast strain on the concentration and composition of wine tannins. The use of S. cerivisiae RC212 results in wines with high concentration of total pigment, free anthocyanin, non-bleachable pigment, and total tannin, and showed high color
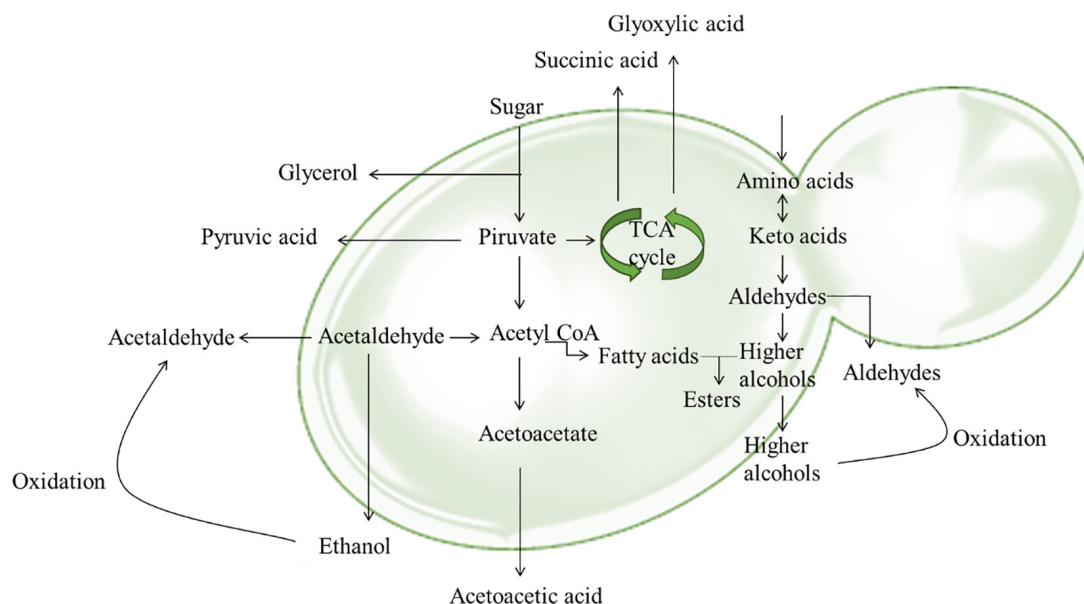
FIGURE 3 | Main pathways involved in the formation of anthocyanin derivatives and polymeric pigments.

density. Moreover, the sequential inoculation of *S. cerivisiae* RC212 and *T. delbruekii* allowed to obtain wines with a high degree of tannin polymerization. Different behavior of two *S. cerevisiae* strains, and a *S. bayanus* strain on the phenolic profile of Aurora white wine have been reported by Samoticha et al. (2019). *S. cerevisiae* strains produced higher amounts of polyphenols. In fact, resulting wines had a content of total polyphenols of about 300 mg/L and a high antioxidant capacity. Similarly, Grieco et al. (2019) highlighted the importance to select autochthonous tailored yeast strains to modulate the phenolic composition of Negroamaro and Primitivo wines. A recent study revealed that the use of *M. pulcherrima*, *Z. bailii*, *Candida zeylanoides*, and *T. delbrueckii* increased the content of monomeric anthocyanin in Tempranillo wines improving their color and health properties (Escribano-Viana et al., 2019). Similar results concerning *T. delbrueckii* were obtained by Balmaseda et al. (2021). These authors revealed that the use of *T. delbrueckii* improved the volatile complexity and polyphenolic composition of wines and enabled spontaneous MLF. Moreover, some yeast metabolites can react with anthocyanins forming more stable pigments. Visitin A is formed during the fermentation process and aging from pyruvic acid produced by yeasts and malvidin-3-*O*-glucoside, while visitin B from malvidin-3-*O*-glucoside and acetaldehyde (Morata et al., 2003).

Pyruvate is produced during the catabolism of sugars and can be metabolized into acetaldehyde, or used in the formation of acetyl-CoA (Morata et al., 2003). Acetaldehyde is a byproduct of yeast metabolism, and can be also produced through a non-enzymatic oxidation of ethanol (Danilewicz, 2003). Acetaldehyde concentrations increase during the fermentation process and as a wine is exposed to oxygen. A controlled oxidation of wine is highly recommended since uncontrolled introduction of oxygen can cause alterations of wine color, and the loss

of desirable aromas and formation of undesirable aromas, and even promote aerobic bacteria (Gómez-Plaza and Cano-López, 2011). Generally, acetaldehyde reacts with sulfur dioxide (SO_2), or with other compounds such as tannins influencing wine color (Carlton et al., 2007). In fact, it promotes rapid polymerization between anthocyanins and catechins or tannins, forming stable polymeric pigments resistant to SO_2 bleaching (Timberlake and Bridle, 1976).

Acetaldehyde and pyruvic acid production are strain specific and is particularly evident in NS yeasts. For instance, *Sch. pombe* released a higher concentration of pyruvate than *S. cerevisiae* during fermentation (Morata et al., 2012; Belda et al., 2017). *Torulaspora delbrueckii* produces low amount of acetaldehyde, in comparison with *S. cerevisiae*. This trait is interesting not only in terms of wine color because is related to visitin B production, but also since concentrations above 125 mg/L of acetaldehyde has a negative effect on wine's flavor (Benito et al., 2019). Therefore, the selection NS yeast with a suitable production of pyruvate and acetaldehyde to be used in combination with *S. cerevisiae* could represent a useful strategy to increase visitins production during must fermentation. Some NS yeasts can release up to four times higher concentrations of pyruvic acid or acetaldehyde than *S. cerevisiae*. The combination of tailored NS species/strains could allow the microbial stabilization of wines, avoiding malolactic fermentation and increase the acidity and color perception (Chen et al., 2018; Benito et al., 2019). Moreover, *S. cerevisiae* as well as NS yeasts have hydroxycinnamate decarboxylase (HCDC) activity and are able to produce vinylphenolic pyranoanthocyanins (VPAs) from the chemical interaction between hydroxycinnamic acids and anthocyanins (Morata et al., 2007). Hydroxycinnamic acids could also directly react with anthocyanins and form VPAs without enzymatic support (Schwarz et al., 2003). HCDC activity

has been described in several yeast species (Morata et al., 2012). Recently, this activity was tested in 14 different yeast genera (*Wickerhamomyces*, *Torulaspora*, *Starmerella*, *Pichia*, *Metschnikowia*, *Lachancea*, *Kregervanrija*, *Kluyveromyces*, *Kodamaea*, *Issatchenkia*, *Hanseniaspora*, *Debaryomyces*, *Candida*, *Meyerozyma*) and revealed that *M. guilliermondii* and *W. anomalus* strains had the highest HCDC activity, while *S. servazii*, *M. fructicola*, *K. dobzhanskii*, *H. osmophila*, *C. sake* strains the lowest (Božič et al., 2020). Moreover, Escott et al. (2018) found a higher concentration of stable pigments produced during fermentation with non-*Saccharomyces* yeasts in comparison to pure fermentations with *S. cerevisiae*.

Independent of their adsorption behavior during red wine fermentation, damaged yeast cells showed the same anthocyanin adsorption capacity, indicating that any major differences in anthocyanin adsorption between yeast strains are determined by their ability to maintain cell viability, as well as the cell wall and membrane integrity throughout wine fermentation (Echeverrigaray et al., 2020). Moreover, these results suggest that anthocyanin adsorption binding molecules, probably MPs, did not vary significantly among strains, and that such factors are localized in the inner part of yeast cell walls, as previously suggested by observations by Vasserot et al. (1997), Petruzzi et al. (2015), and Gonçales et al. (2018).

Studies carried out some years ago found that some *S. cerevisiae* strains could stabilize white wine color after exposure to air and light (Suzzi et al., 1985). This ability called "stabilizing power" varied in relation to sulfite production; low sulfite forming strains did not produce stable wines, whereas strains able to stabilize wine color were high sulfite forming ones. However, a direct relationship between SO_2 produced during fermentation and stabilization ability was not always recorded (for a review see Romano and Suzzi, 1993).

Yeasts can also influence wine color releasing organic acids. Englezos et al. (2018) revealed that *Starm. bacillaris* produced pyruvic acid, acting as a natural acidification agent by reducing the wine's pH. The acidogenic nature of *Starm. bacillaris* could have an impact on wine color stability, mainly due to the ability of the pyruvic acid to bind sulfur dioxide and swift the equilibrium of anthocyanins from the colorless to colored form due to the reaction of pyruvic acid with anthocyanins producing stable pigments such as vitisin A (Englezos et al., 2018).

Some studies also focused on the role of flor yeast to decrease browning in white wines (Fabios et al., 2000; Merida et al., 2005). *Saccharomyces cerevisiae* flor yeast or flor velum yeasts can grow at the surface of different wines are involved in their biological aging (David-Vaizant and Alexandre, 2018). During biological aging these yeasts shifts from a fermentative to an oxidative metabolism (diauxic shift) induced by nitrogen and sugar depletion and rise to the wine surface to form multicellular aggregates. This aggregation leads to the build-up of a biofilm, or velum or flor (Legras et al., 2016). Biofilm formation begins under nutrients starvation and is favored by the presence of other carbon sources, such as glycerol and ethyl acetate (Zara et al., 2010). Flor yeasts are able to face wine stresses characterizing mainly induced by ethanol and acetaldehyde (Zara

et al., 2010). This strong adaptation is probably related to DNA mutations responsible for mitochondrial DNA polymorphism and chromosomal rearrangements (Legras et al., 2016). Some studies suggested that probably these yeasts are able to protect wine from browning since they consume the oxygen through their aerobic metabolism and because they retain brown pigments on the cell wall. Some authors reported in presence of flor yeasts a gradual disappearance of catechin. This phenomenon may be due to the production of acetaldehyde during flor yeasts aerobic growth (Cortes et al., 1998), thereby favoring the formation of oligomers (Fulcrand et al., 1998).

INFLUENCE OF LACTIC ACID BACTERIA ON WINE COLOR

Lactic acid bacteria are responsible of MLF. During this process L-malic acid is converted into L-lactic acid and reduce the acidity of wine. The consumption of L-malic acid reduces the risk of the wine spoilage and improve the palatability of wine (Sumby et al., 2019). In addition, aroma precursors in wine can be further hydrolyzed into free-form volatiles under malolactic fermentation (Lonvaud-Funel, 1999). Color loss is common in wines that have undergone MLF (Burns and Osborne, 2013). In fact, these wines, independently from pH, are characterized by lower levels of polymeric pigments, lower visitin A and B content and a higher concentration of monomeric anthocyanins than wines that did not undergo MLF (**Figure 4**). This phenomenon has been explained by LAB ability to utilize acetaldehyde and pyruvic acid during MLF (Devi and Ka, 2019). This compound is essential for ethylene-linked pigments formation, which are more stable than their respective monomeric anthocyanins and show better colorimetric properties (Forino et al., 2020). Moreover, pyruvic acid can react with malvidin-3-glucoside, forming pyranoanthocyanins e.g., visitin A and visitin B (Waterhouse et al., 2016). Moreover, Devi et al. (2019) showed that *O. oeni* and *Lactiplantibacillus plantarum* strains are able to adsorb delphinidin-3-glucoside, malvidin-3-glucoside and peonidin-3-glucoside on the cell wall and can also produce β-glycosidase enzymes. However, it has been reported that some LAB strains belonging to *Lactiplantibacillus, Leuconostoc, Pediococcus,* and *Streptococcus* genera are able to produce acetaldehyde (Liu and Pilone, 2000). Wang et al. (2018) showed that *Lpb. plantarum* inoculated wine after MLF showed higher level of pyranoanthocyanins, whereas the use of *O. oeni* strains resulted in less formation of pyranoanthocyanins in wine. Moreover, *Lpb. plantarum* strains increased the accumulation of acetaldehyde in wine model medium and wine during malolactic fermentation. The influence on wine color depends also on the yeast/bacteria inoculation strategy. In fact, when MLF is performed with sequential inoculation a higher color loss is observed (Devi and Ka, 2019). Actually, some studies are focused on the long-term effects of MLF on wine color. Izquierdo-Cañas et al. (2016) reported that 9 months after the end of MLF resulting in color intensity loss and lower acylated and non-acylated anthocyanins levels. However, an increase of the pyranoanthocyanin concentration was observed.
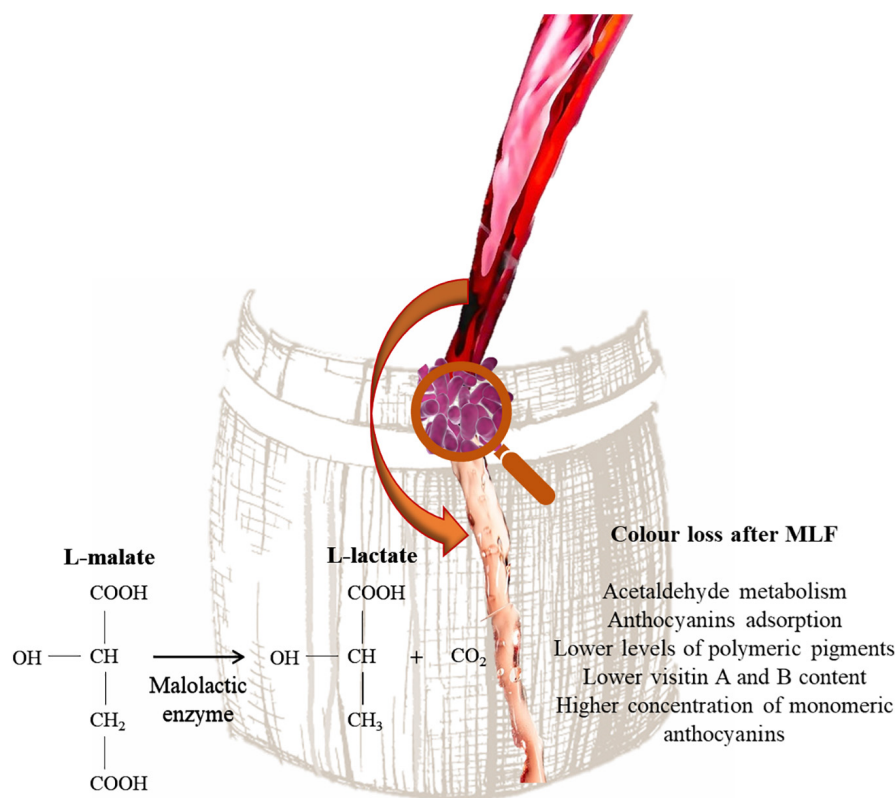
FIGURE 4 | Influence of MLF on wine color.

CONCLUSION

Wine characteristics depend on grape berry phenolic composition, and on microbial activities during fermentation. Yeasts belonging to the genus *Saccharomyces* are considered the main actors of wine fermentation. However, non-*Saccharomyces* yeasts can be exploited as potential starters in mixed fermentations with *S. cerevisiae*. The understanding and managing of yeasts, their diversity and effects on wine quality can be optimized, resulting in better organoleptic characteristics, such as color and aroma. Yeasts can impact wine color through at least 3 different mechanisms: (i) release of metabolites which could participate in the red wine color stabilization process and increase the content of stable pigments; (ii) presence of enzymatic activities such as glycosidase and pectinase; (iii) adsorption of phenolic compounds by yeast cell wall, especially anthocyanins and tannins, which largely leads to the loss of red wine color and reduction of astringency. Different strains of *S. cerevisiae* and NS yeasts have been found to influence wine color in a different way, mostly due to their variations in the forementioned three mechanisms. In this sense population studies might be useful. In fact, the ratio of low and high adsorbing cell populations varied among wine yeast strains, and is related to yeast fermentative life-span or cell viability. Tailored yeast strains can affect and stabilize wine color and pigments. The wide diversity of effects on polyphenols and on the final wine color in single and mixed fermentations carried out by *S. cerevisiae* and NS strains indicates the great importance of these studies for the future of winemaking. Moreover, the influence of LAB on wine color should be also considered. A key point is the establishment of the right time for promoting MLF to prevent consumption of pyruvic acid by LAB and to promote vitisin synthesis. Some yeast-derived compounds such as mannoproteins can be stimulatory for *O. oeni*. Further studies are necessary to clarify the regulation of mannoprotein metabolism in LAB and to evaluate the effect of mannoproteins released by different yeasts on LAB fitness and MLF kinetics. Moreover, a better knowledge on yeasts/bacteria interactions during fermentation and on the effects of inoculation strategies should be achieved to improve wine color and contribute to consumers' purchasing decision.

AUTHOR CONTRIBUTIONS

RT: conceptualization, writing—review and editing, and funding acquisition. GS: review and editing. GP: writing—original draft and writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

REFERENCES

Aguilar-Uscanga, B., and François, J. M. (2003). A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. *Lett. Appl. Microbiol.* 37, 268–274. doi: 10.1046/j.1472-765X.2003. 01394.x

Álvarez-Pérez, J. M., Campo, E., San-Juan, F., Coque, J. J. R., Ferreira, V., and Hernández-Orte, P. (2012). Sensory and chemical characterisation of the aroma of prieto picudo rosé wines: the differential role of autochthonous yeast strains on aroma profiles. *Food Chem.* 133, 284–292. doi: 10.1016/j.foodchem.2012.01. 024

Arévalo Villena, M., Úbeda Iranzo, J., Cordero Otero, R., and Briones Pérez, A. (2005). Optimization of a rapid method for studying the cellular location of β-glucosidase activity in wine yeasts. *J. Appl. Microbiol.* 99, 558–564. doi: 10.1111/j.1365-2672.2005.02627.x

Balmaseda, A., Aniballi, L., Rozès, N., Bordons, A., and Reguant, C. (2021). Use of yeast mannoproteins by *Oenococcus oeni* during malolactic fermentation under different oenological conditions. *Foods* 10:1540. doi: 10.3390/foods1007 1540

Barbosa, C., Lage, P., Esteves, M., Chambel, L., Mendes-Faia, A., and Mendes-Ferreira, A. (2018). Molecular and phenotypic characterization of *Metschnikowia pulcherrima* strains from Douro Wine Region. *Fermentation* 4:8. doi: 10.3390/fermentation4010008

Belda, I., Navascués, E., Marquina, D., Santos, A., Calderón, F., and Benito, S. (2016). Outlining the influence of non-conventional yeasts in wine ageing over lees. *Yeast* 33, 329–338. doi: 10.1002/yea.3165

Belda, I., Ruiz, J., Beisert, B., Navascués, E., Marquina, D., Calderón, F., et al. (2017). Influence of *Torulaspora delbrueckii* in varietal thiol (3-SH and 4-MSP) release in wine sequential fermentations. *Int. J. Food Microbiol.* 257, 183–191. doi: 10.1016/j.ijfoodmicro.2017.06.028

Bellon, J. R., Schmid, F., Capone, D. L., Dunn, B. L., and Chambers, P. J. (2013). Introducing a new breed of wine yeast: interspecific hybridisation between a commercial *Saccharomyces cerevisiae* wine yeast and *Saccharomyces mikatae.* *PLoS One* 8:e62053. doi: 10.1371/journal.pone.0062053

Bely, M., Stoeckle, P., Masneuf-Pomarède, I., and Dubourdieu, D. (2008). Impact of mixed *Torulaspora delbrueckii-Saccharomyces cerevisiae* culture on high-sugar fermentation. *Int. J. Food Microbiol.* 122, 312–320. doi: 10.1016/j.ijfoodmicro. 2007.12.023

Benito, Á, Calderón, F., and Benito, S. (2019). The influence of non-*Saccharomyces* species on wine fermentation quality 540 parameters. *Fermentation* 5, 1–18. doi: 10.3390/fermentation5030054

Benito, Á, Calderón, F., Palomero, F., and Benito, S. (2016). Quality and composition of airén wines fermented by sequential inoculation of *Lachancea thermotolerans* and *Saccharomyces cerevisiae.* *Food Technol. Biotechnol.* 54, 135–144. doi: 10.17113/ftb.54.02.16.4220

Benito, S. (2018). The impact of *Torulaspora delbrueckii* yeast in winemaking. *Appl. Microbiol. Biotechnol.* 102, 3081–3094. doi: 10.1007/s00253-018-8849-0

Benito, S., Palomero, F., Calderón, F., Palmero, D., and Suárez-Lepe, J. A. (2014). Selection of appropriate *Schizosaccharomyces* strains for winemaking. *Food Microbiol.* 42, 218–224. doi: 10.1016/j.fm.2014.03.014

Benito, S., Palomero, F., Morata, A., Calderón, F., and Suárez-Lepe, J. A. (2012). New applications for *Schizosaccharomyces pombe* in the alcoholic fermentation of red wines. *Int. J. Food Sci. Technol.* 47, 2101–2108. doi: 10.1111/j.1365-2621. 2012.03076.x

Bokulich, N. A., Ohta, M., Richardson, P. M., and Mills, D. A. (2013). Monitoring seasonal changes in winery-resident microbiota. *PLoS One* 8:e66437. doi: 10. 1371/journal.pone.0066437

Božič, J. T., Butinar, L., Albreht, A., Vovk, I., Korte, D., and Vodopivec, B. M. (2020). The impact of *Saccharomyces* and non-*Saccharomyces* yeasts on wine colour: a laboratory study of vinylphenolic pyranoanthocyanin formation and anthocyanin cell wall adsorption. *LWT Food Sci. Technol.* 123:109072. doi: 10.1016/j.lwt.2020.109072

Braconi, D., Amato, L., Bernardini, G., Arena, S., Orlandini, M., Scaloni, A., et al. (2011). Surfome analysis of a wild-type wine Saccharomyces cerevisiae strain. *Food Microbiol.* 28, 1220–1230. doi: 10.1016/j.fm.2011.04.009

Burns, T. R., and Osborne, J. P. (2013). Impact of malolactic fermentation on the color and color stability of Pinot noir and Merlot wine. *Am. J. Enol. Vitic.* 64, 370–377. doi: 10.5344/ajev.2013.13001

Cabib, E., Blanco, N., and Arroyo, J. (2012). Presence of a large beta (1- 3) glucan linked to chitin at the Saccharomyces cerevisiae mother-bud neck suggests involvement in localized growth control. *Eukaryot. Cell* 11, 388–400. doi: 10.1128/EC.05328-11

Carew, A. L., Smith, P., Close, D. C., Curtin, C., and Dambergs, R. G. (2013). Yeast effects on pinot noir wine phenolics, color, and tannin composition. *J. Agric. Food Chem.* 61, 9892–9898. doi: 10.1021/jf4018806

Caridi, A., Cufari, A., Lovino, R., Palumbo, R., and Tedesco, I. (2004). Influence of yeast on polyphenol composition of wine. *Food Technol. Biotechnol.* 42, 37–40.

Caridi, A., Sidari, R., Solieri, L., Cufari, A., and Giudici, P. (2007). Wine colour adsorption phenotype: an inheritable quantitative trait loci of yeasts. *J. Appl. Microbiol.* 103, 735–742. doi: 10.1111/j.1365-2672.2007.03301.x

Carlton, W. K., Gump, B., Fugelsang, K., and Hasson, A. S. (2007). Monitoring acetaldehyde concentrations during micro-oxygenation of red wine by headspace solid-phase microextraction with on-fiber derivatization. *J. Agric. Food Chem.* 55, 5620–5625. doi: 10.1021/jf070243b

Chen, J., and Seviour, R. (2007). Medicinal importance of fungal beta-(1->3), (1->6)-glucans. *Mycol. Res.* 111, 635–652. doi: 10.1016/j.mycres.2007.02.011

Chen, K., Escott, C., Loira, I., del Fresno, J. M., Morata, A., Tesfaye, W., et al. (2018). Use of non-*Saccharomyces* yeasts and oenological tannin in red winemaking: influence on colour, aroma and sensorial properties of young wines. *Food Microbiol.* 69, 51–63. doi: 10.1016/j.fm.2017.07.018

Claus, H., and Mojsov, K. (2018). Enzymes for wine fermentation: current and perspective applications. *Fermentation* 4:52. doi: 10.3390/fermentation4030052

Comitini, F., Gobbi, M., Domizio, P., Romani, C., Lencioni, L., Mannazzu, I., et al. (2011). Selected non-*Saccharomyces* wine yeasts in controlled multistarter fermentations with *Saccharomyces cerevisiae.* *Food Microbiol.* 28, 873–882. doi: 10.1016/j.fm.2010.12.001

Cortes, M. B., Moreno, J., Zea, L., Moyano, L., and Medina, M. (1998). Changes in aroma compounds of Sherry wines during their biological aging carried out by *Saccharomyces cerevisiae* races bayanus and capensis. *J. Agric. Food Chem.* 46, 2389–2394. doi: 10.1021/jf970903k

Danilewicz, J. C. (2003). Review of reaction mechanisms of oxygen and proposed intermediate reaction products in wine: central role of iron and copper. *Am. J. Enol. Vitic.* 54, 73–85.

Darriet, P., Thibon, C., and Dubourdieu, D. (2012). "Aroma and aroma precursors in grape berry," in *Aroma and Aroma Precursors in Grape Berry*, eds M. Hernâni Gerós, C. Manuela, and D. Serge Bentham (Sharjah: Science Publishers). doi: 10.2174/97816080536051120101011

David-Vaizant, V., and Alexandre, H. (2018). Flor Yeast diversity and dynamics in biologically aged wines. *Front. Microbiol.* 9:2235. doi: 10.3389/fmicb.2018. 02235

De Iseppi, A., Lomolino, G., Marangon, M., and Curioni, A. (2020). Current and future strategies for wine yeast lees valorization. *Food Res. Int.* 137:109352. doi: 10.1016/j.foodres.2020.109352

Devi, A., Aiyappaa, A. A. K., and Waterhouse, A. L. (2019). Adsorption and biotransformation of anthocyanin glucosides and quercetin glycosides by *Oenococcus oeni* and *Lactobacillus plantarum* in model wine solution. *J. Sci. Food. Agric.* 100, 2110–2120. doi: 10.1002/jsfa.10234

Devi, A., and Ka, A.-A. (2019). Yeast-bacterial interactions during malolactic inoculations affecting anthocyanin adsorption and content in Shiraz wine. *Am. J. Enol. Vitic.* 71, 105–112. doi: 10.5344/ajev.2019.19033

Domizio, P., Lencioni, L., Calamai, L., Portaro, L., and Bisson, L. F. (2018). Evaluation of the yeast Schizosaccharomyces japonicus for use in wine production. *Am. J. Enol. Vitic.* 69, 266–277. doi: 10.5344/ajev.2018.18004

Domizio, P., Liu, Y., Bisson, L. F., and Barile, D. (2014). Use of non-Saccharomyces wine yeasts as novel sources of mannoproteins in wine. *Food Microbiol.* 43, 5–15. doi: 10.1016/j.fm.2014.04.005

Domizio, P., Liu, Y., Bisson, L. F., and Barile, D. (2017). Cell wall polysaccharides released during the alcoholic fermentation by *Schizosaccharomyces pombe* and *S. japonicus*: quantification and characterization. *Food Microbiol.* 61, 136–149. doi: 10.1016/j.fm.2016.08.010

Duc, C., Pradal, M., Sanchez, I., Noble, J., Tesnière, C., and Blondin, B. (2017). A set of nutrient limitations trigger yeast cell death in a nitrogen-dependent manner during wine alcoholic fermentation. *PLoS One* 12:e0184838. doi: 10. 1371/journal.pone.0184838

Echeverrigaray, S., Menegotto, M., and Longaray Delamare, A. P. (2019). A simple and reliable method for the quantitative evaluation of anthocyanin adsorption

by wine yeasts. *J. Microbiol. Methods* 157, 88–92. doi: 10.1016/j.mimet.2018.12.016

Echeverrigaray, S., Scariot, F. J., Menegotto, M., and Longaray Delamare, A. P. (2020). Anthocyanin adsorption by *Saccharomyces cerevisiae* during wine fermentation is associated to the loss of yeast cell wall/membrane integrity. *Int. J. Food Microbiol.* 314:108383. doi: 10.1016/j.ijfoodmicro.2019.108383

Englezos, V., Rantsiou, K., Cravero, F., Torchio, F., Giacosa, S., Ortiz-Julien, A., et al. (2018). Volatile profiles and chromatic characteristics of red wines produced with *Starmerella bacillaris* and *Saccharomyces cerevisiae*. *Food Res. Int.* 109, 298–309. doi: 10.1016/j.foodres.2018.04.027

Escot, S., Feuillat, M., Dulau, L., and Charpentier, C. (2001). Release of polysaccharides by yeast and the influence of polysaccharides on colour stability and wine astringency. *Aust. J. Grape Wine Res.* 7, 153–159. doi: 10.1111/j.1755-0238.2001.tb00204.x

Escott, C., del Fresno, J. M., Loira, I., Morata, A., Tesfaye, W., González, M. C., et al. (2018). Formation of polymeric pigments in red wine through sequential fermentation of flavanol-enriched musts with non-*Saccharomyces* yeasts. *Food Chem.* 239, 975–983. doi: 10.1016/j.foodchem.2017.07.037

Escribano-Viana, R., Portu, J., Garijo, P., López, R., Santamaría, P., López-Alfaro, I., et al. (2019). Effect of the sequential inoculation of non-*Saccharomyces*/*Saccharomyces* on the anthocyans and stilbenes composition of Tempranillo wines. *Front. Microbiol.* 10:773. doi: 10.3389/fmicb.2019.00773

Fabios, M., Lopez-Toledano, A., Mayen, M., Merida, J., and Medina, M. (2000). Phenolic compounds and browning in Sherry wines subjected to oxidativeand biological ageing. *J. Agric. Food Chem.* 48, 2155–2159. doi: 10.1021/jf9908502

Fernando, J. G., Pedro, A. R. F., Dulcineia, F. W., Susana, M. C., Silvia, M. R., and Manuel, A. C. (2018). Interaction of wine mannoproteins and arabinogalactans with anthocyanins. *Food Chem.* 243, 1–10.

Ferrando, N., Araque, I., Ortís, A., Thornes, G., Bautista-Gallego, J., Bordons, A., et al. (2020). Evaluating the effect of using non-Saccharomyces on Oenococcus oeni and wine malolactic fermentation. *Food Res. Int.* 138(Pt B), 109779. doi: 10.1016/j.foodres.2020.109779

Forino, M., Picariello, L., Lopatriello, A., Moio, L., and Gambuti, A. (2020). New insights into the chemical bases of wine color evolution and stability: the key role of acetaldehyde. *Eur. Food Res. Technol.* 246, 733–743. doi: 10.1007/s00217-020-03442-x

Francois, J. M. (2016). Cell surface interference with plasma membrane and transport processes in yeasts. *Adv. Exp. Med. Biol.* 892, 11–31. doi: 10.1007/978-3-319-25304-6_2

Fulcrand, H., Benabdeljalil, C., Rigaud, J., Chenyier, V., and Moutounet, M. (1998). A new class of wine pigments generated by reaction between pyruvic acid and grape anthocyanins. *Phytochemistry* 47, 1401–1407. doi: 10.1016/S0031-9422(97)00772-3

Giovinazzo, G., Carluccio, M. A., and Grieco, F. (2019). "Wine polyphenols and health," in *Bioactive Molecules in Food—Reference Series in Phytochemistry*, eds J. M. Mérillon and K. G. Ramawat (Basel: Springer). doi: 10.1007/978-3-319-78030-6_81

Gobbi, M., Comitini, F., Domizio, P., Romani, C., Lencioni, L., Mannazzu, I., et al. (2013). *Lachancea thermotolerans* and *Saccharomyces cerevisiae* in simultaneous and sequential co-fermentation: a strategy to enhance acidity and improve the overall quality of wine. *Food Microbiol.* 33, 271–281. doi: 10.1016/j.fm.2012.10.004

Goldner, C. M., and Zamora, C. M. (2007). Sensory characterization of *Vitis vinifera* cv. Malbec wines from seven viticulture regions of Argentina. *J. Sens. Stud.* 22, 520–532. doi: 10.1111/j.1745-459X.2007.00123.x

Gómez-Plaza, E., and Cano-López, M. (2011). A review on microoxygenation of red wines: claims, benefits and the underlying chemistry. *Food Chem.* 125, 1131–1140. doi: 10.1016/j.foodchem.2010.10.034

Gonçales, F. J., Fernandes, P. A. R., Wessel, D. F., Cardoso, S. M., Rocha, S. M., and Coimbra, M. (2018). Interaction of wine mannoproteins and arabinogalactans with anthocyanins. *Food Chem.* 243, 1–10. doi: 10.1016/j.foodchem.2017.09.097

Gonzalez, S. S., Barrio, E., Gafner, J., and Querol, A. (2006). Natural hybrids from *Saccharomyces cerevisiae*, *Saccharomyces bayanus* and *Saccharomyces kudriavzevii* in wine fermentations. *FEMS Yeast Res.* 6, 1221–1234. doi: 10.1111/j.1567-1364.2006.00126.x

Grieco, F., Carluccio, M. A., and Giovinazzo, G. (2019). Autochthonous *Saccharomyces cerevisiae* starter cultures enhance polyphenols content, antioxidant activity, and anti-inflammatory response of Apulian red wines. *Foods* 8:453. doi: 10.3390/foods8100453

Herderich, M. J., Siebert, T. E., Parker, M., Capone, D. L., Mayr, C., Zhang, P., et al. (2013). Synthesis of the ongoing works on rotundone, an aromatic compound responsible of the peppery notes in wines. *Internet J. Enol. Vitic.* 6, 1–6.

Izquierdo-Cañas, P. M., García-Romero, E., Mena-Morales, A., and Gómez-Alonso, S. (2016). Effects of malolactic fermentation on colour stability and phenolic composition of Petit Verdot red wines. *Wine Stud.* 5:5795. doi: 10.4081/ws.2016.5795

Jaganath, I. B., and Crozier, A. (2010). "Dietary flavonoids and phenolic compounds," in *Plant Phenolics and Human Health*, ed. C. G. Fraga (Hoboken, NY: John Wiley & Sons, Inc).

Jagatić Korenika, A. M., Tomaz, I., Preiner, D., Plichta, V., and Jeromel, A. (2021). Impact of commercial yeasts on phenolic profile of Plavac Mali wines from Croatia. *Fermentation* 7:92. doi: 10.3390/fermentation7020092

Jolly, N. P., Varela, C., and Pretorius, I. S. (2014). Not your ordinary yeast: non-Saccharomyces yeasts in wine production uncovered. *FEMS Yeast Res.* 14, 215–237. doi: 10.1111/1567-1364.12111

Klis, F. M., Boorsma, A., and De Groot, P. W. J. (2006). Cell wall construction in *Saccharomyces cerevisiae*. *Yeast* 23, 185–202. doi: 10.1002/yea.1349

Klis, F. M., Mol, P., Helligwerf, K., and Brul, S. (2002). Dynamics of cell wall structure of *Saccharomyces cerevisiae*. *FEMS Microb. Rev.* 26, 239–256. doi: 10.1111/j.1574-6976.2002.tb00613.x

Kogani, G., Pajtinka, M., Babincova, M., Miadokova, E., Rauko, P., Slamenova, D., et al. (2008). Yeast cell wall polysaccharides as antioxidants and antimutagens: can they fight cancer? Minireview. *Neoplasma* 55:387.

Kollár, R., Reinhold, B. B., Petráková, E., Yeh, H. J., Ashwell, G., Drgonová, J., et al. (1997). Architecture of the yeast cell wall. Beta (1–>6)-glucan interconnects mannoprotein, beta(1–>)3-glucan, and chitin. *J Biol. Chem.* 272, 17762–17775. doi: 10.1074/jbc.272.28.17762

Legras, J. L., Moreno-Garcia, J., Zara, S., Zara, G., Garcia-Martinez, T., Mauricio, J. C., et al. (2016). Flor Yeast: new perspectives beyond wine aging. *Front. Microbiol.* 7:503. doi: 10.3389/fmicb.2016.00503

Lencioni, L., Taccari, M., Ciani, M., and Domizio, P. (2018). Zygotorulaspora florentina and *Starmerella bacillaris* in multistarter fermentation with *Saccharomyces cerevisiae* to reduce volatile acidity of high sugar musts. *Aust. J. Grape Wine Res.* 24, 368–372. doi: 10.1111/ajgw.12327

Lesage, G., and Bussey, H. (2006). Cell wall assembly in Saccharomyces cerevisiae. *Microbiol. Mol. Biol. Rev.* 70, 317–343. doi: 10.1128/MMBR.00038-05

Liu, S., Laaksonen, O., and Yang, B. (2019). Volatile composition of bilberry wines fermented with non-*Saccharomyces* and *Saccharomyces* yeasts in pure, sequential and simultaneous inoculations. *Food Microbiol.* 80, 25–39. doi: 10.1016/j.fm.2018.12.015

Liu, S. Q., and Pilone, G. J. (2000). An overview of formation and roles of acetaldehyde in winemaking with emphasis on microbiological implications. *Intern. J. Food Sci. Technol.* 35, 49–61. doi: 10.1046/j.1365-2621.2000.00341.x

Lonvaud-Funel, A. (1999). Lactic acid bacteria in the quality improvement and depreciation of wine. *Ant. Van Leeuw.* 76, 317–331. doi: 10.1023/A:1002088931106

Mansfield, A. K., Zoecklein, B. W., and Whiton, R. S. (2002). Quantification of glycosidase activity in selected strains of *Brettanomyces bruxellensis* and *Oenococcus oeni*. *Am. J. Enol. Vitic.* 53, 303–307.

Manzanares, P., Rojas, V., Genovés, S., and Vallés, S. (2000). A preliminary search for anthocyanin-ß-D-glucosidase activity in non-*Saccharomyces* wine yeasts. *Int. J. Food Sci. Technol.* 35, 95–103. doi: 10.1046/j.1365-2621.2000.00364.x

Mazauric, J. P., and Salmon, J. M. (2005). Interactions between yeast lees and wine poly- phenols during simulation of wine aging: I. Analysis of remnant polyphenolic com- pounds in the resulting wines. *J. Agric. Food Chem.* 53, 5647–5653. doi: 10.1021/jf050308f

Mazauric, J. P., and Salmon, J. M. (2006). Interactions between yeast lees and wine poly- phenols during simulation of wine aging: II. Analysis of desorbed polyphenol com- pounds from yeast lees. *J. Agric. Food Chem.* 54, 3876–3881. doi: 10.1021/jf060037o

Medina, K., Boido, E., Dellacassa, E., and Carrau, F. (2005). Yeast interactions with anthocyanins during red wine fermentation. *Am. J. Enol. Vitic.* 56, 104–109.

Mekoue Nguela, J., Vernhet, A., Julien-Ortiz, A., Sieczkowski, N., and Mouret, J. R. (2019). Effect of grape must polyphenols on yeast metabolism during alcoholic fermentation. *Food Res. Int.* 121, 161–175. doi: 10.1016/j.foodres.2019.03.038

Merida, J., Lopez-Toledano, A., Marquez, T., Millan, C., Ortega, J. M., and Medina, M. (2005). Retention of browning compounds by yeasts involved in the winemaking of sherry type wines. *Biotechnol. Lett.* 27, 1565–1570. doi: 10.1007/s10529-005-1795-9

Morata, A., Benito, S., Loira, L., Palomero, E, González, M. C., and Suárez-Lepe, J. A. (2012). Formation of pyranoanthocyanins by *Schizosaccharomyces pombe* during the fermentation of red must. *Int. J. Food Microbiol.* 159, 47–53. doi: 10.1016/j.ijfoodmicro.2012.08.007

Morata, A., Gómez-Cordovés, M. C., Calderón, F., and Suárez, J. A. (2006). Effects of pH, temperature and SO₂ on the formation of pyranoanthocyanins during red wine fermentation with two species of *Saccharomyces*. *Int. J. Food Microbiol.* 106, 123–129. doi: 10.1016/j.ijfoodmicro.2005.05.019

Morata, A., Gómez-Cordovés, M. C., Colomo, B., and Suárez, J. A. (2003). Pyruvic acid and acetaldehyde production by different strains of *Saccharomyces cerevisiae*: relationship with Vitisin A and B formation in red wines. *J. Agric. Food Chem.* 51, 7402–7409. doi: 10.1021/jf0304167

Morata, A., González, C., and Suárez-Lepe, J. A. (2007). Formation of vinylphenolic pyranoanthocyanins by selected yeasts fermenting red grape musts supplemented with hydroxycinnamic acids. *Int. J. Food Microbiol.* 116, 144–152. doi: 10.1016/j.ijfoodmicro.2006.12.032

Morata, A., Loira, I., Escott, J. M., del Fresno, J. M., Bañuelos, M. A., and Suárez-Lepe, J. A. (2019). Applications of Metschnikowia pulcherrima in wine biotechnology. *Fermentation* 5:63. doi: 10.3390/fermentation5030063

Morata, A., Loira, I., Heras, J. M., Callejo, M. J., Tesfaye, W., González, C., et al. (2016). Yeast influence on the formation of stable pigments in red winemaking. *Food Chem.* 197, 686–691. doi: 10.1016/j.foodchem.2015.11.026

Nardi, T. (2020). Microbial resources as a tool for enhancing sustainability in winemaking. *Microorganisms* 8:507. doi: 10.3390/microorganisms8040507

Nguyen, T. H., Fleet, G. H., and Rogers, P. L. (1998). Composition of the cell walls of several yeast species. *Appl. Microbiol. Biotechnol.* 50, 206–212. doi: 10.1007/s002530051278

Olejar, K. J., Fedrizzi, B., and Kilmartin, P. A. (2015). Influence of harvesting technique and maceration process on aroma and phenolic attributes of Sauvignon blanc wine. *Food Chem.* 183, 181–189. doi: 10.1016/j.foodchem.2015.03.040

Padilla, B., Gil, J. V., and Manzanares, P. (2016). Past and future of Non-*Saccharomyces* yeasts: from spoilage microorganisms to biotechnological tools for improving wine aroma complexity. *Front. Microbiol.* 7:411. doi: 10.3389/fmicb.2016.00411

Palomero, F., Morata, A., Benito, S., Calderón, F., and Suárez-Lepe, J. A. (2009). New genera of yeasts for over-lees aging of red wine. *Food Chem.* 112, 432–441. doi: 10.1016/j.foodchem.2008.05.098

Peter, J., De Chiara, M., Friedrich, A., Yue, J. X., Pflieger, D., Bergström, A., et al. (2018). Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* 556, 339–344. doi: 10.1038/s41586-018-0030-5

Petruzzi, L., Baiano, A., De Gianni, A., Sinigaglia, M., Corbo, M. R., and Bevilacqua, A. (2015). Differencial adsorption of Ochratoxin A and anthocyanins by inactivated yeasts and yeast cell walls during simulation of wine aging. *Toxins* 7, 4350–4365. doi: 10.3390/toxins7104350

Pfliegler, W. P., Pusztahelyi, T., and Pocsi, I. (2015). Mycotoxins - prevention and decontamination by yeasts. *J. Basic Microbiol.* 55, 805–818. doi: 10.1002/jobm.201400833

Renault, P., Coulon, J., Moine, V., Thibon, C., and Bely, M. (2016). Enhanced 3-sulfanylhexan-1-ol production in sequential mixed fermentation with *Torulaspora delbrueckii/Saccharomyces cerevisiae* reveals a situation of synergistic interaction between two industrial strains. *Front. Microbiol.* 7:293. doi: 10.3389/fmicb.2016.00293

Renault, P., Miot-Sertier, C., Marullo, P., Lagarrigue, L., Lonvaud-Funel, A., and Bely, M. (2009). Genetic characterization and phenotypic variability in *Torulaspora delbrueckii* species: potential applications in the wine industry. *Int. J. Food Microbiol.* 134, 201–210. doi: 10.1016/j.ijfoodmicro.2009.06.008

Ribereau-Gayon, P., Glories, Y., and Maujean, A. (2006). *Handbook of Enology*. West Sussex: John Wiley & Sons, Ltd. doi: 10.1002/0470010398

Rinaldi, A., Coppola, M., and Moio, L. (2019). Aging of aglianico and sangiovese wine on mannoproteins: effect on astringency and color. *LWT Food Sci. Technol.* 105, 233–241. doi: 10.1016/j.lwt.2019.02.034

Romano, P., and Suzzi, G. (1993). "Sulphur dioxide and wine microorganisms," in *Wine Microbiology and Biotechnology*, ed. G. Fleet Harwood (Chur: Academic Publisher GmbH).

Sadoudi, M., Tourdot-Maréchal, R., Rousseaux, S., Steyer, D., Gallardo-Chacón, J. J., Ballester, J., et al. (2012). Yeast–yeast interactions revealed by aromatic profile analysis of Sauvignon Blanc wine fermented by single or co-culture of non-*Saccharomyces* and *Saccharomyces* yeasts. *Food Microbiol.* 32, 243–253. doi: 10.1016/j.fm.2012.06.006

Samoticha, J., Wojdyło, A., Chmielewska, J., and Nofer, J. (2019). Effect of different yeast strains and temperature of fermentation on basic enological parameters, polyphenols and volatile compounds of Aurore white wine. *Foods* 8:599. doi: 10.3390/foods8120599

Schiavone, M., Déjean, S., Sieczkowski, N., Castex, M., Dague, E., and François, J. M. (2017). Integration of biochemical, biophysical and transcriptomics data for investigating the structural and nanomechanical properties of the yeast cell wall. *Front. Microbiol.* 8:1806. doi: 10.3389/fmicb.2017.01806

Schiavone, M., Vax, A., Formosa, C., Martin-Yken, H., Dague, E., and Francois, J. M. (2014). A combined chemical and enzymatic method to determine quantitatively the polysaccharide components in the cell wall of yeasts. *FEMS Yeast Res.* 14, 933–947. doi: 10.1111/1567-1364.12182

Schwarz, M., Wabnitz, T. C., and Winterhalter, P. (2003). Pathway leading to the formation of anthocyanin-vinylphenol adducts and related pigments in red wines. *J. Agric. Food Chem.* 51, 3682–3687. doi: 10.1021/jf0340963

Setati, M. E., Jacobson, D., Andong, U. C., and Bauer, F. F. (2012). The vineyard yeast microbiome, a mixed model microbial map. *PLoS One* 7:e52609. doi: 10.1371/journal.pone.0052609

Simonin, S., Alexandre, H., Nikolantonaki, M., Coelho, C., and Tourdot-Maréchal, R. (2018). Inoculation of *Torulaspora delbrueckii* as a bio-protection agent in winemaking. *Food Res. Int.* 107, 451–461. doi: 10.1016/j.foodres.2018.02.034

Stanley, D., Bandara, A., Fraser, S., Cambers, P. J., and Stanley, G. A. (2010). The ethanol stress response and ethanol tolerance of *Saccharomyces cerevisiae*. *J. Appl. Microbiol.* 109, 13–24. doi: 10.1111/j.1365-2672.2009.04657.x

Sumby, K., Bartle, L., Grbin, P., and Jiranek, V. (2019). Measures to improve wine malolactic fermentation. *Appl. Microbiol. Biotechnol.* 103, 2033–2051. doi: 10.1007/s00253-018-09608-8

Suzzi, G., Romano, P., and Zambonelli, C. (1985). Saccharomyces strain selection in minimizing SO2 requirement during vinification. *Am. J. Enol. Vitic.* 36, 199–202.

Timberlake, C. F., and Bridle, P. (1976). Interactions between anthocyanins, phenolic compounds, and acetaldehyde and their significance in red wines. *Am. J. Enol. Vitic.* 27, 97–105.

Tofalo, R., Patrignani, F., Lanciotti, R., Perpetuini, G., Schirone, M., Di Gianvito, P., et al. (2016). Aroma profile of Montepulciano d'Abruzzo wine fermented by single and co-culture starters of autochthonous *Saccharomyces* and non-*Saccharomyces* yeasts. *Front. Microbiol.* 7:610. doi: 10.3389/fmicb.2016.00610

Tofalo, R., Schirone, M., Torriani, S., Rantsiou, K., Cocolin, L., Perpetuini, G., et al. (2012). Diversity of *Candida zemplinina* strains from grapes and Italian wines. *Food Microbiol.* 29, 18–26. doi: 10.1016/j.fm.2011.08.014

Van Breda, V., Jolly, N., and Van Wyk, J. (2013). Characterisation of commercial and natural *Torulaspora delbrueckii* wine yeast strains. *Int. J. Food Microbiol.* 163, 80–88. doi: 10.1016/j.ijfoodmicro.2013.02.011

Vasserot, Y., Caillet, S., and Maujean, A. (1997). Study of anthocyanin adsorption by yeast lees. Effect of some physicochemical parameters. *Am. J. Enol. Vitic.* 48, 433–437.

Vejarano, R. (2020). Non-*Saccharomyces* in winemaking: source of mannoproteins, nitrogen, enzymes, and antimicrobial compounds. *Fermentation* 6:76. doi: 10.3390/fermentation6030076

Verstrepen, K. J., and Fink, G. R. (2009). Genetic and epigenetic mechanisms underlying cell-surface variability in protozoa and fungi. *Annu. Rev. Genet.* 43:1. doi: 10.1146/annurev-genet-102108-134156

Vilela, A. (2018). *Lachancea thermotolerans*, the non-*Saccharomyces* yeast that reduces the volatile acidity of wines. *Fermentation* 4:56. doi: 10.3390/fermentation4030056

Virdis, C., Sumby, K., Bartowsky, E., and Jiranek, V. (2021). Lactic acid bacteria in wine: technological advances and evaluation of their functional role. *Front. Microbiol.* 11:612118. doi: 10.3389/fmicb.2020.612118

Visioli, F., Panaite, S. A., and Tomé-Carneiro, J. (2020). Wine's phenolic compounds and health: a pythagorean view. *Molecules* 25:4105. doi: 10.3390/molecules25184105

Wang, S., Li, S., Zhao, H., Gu, P., Chen, Y., Zhang, B., et al. (2018). Acetaldehyde released by *Lactobacillus plantarum* enhances accumulation of pyranoanthocyanins in wine during malolactic fermentation. *Food Res. Int.* 108, 254–263. doi: 10.1016/j.foodres.2018.03.032

Waterhouse, A. L., Sacks, G. L., and Jeffery, D. W. (2016). *Understanding Wine Chemistry.* New York, NY: JohnWiley & Sons, Incorporated. doi: 10.1002/9781118730720

Yin, Q. Y., De Groot, P. W. J., De Jong, L., Klis, F. M., and De Koster, C. G. (2007). Mass spectrometric quantitation of covalently bound cell wall proteins in *Saccharomyces cerevisiae. FEMS Yeast Res.* 7, 887–896. doi: 10.1111/j.1567-1364.2007.00272.x

Yue, X. F., Jing, S. S., Ni, X. F., Zhang, K. K., Fang, Y. L., Zhang, Z. W., et al. (2021). Anthocyanin and phenolic acids contents influence the color stability and antioxidant capacity of wine treated with mannoprotein. *Front. Nutr.* 8:691784. doi: 10.3389/fnut.2021.691784

Zara, S., Gross, M. K., Zara, G., Budroni, M., and Bakalinsky, A. T. (2010). Ethanol-independent biofilm formation by a flor wine yeast strain of *Saccharomyces cerevisiae. Appl. Environ. Microbiol.* 76, 4089–4091. doi: 10.1128/AEM.00111-10

Zhang, P., Ma, W., Meng, Y., Zhang, Y., Jin, G., and Fang, Z. (2021). Wine phenolic profile altered by yeast: mechanisms and influences. *Compr. Rev. Food Sci. Food Saf.* 20, 3579–3619. doi: 10.1111/1541-4337.12788

Differential Survivability of Two Genetically Similar *Salmonella* Thompson Strains on Pre-harvest Sweet Basil (*Ocimum basilicum*) Leaves

Ye Htut Zwe, Michelle Mei Zhen Ten, Xinyi Pang, Chun Hong Wong and Dan Li*

Department of Food Science and Technology, National University of Singapore, Singapore, Singapore

Although conventionally considered an animal pathogen, recent evidence increasingly suggests that fresh produce may act as significant transmission vehicles and alternative hosts to *Salmonella*. This study reports the differential survivability of two genetically similar *Salmonella* Thompson strains (ST 889B and ST 688C) on the adaxial surface of pre-harvest basil (*Ocimum basilicum*) leaves. Upon inoculation, two distinct phenomena, a dried water-print or a macroscopic lesion, were observed within 24 h. ST 889B survived better than ST 688C on healthy-looking leaves without lesions, possibly due to its higher biofilm-forming ability. Both strains survived better on the leaves with lesions than on the healthy-looking leaves (ST 688C: 4.39 ± 0.68 vs. 2.18 ± 0.29; ST 889B: 4.78 ± 0.12 vs. 2.83 ± 0.18 log CFU per sample at 6 days post-inoculation). ST 889B caused the formation of lesions at a higher frequency [70/117 leaves (59.8%)] than ST 688C [35/96 leaves (36.5%)]. Thus, we highlighted two distinct *Salmonella* survival strategies in the basil pathosystem and demonstrated gene expression polymorphism (variations in the expression of the same set of genes) as an indispensable strategy in the colonization of plants as hosts by the human pathogens.

Keywords: *Salmonella* Thompson, pre-harvest basil, hypersensitive response, expression polymorphism, type III secretion system

INTRODUCTION

Salmonella enterica, which causes non-typhoidal salmonellosis, is an important foodborne pathogen estimated to cause up to 11% of the 9.4 million foodborne illnesses in the United States annually. It is also one of the most virulent and fatal foodborne pathogens, causing up to 35 and 28% of all foodborne-related hospitalizations and deaths, respectively, in the United States (Scallan et al., 2011). Its impact is also felt in smaller nations like Singapore, where 68% of all reported food- and waterborne disease cases consisted of non-typhoidal salmonellosis (Zwe and Yuk, 2017). Fresh produce has been identified as an important vehicle for the transmission of *Salmonella*. Culinary herbs such as basil have drawn particular attention for their food safety since once fresh culinary herbs are contaminated, completely removing or killing pathogens is unlikely. Large-scale *S.* Senftenberg

Function	Primers	5' to 3'	Product size
Biofilm	bcsA_F	GTCCCACATATCGTTACCGTCCTG	119
	bcsA_R	CGCCGCATCATTTCTTCTCCC	
	wcaD_F	CTGCCGCCAGTAAGGATAAT	111
	wcaD_R	CATTTCCGCGCATAACCACC	
	csgB_F1	GGTCAAGTCGGCACGGATAA	103
	csgB_R1	GGTCGACTTTCGCCCGATTA	
	bapA_F	CGGTGAATTCGTCGTTACGC	115
	bapA_R	GTCGGAAGCGGGAAAATTCG	
SPI-1 T3SS	prgH_F	GCTAAAACCTGACGAGCCCA	127
	prgH_R	CCGCAGAGCTCGATTCGTTA	
PAMP signature	fliC_F	GTTCAACGGCGTGAAAGTCC	115
	fliC_R	ACCCAGGGTCTGAGAGTTGA	
Housekeeping	hcaT_F	CCTGCAAACGAATCACCT	113
	hcaT_R	GCCGTGGCTGATTGTGATA	

SPI-1, Salmonella pathogenicity island-1; T3SS, type III secretion system; PAMP, pathogen-associated molecular patterns.

and *S.* Anatum outbreaks have been reported in recent history due to contaminated fresh basil consumption in multiple countries (Pezzoli et al., 2008; Pakalniskiene et al., 2009). Therefore, it is crucial to understand how *Salmonella* survives in the pre-harvest basil plants for its food safety implications.

In this study, we report the differential capabilities of two *S.* Thompson isolates in colonizing and surviving on the basil plants. ST 889B was isolated from basil in Belgium in 2013 (Delbeke et al., 2015), and ST 688C was a clonal subculture of the 1999 California coriander outbreak clinical-isolate RM1894 (Li and Uyttendaele, 2018). These two *S.* Thompson isolates were of near-identical genetic makeup and thus served as valuable study materials to examine the plant–pathogen interactions between *Salmonella* and fresh produce, offering insights that may be critical to improving the microbial safety of plant produce.

MATERIALS AND METHODS

Bacterial Strain and Culture Conditions

ST 688C is a clonal subculture of the *S.* Thompson strain RM1984 clinically isolated from a salmonellosis patient during the California coriander outbreak in 1999, kindly provided by Dr. Maria Brandl (U.S. Department of Agriculture, Agricultural Research Service, Albany, CA, United States). ST 889B was previously isolated from basil (Delbeke et al., 2015), kindly provided by Prof. Mieke Uyttendaele (Ghent University). The frozen cultures of bacteria were activated by transferring twice consecutively (37°C, 24 h) in sterile tryptic soy broth (TSB, Oxoid, Basingstoke, Hampshire, England). The stock cultures of bacteria were maintained on tryptic soy broth (TSA, Oxoid) plates stored at 4°C. A single colony from TSA was transferred into 10-ml TSB and incubated at 37°C for 24 h. A 1-ml portion of the bacterial suspension was centrifuged at 9,000 × *g* at 4°C, washed twice in phosphate-buffered saline (PBS; Vivantis, Inc., Oceanside, CA, United States), serially diluted in 0.1% peptone

water (PW; Oxoid) to appropriate concentration to obtain the working cultures of bacteria for downstream experiments.

Whole Genome Sequencing and Bioinformatic Analysis

The DNA from the 24-h cultures of bacteria were extracted using the GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific, Vilnius, Lithuania) following the manufacturer's instructions. Whole genome sequencing (WGS) was performed by NovogeneAIT Genomics Singapore Pte Ltd. using Illumina Hiseq4000, and 350 bp insert DNA library preparation.

Raw reads were assembled into contigs using Assembler 1.2 publicly available at the Center for Genomic Epidemiology by the Technical University of Denmark[1] using default parameters. Assembled contigs were analyzed for resistance determinants, multilocus sequence type (MLST), and core genome multilocus sequence type (cgMLST) using ResFinder 4.0 (Zankari et al., 2012; ResFinder Database 2020-04-08 and PointFinder Database 2019-07-02), MLST 2.0 (Larsen et al., 2012), and cgMLSTFinder 1.1 (Alikhan et al., 2018), respectively, using default parameters. The average nucleotide identity between ST 688C and ST 889B was analyzed using the ANI Calculator (Yoon et al., 2017). The contigs were deposited to the GenBank, National Center for Biotechnology Information (NCBI)[2] under accession numbers JAFFRV000000000 (ST 688C) and JAFDOS000000000 (ST 889B) of BioProject PRJNA697834.

Single-nucleotide polymorphism tree for phylogenetic analysis was performed using the CSI Phylogeny 1.4 (Kaas et al., 2014) using default parameters. All known genome assemblies of 46 *S.* Thompson strains were downloaded from the NCBI. Three genomes were found to be outgroups from the others in an initial trial and were subsequently excluded. A group of 43 genomes, ST 688C, and ST 889B, were used to construct a phylogenetic tree based on SNPs with *S.* Thompson RM6836 as the reference. The resulting tree was visualized using the public available FigTree 1.4.4 software.[3] Bootstrap values displayed were automatically calculated by the proprietary algorithm of the CSI Phylogeny 1.4 software.

Salmonella Survivability in Sterile Basil Juice

A 100 g portion of *Ocimum basilicum* sweet basil leaves and stems from the freshly bought basil plants were immediately harvested and added to 200 ml of deionized (DI) water and homogenized in a food processor. The solid particles in the resultant suspension were removed by first filtering through a mesh strainer, followed by multiple rounds of centrifugation (9,000 × *g*, 4°C, 10 min) until no visible solid particles were present throughout the liquid. The basil juice was sterilized by passing through 0.22 μm filters.

A 100 μl portion of the working culture of ST 688C or ST 889B (*ca.* 10⁹ CFU/ml) in 0.1% PW was inoculated into 30 ml

[1]https://cge.cbs.dtu.dk/services/Assembler/
[2]https://www.ncbi.nlm.nih.gov/
[3]http://tree.bio.ed.ac.uk/software/figtree/

of sterile basil juice in 50-ml falcon tubes. The samples were stored loosely capped at 25°C over 3 days. A 100 μl portion of the sample was drawn from each sample at 0 (immediately after inoculation), 1, 2, and 3 days post-inoculation (dpi), serially diluted in 0.1% PW and enumerated on xylose lysine deoxycholate (XLD; Oxoid) agar.

Biofilm-Forming Ability Assay

As described by Lee et al. (2011), crystal violet biofilm assay was employed with slight modifications to determine the biofilm-forming ability of ST 688C and ST 889B. Briefly, working cultures of bacterial cells were serially diluted in 0.1% PW while the final dilution was performed in either 1/10 TSB (TSB:water of 1:9) or sterile basil juice to a final concentration of 10^5 CFU/ml. For every combination of bacterial strain and medium, a 200 μl portion of the sample per well was added to six wells in a 96-well plate. After incubating at 25°C for 24 h, each well was washed three times with sterile DI water, stained with 200 μl of 0.1% (wt/vol) crystal violet (Sigma-Aldrich, St. Louis, MO, United States) solution for 20 min at room temperature, rewashed three times with DI water, and incubated with 200 μl of 95% ethanol at 4°C for 45 min to extract dye from the biofilm. Absorption at 595 nm was measured using a Multiskan FC microplate photometer (Thermo Fisher Scientific, Shanghai, China).

Construction of GFP-Tagged Strains and Fluorescence Confocal Microscopy

The procedures for preparing competent *Salmonella* cells and the construction of GFP-tagged *Salmonella* strains through the transformation with pGFPuv plasmid have been previously described (Ma et al., 2011). Briefly, a 250 μl portion of overnight cultures of ST 889B incubated at 37°C in at 37°C Luria-Bertani (LB, Invitrogen, CA, United States) broth was transferred into 25 ml of LB broth and incubated at 37°C for 3 h in the shaker incubator shaking at 200 rpm. The resultant bacterial suspension was then cooled in ice for 10 min, centrifuged (4,000 rpm, 4°C, 10 min), resuspended in 10 ml of 0.1 M $CaCl_2$, centrifuged, and resuspended again in 2 ml of 0.1 M $CaCl_2$ with 15% glycerol medium. The pGFPuv plasmid was extracted using the Qiagen plasmid miniprep kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. The transformation was performed by the heat-shock method. Briefly, 5 μl of the plasmid was mixed into 50 μl of competent cells and kept on ice for 30 min before transferring to a 42°C incubator for 30 s and transferring back on ice for 5 min. A 950 μl portion of LB broth was added to the heat-shocked bacterial-plasmid mix and incubated at 37°C for 1 h while shaking at 250 rpm. The suspension was plated onto LB agar with 100 μg/ml ampicillin to recover the transformants. The GFP-tagged bacteria were grown in TSB with 100 μg/ml ampicillin before washing and appropriately diluted as previously described for inoculation. The observation of GFP-tagged bacteria was performed with fluorescence confocal microscopy Olympus Fluoview FV1000 using the emission band 505–525 nm and the autofluorescence from the leaves, using emission band 525–650 nm.

Salmonella Survivability on Pre-harvest Basil (*Ocimum basilicum*) Leaves

Pre-harvest *O. basilicum* sweet basil plants (*ca.* 2 months of age) were purchased from a local supermarket. A 300-W LED lamp (Melonfarm) was used in the 32″ × 32″ × 63″ grow tent (MelonFarm, China) to supply light with a photoperiod of 18 h and a dark period of 6 h. The temperature (26 ± 2°C) and relative humidity (58 ± 8%) in the grow tent were monitored using an EL-USB-2-LCD+ logger (Lascar Electronics Ltd., United Kingdom). Newly purchased plants were subjected to a 24-h acclimatization period to the new environment in the grow tent before inoculation with bacteria.

A 100 μl portion working culture of ST 688C or ST 889B (*ca.* 10^9 CFU/ml) in PBS was inoculated onto the adaxial side of an inverted basil leaf (to form a "cup-like" depression to assist inoculation) in a roughly 1 cm² area and dried ambiently. Immediately after drying (0 dpi) and at other time points (1 and 6 dpi), leaves for each bacterial strain were harvested to be enumerated. The 1 cm² portion from each leaf where the inoculum has dried was cut and mixed with 5 ml of PBS in a small vacuum-sealed stomacher pouch. The leaf's surface was lightly massaged for 1 min, followed by a light crushing using a pestle to release any bacterial cells that may be internalized within the leaf structure. The resultant suspension was then serially diluted to the appropriate concentration in 0.1% PW and spread plated on XLD (Oxoid) agar plates and incubated at 37°C for 24 h.

Gene Expression Assay

RNA from freshly dried bacterial inoculum on basil leaves was extracted using the RNeasy Mini Kit (Qiagen) and was reverse transcribed using the GoScript™ Reverse Transcriptase kit (Promega Corp., Madison, WI, United States) following the manufacturer's instructions to obtain the cDNA. The primers used in this study were designed using the Primer-BLAST tool[4] and are described in **Table 1**. The single-copy chromosomal gene *hcaT* was used as the reference gene (Barak et al., 2007). The assay was performed in 10 μl reaction volumes with 5 μl of GoTaq™ qPCR Master Mix (Promega), 0.25 μl each of the forward and reverse primer at 10 μmol/l concentration, and 1 μl of cDNA template. Threshold cycle (C_T) values were measured using the StepOnePlus Real-Time PCR System (Applied Biosystems, Carlsbad, CA, United States). The thermal cycling conditions were as previously described (Yang et al., 2014). The fold change in gene expression was calculated as previously described (Pfaffl, 2001).

Statistical Analysis

Each bacterial count data point was derived from three independent experiments with two biological replicates ($n = 6$). Biofilm-forming assay data consisted of two independent experiments with six wells each. Fold change in gene expression data was derived from two independent experiments, each consisting of RNA extracted from three biological samples

---

[4] https://www.ncbi.nlm.nih.gov/tools/primer-blast/

which were polled together. Unpaired t-test, which can be accessed online at https://www.graphpad.com was used to compare bacterial count, biofilm-forming ability, and fold change in gene expression data set.

RESULTS AND DISCUSSION

WGS Analysis

Whole Genome Sequencing analysis revealed that ST 688C and ST 889B had identical MLST (Larsen et al., 2012) and cgMLST (Alikhan et al., 2018) of 26 and 5529, respectively. They also carried the identical chromosomal mutations *gyrA* S83Y and *parC* T57S. The acquired resistance gene *aac(6′)-Iaa* was present with an ID match of 97.72%, and both genomes featured identical point mutations in ten bases. The exact sequence of the *aac(6′)-Iaa* gene in the form of ResFinder 4.0 (Zankari et al., 2012) output text file from both ST 688C and ST 889B is available as **Supplementary Figure S1**. The two genomes shared an average nucleotide identity of 99.98% and an identical GC content of 52.23%. A SNP tree was constructed by including all known *S.* Thompson genomes from the NCBI database to investigate their phylogeny (**Figure 1**). ST 688C was genetically similar to the RM1894 outbreak strain and ST 889B. The SNP differences between the genomes were as follows: ST 688C and RM1894 (6 SNPs), ST 688C and ST 889B (9 SNPs), and ST 889B and RM1894 (9 SNPs). These three isolates formed a distinct clade (**Figure 1**; colored) whose most genetically similar clade consists of *S.* Thompson strains isolated exclusively from the United States from 2009 to 2012 except for the *S.* Thompson ATCC BAA 1738. In short, ST 688C and ST 889B are genetically similar to each other and the original RM1894 outbreak strain.

Inoculation of *S.* Thompson Onto Pre-harvest Basil Leaves

Upon inoculation of ST 688C and ST 889B on the adaxial surface of pre-harvest basil leaves, it was observed that the bacterial inoculum might either dry up on the leaf surface without eliciting any visible response from the leaf, i.e., no macroscopic lesions (referred to as healthy-looking leaf samples hereafter) as shown in **Figure 2A** or it may result in the formation of macroscopic lesions as shown in **Figure 3A**. The lesions fully developed within 24-h post-inoculation and were restricted to the leaf's portion in contact with the inoculum. Once fully formed, no further enlargement of the lesions was observed throughout the experiment (6 days).

Survivability of *S.* Thompson on Healthy-Looking Leaves

On the healthy-looking leaves, fluorescence confocal microscopy revealed that despite many GFP-tagged bacteria (emitting green fluorescence signal) at 1-day post-inoculation (dpi), very few remained by 6 dpi (**Figure 2A**). ST 889B was found to have

survived significantly ($p < 0.05$) better at both 1 (6.68 ± 0.17 vs. 6.01 ± 0.51 log CFU/sample) and 6 dpi (2.83 ± 0.18 vs. 2.18 ± 0.29 log CFU/sample) than ST 688C (**Figure 2B**).

As basil is known to produce antimicrobial substances such as linalool (Bagamboula et al., 2004; Beier et al., 2014), we first investigated if the two *S.* Thompson isolates had differential resistance to the potential antimicrobial basil-derived substances in freshly prepared sterile basil juice. The survivability of the two *S.* Thompson strains in sterile basil juice was determined over 3 days, and no significant differences in bacterial counts between the two strains were observed ($p > 0.05$, **Figure 2C**). On average, the counts of both strains reduced from ~5.6 log CFU/ml initially to ~4.2 log CFU/ml by 3 dpi.

Biofilm formation can increase *Salmonella*'s resistance toward stresses such as desiccation (Gibson et al., 2006; White et al., 2006, 2010) and, consequently, its ability to persist on abiotic surfaces and non-host environments (Gibson et al., 2006; White et al., 2006, 2010; Vestby et al., 2009). There is accumulating evidence showing that *Salmonella* can adhere to and from biofilms on plant surfaces, leading to persistent survival and colonization (Yaron and Römling, 2014; Maruzani et al., 2019). Accordingly, we compared the two *S.* Thompson isolates' biofilm-forming ability in both bacterial culture broth (1/10 tryptone soya broth, TSB; 1:9 TSB:DI water) and freshly prepared basil juice. Indeed, ST 889B showed significantly higher ($p < 0.01$) biofilm-forming ability than ST 688C in both 1/10 TSB and basil juice (**Figure 2D**). Accordingly, ST 889B also showed a higher mean fold change in gene expression level across the four biofilm formation genes tested in this study: *bapA* (1.95-fold), *bcsA* (2.49-fold), *csgB* (3.15-fold), and *wcaD* (1.86-fold) as compared to ST 688C after the inoculum have dried on pre-harvest basil leaves (**Figure 2E**).

These results suggest that in the scenario where *Salmonella* failed to elicit a visible reaction from the basil leaves upon inoculation, their survival strategy is more comparable to that of on an inert or abiotic surface, i.e., mainly characterized by biofilm formation to resist desiccation and starvation.

Survivability of *S.* Thompson in Leaves With Lesions

In this study, we observed the formation of distinctive macroscopic cell death lesions upon direct inoculation of *S.* Thompson on some basil leaves within 24 h (**Figure 3A**). In such samples, ST 688C survived slightly better than ST 889B at 1 dpi initially (6.99 ± 0.13 vs. 6.29 ± 0.15 log CFU/sample, $p < 0.05$) but this difference in survivability was no longer present at 6 dpi ($p > 0.05$, **Figure 3B**). More notably, for both *Salmonella* strains, the bacterial counts were drastically higher in the leaf samples with lesions than the healthy-looking leaf samples at 6 dpi (ST 688C: 4.39 ± 0.68 vs. 2.18 ± 0.29; ST 889B: 4.78 ± 0.12 vs. 2.83 ± 0.18 log CFU/sample, $p < 0.01$, **Figures 2B, 3B**). The frequency at which the lesions formed on pre-harvest basil leaves was significantly lower ($p < 0.05$) for the clinical-isolate ST 688C at 36.5% (35/96 tested leaves across 10 basil plants) as compared to

FIGURE 1 | Maximum likelihood SNP tree illustrating the phylogenetic relationship between ST 688C, ST 889B, the 1999 California coriander outbreak clinical strain RM1984 (three isolates of interest highlighted in color) and all other publicly available closely related members of *S.* Thompson isolates. Genomes downloaded from NCBI were denoted with their Biosample accession numbers followed by the place and year of isolation, if available. Numbers denote bootstrap values.

that of the basil-isolate ST 889B at 59.8% (70/117 tested leaves across 12 basil plants; **Figure 3C**). In other words, the basil-isolate ST 889B could induce the formation of these cell death lesions in basil leaves significantly better than the clinical counterpart ST 688C. Our fluorescence confocal microscopy of the lesions confirmed that although *Salmonella* grew luxuriantly, they were confined within the lesions while the healthy tissue was devoid of *Salmonella* cells (**Figure 3A**).

Plants, lacking the highly specific acquired immune systems of higher animals, feature a generalized multilayer immune system designed to resist invasion from a large variety of potential pathogens. The first layer of the plant immune system is the recognition of pathogen-associated molecular patterns (PAMPs), after which the plant mounts an immune response called the PAMP-triggered immunity (PTI; Monaghan and Zipfel, 2012; Garcia and Hirt, 2014). In response, pathogens well-adapted to the host plant can use bacterial delivery systems such as the type III secretion system (T3SS) to suppress the host PTI response (Garcia and Hirt, 2014). Upon failure of PTI, plants typically resort to the second layer of their immune system called the effector-triggered immunity (ETI). This typically leads to the hypersensitive response (HR), characterized by localized cell death lesions, designed to restrict the spread and growth of pathogens (Morel and Dangl, 1997; Heath, 2000a,b; Heidrich et al., 2012).

We next dedicated our efforts to shedding light on the possible underlying basis of ST 889B's ability to induce the formation of lesions in basil leaves more successfully. Since PTI precedes ETI, which leads to HR, a successful evasion or suppression of PTI must occur for the subsequent induction of HR. It has been reported that in *Salmonella*, flg22 epitome at the N-terminal region of bacterial flagellin is the most well-characterized PAMP (Gómez-Gómez and Boller, 2000; Zipfel et al., 2004) that leads to PTI in plants. Hence, we investigated if any downregulation of the flagellin gene *fliC* by ST 889B to mask its presence and evade PTI was at work. No significant difference was identified in the fold change in the expression level of the *fliC* gene between the two strains (**Figure 3D**) upon inoculation onto the basil leaves. Hence, the difference in *fliC* gene expression levels between the two strains was promptly eliminated as a possible basis for their differential ability to induce HR in basil leaves.

Subsequently, we explored the possible involvement of the T3SS in the formation of these cell death lesions. Although the role of T3SS and its effectors of a plant pathogen like *Pseudomonas syringae* is very well established (Jones and Dangl, 2006) that of the *Salmonella* T3SS in colonization, proliferation and eliciting the HR in plants is, unfortunately, less unambiguous. The *spaS−* and *sipB−* T3SS mutants showed increased colonization in the alfalfa roots and wheat seedlings (Iniguez et al., 2005) while the *prgH−* and *ssaV−* T3SS mutant
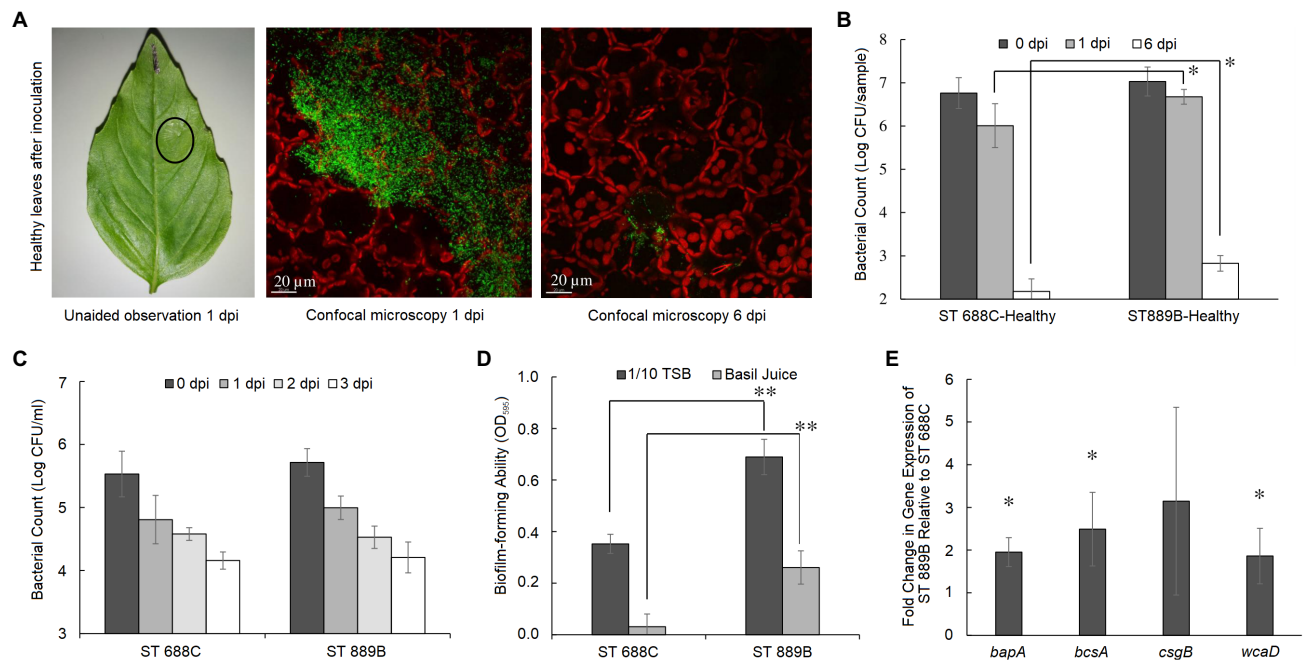
FIGURE 2 | (A) A representative healthy-looking basil leaf as seen unaided at 1 dpi, and under confocal microscopy at 1 and 6 dpi (from left to right) where red signals represent leaf cells while green signal represents GFP-tagged *Salmonella* cells; **(B)** bacterial counts of ST 688C and ST 889B on healthy-looking leaf tissue at 0, 1, and 6 dpi; and **(C)** in sterile basil juice at 0, 1, 2, and 3 dpi; **(D)** biofilm forming ability of ST 688C and ST 889B in 1/10 TSB and sterile basil juice; **(E)** fold change in gene expression of biofilm formation genes in ST 889B relative to ST 688C. Values linked by lines and denoted by * and ** represent significant difference at $p < 0.05$ and < 0.01, respectively, except in **(E)**, where * denotes significant difference ($p < 0.05$) to a theoretical mean of 1, which is the relative gene expression level of ST 688C. Bacterial counts in **(B,C)** were the average of three experiments, while that of biofilm forming ability in **(D)** and fold change in gene expression in **(E)** were the average of two experiments. Error bars denote standard deviation.

counts decreased in *Arabidopsis thaliana* (Schikora et al., 2011). Chalupowicz et al. (2018), on the other hand, observed no significant difference in survivability of wild type *S.* Typhimurium ATCC 14028 to those of various T3SS mutants (*invA⁻*, *ssaV⁻*, *sifA⁻*, and *sipB⁻*) in lettuce leaves.

More pertinently, the wild type *S.* Typhimurium ATCC 14028 also failed to elicit HR in beetroot tissue and pepper leaves when directly inoculated. However, its T3SS effectors SipB and SifA successfully elicited HR when they were delivered instead by plant pathogens *Erwinia amylovora* and *Xanthomonas euvesicatoria*, suggesting that *Salmonella* T3SS effectors, if successfully translocated into plant cells, can cause HR, but the *Salmonella* T3SS machinery is incapable of doing so (Chalupowicz et al., 2018). In contrast, we observed the development of HR cell death lesions within 24 h by direct inoculation of *S.* Thompson onto the basil leaves. This may be due to the high initial inoculum (10^8 CFU per leaf sample) used in this study. The T3SS has been shown to be an important factor for virulence at high ($\sim 10^8$ CFU per plant) but not low titers of *Salmonella* in tomato plants (De Moraes et al., 2017) and *Arabidopsis* (Schikora et al., 2011), although the underlying reason was unclear. High bacterial inoculum may have allowed *Salmonella* to secrete T3SS effectors at levels sufficient to suppress PTI and elicit HR from the basil leaf without translocating into the plant cells directly. Alternatively, the exact outcome of the interaction

may also be contingent upon the specific plant species, cultivar., and the *Salmonella* serovar in question (Klerks et al., 2007; Barak et al., 2011). Further research involving multiple plant pathosystems with multiple *Salmonella* serovars need to be conducted to shed light on this matter.

Despite prevailing mixed opinions and results, particularly on the *Salmonella* T3SS and the plant HR as outlined above, generally, the role of the T3SS in triggering HR is well established in many plant pathosystems (Jones and Dangl, 2006). Hence, in this study, we compared the gene expression level of *Salmonella* T3SS gene *prgH*, which has been employed as the reporter indicative of the overall T3SS activity in *Salmonella* (Clark et al., 2009; Wang et al., 2017) in the two *S.* Thompson strains immediately upon inoculation to the basil leaves. We observed a significantly higher ($p < 0.05$) 4.33-fold change in the mean expression level of *prgH* in ST 889B than ST 688C (**Figure 3D**). First, these results allude to possible involvement of the T3SS in suppressing the PTI and subsequently eliciting HR, specifically in the *S.* Thompson and basil plant pathosystem. Second, a higher T3SS activity may be responsible for the enhanced ability of ST 889B over ST 688C to cause HR in basil leaves.

The observed differences in the gene expression levels despite remarkably similar genetic makeup, which ultimately led to the differences in phenotype and survivability on pre-harvest basil leaves, might in part be attributed to the
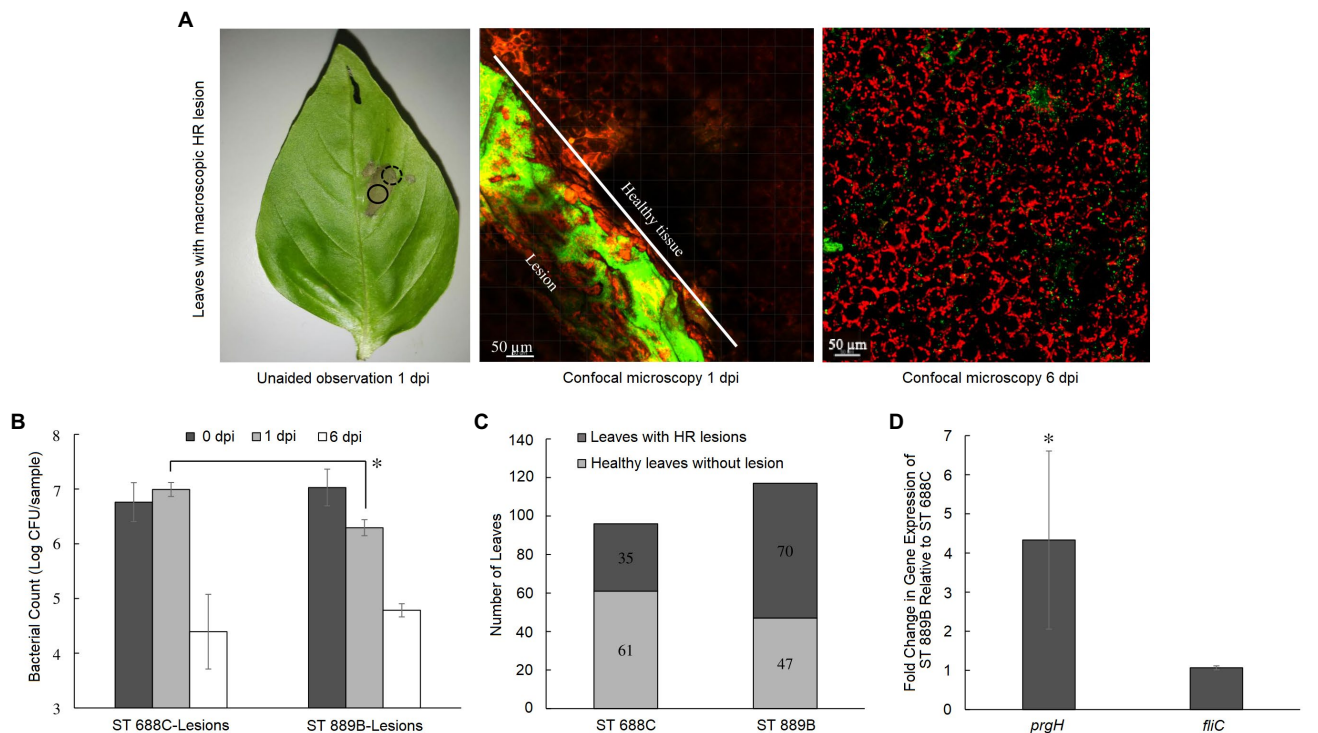
FIGURE 3 | (A) A representative basil leaf with macroscopic HR lesion as seen unaided at 1 dpi, and under confocal microscopy at 1 (dotted circle showing the border between healthy tissue and the lesion separated by the white line) and 6 dpi (full circle showing the lesion; from left to right) where red signals represent leaf cells while green signal represents GFP-tagged *Salmonella* cells; **(B)** bacterial counts of ST 688C and ST 889B on leaf tissue with HR lesion at 0, 1, and 6 dpi; **(C)** number of resultant healthy-looking leaves without HR and unhealthy leaves with HR following inoculation with ST 688C or ST 889B; **(D)** fold change in gene expression of select HR-related genes in ST 889B relative to ST 688C. Values linked by lines and denoted by * represent significant difference at $p < 0.05$ except in **(D)**, where * denotes significant difference ($p < 0.05$) to a theoretical mean of 1, which is the relative gene expression level of ST 688C. Bacterial counts were the average of three experiments in **(B)**, while that of the fold change in gene expression in **(D)** was of two experiments. Data on the number of leaves with and without HR lesions upon inoculation with ST 688C and ST 889B **(C)** were derived across 10 and 12 different pre-harvest basil plants, respectively. Error bars denote standard deviation.

inherently different natures of the two *S*. Thompson strains. It is known that pathogenic bacteria such as *Salmonella* spp. possesses two distinct "lifestyles," one tuned for virulence and one for environmental persistence, with each its own set of gene expression profiles as comprehensively reviewed elsewhere (Desai and Kenney, 2019). ST688C, being isolated from a human host with clinical salmonellosis, might have its gene expression profile tuned to that of human host virulence, which may partially explain its lower biofilm formation ability. Conversely, ST889B, being isolated from basil, might have its expression profile tuned to that of environmental persistence. However, whether allowing the clinical ST688C strain to be grown and adapted in basil over an extended time will improve its survivability to match that of ST889B is currently unknown and is a subject of further study.

CONCLUSION

In this study, we demonstrated that *S*. Thompson was capable of causing cell death lesions congruent with HR elicited from the basil plant. Furthermore, given the remarkable genetic similarity between ST 889B and ST 688C, the observed differences in phenotypes such as biofilm-forming ability, virulence (lesion formation rate), and survivability on pre-harvest basil could be attributed to the variations in expression levels of the same set of genes rather than to the differences in their genetic makeup. We also highlighted the two distinct *Salmonella* survival strategies in the basil pathosystem (biofilm formation on the healthy-looking leaf surface and elicitation of cell death lesions with subsequent invasion). Gene expression polymorphism may be an important strategy in the adaptation by human pathogens to survive and persist in plants, as demonstrated in this *S*. Thompson-basil model.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found at: https://www.ncbi.nlm.nih.gov/search/all/?term=PRJNA697834.

AUTHOR CONTRIBUTIONS

FUNDING

SUPPLEMENTARY MATERIAL

REFERENCES

Alikhan, N.-F., Zhou, Z., Sergeant, M. J., and Achtman, M. (2018). A genomic overview of the population structure of *Salmonella. PLoS Genet.* 14:e1007261. doi: 10.1371/journal.pgen.1007261

Bagamboula, C., Uyttendaele, M., and Debevere, J. (2004). Inhibitory effect of thyme and basil essential oils, carvacrol, thymol, estragol, linalool and p-cymene towards *Shigella sonnei* and *S. flexneri. Food Microbiol.* 21, 33–42. doi: 10.1016/S0740-0020(03)00046-7

Barak, J. D., Jahn, C. E., Gibson, D. L., and Charkowski, A. O. (2007). The role of cellulose and O-antigen capsule in the colonization of plants by *Salmonella enterica. Mol. Plant-Microbe Interact.* 20, 1083–1091. doi: 10.1094/MPMI-20-9-1083

Barak, J. D., Kramer, L. C., and Hao, L.-Y. (2011). Colonization of tomato plants by *Salmonella enterica* is cultivar dependent, and type 1 trichomes are preferred colonization sites. *Appl. Environ. Microbiol.* 77, 498–504. doi: 10.1128/AEM.01661-10

Beier, R. C., Byrd Ii, J. A., Kubena, L. F., Hume, M. E., Mcreynolds, J. L., Anderson, R. C., et al. (2014). Evaluation of linalool, a natural antimicrobial and insecticidal essential oil from basil: effects on poultry. *Poult. Sci.* 93, 267–272. doi: 10.3382/ps.2013-03254

Chalupowicz, L., Nissan, G., Brandl, M. T., Mcclelland, M., Sessa, G., Popov, G., et al. (2018). Assessing the ability of *Salmonella enterica* to translocate type III effectors into plant cells. *Mol. Plant-Microbe Interact.* 31, 233–239. doi: 10.1094/MPMI-07-17-0166-R

Clark, L., Martinez-Argudo, I., Humphrey, T. J., and Jepson, M. A. (2009). GFP plasmid-induced defects in *Salmonella* invasion depend on plasmid architecture, not protein expression. *Microbiology* 155, 461–467. doi: 10.1099/mic.0.025700-0

De Moraes, M. H., Desai, P., Porwollik, S., Canals, R., Perez, D. R., Chu, W., et al. (2017). *Salmonella* persistence in tomatoes requires a distinct set of metabolic functions identified by transposon insertion sequencing. *Appl. Environ. Microbiol.* 83:e03028-16. doi: 10.1128/AEM.03028-16

Delbeke, S., Ceuppens, S., Jacxsens, L., and Uyttendaele, M. (2015). Microbiological analysis of pre-packed sweet basil (*Ocimum basilicum*) and coriander (*Coriandrum sativum*) leaves for the presence of *Salmonella* spp. and Shiga toxin-producing *E. coli. Int. J. Food Microbiol.* 208, 11–18. doi: 10.1016/j.ijfoodmicro.2015.05.009

Desai, S. K., and Kenney, L. J. (2019). Switching lifestyles is an in vivo adaptive strategy of bacterial pathogens. *Front. Cell. Infect. Microbiol.* 9:421. doi: 10.3389/fcimb.2019.00421

Garcia, A. V., and Hirt, H. (2014). *Salmonella enterica* induces and subverts the plant immune system. *Front. Microbiol.* 5:141. doi: 10.3389/fmicb.2014.00141

Gibson, D., White, A., Snyder, S., Martin, S., Heiss, C., Azadi, P., et al. (2006). *Salmonella* produces an O-antigen capsule regulated by AgfD and important for environmental persistence. *J. Bacteriol.* 188, 7722–7730. doi: 10.1128/JB.00809-06

Gómez-Gómez, L., and Boller, T. (2000). FLS2: an LRR receptor–like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis. Mol. Cell* 5, 1003–1011. doi: 10.1016/S1097-2765(00)80265-8

Heath, M. C. (2000a). "Hypersensitive response-related death," in *Programmed Cell Death in Higher Plants.* eds. E. Lam, H. Fukuda and J. Greenberg (Springer: Netherlands), 77–90.

Heath, M. C. (2000b). Nonhost resistance and nonspecific plant defenses. *Curr. Opin. Plant Biol.* 3, 315–319. doi: 10.1016/s1369-5266(00)00087-x

Heidrich, K., Blanvillain-Baufumé, S., and Parker, J. E. (2012). Molecular and spatial constraints on NB-LRR receptor signaling. *Curr. Opin. Plant Biol.* 15, 385–391. doi: 10.1016/j.pbi.2012.03.015

Iniguez, A. L., Dong, Y., Carter, H. D., Ahmer, B. M., Stone, J. M., and Triplett, E. W. (2005). Regulation of enteric endophytic bacterial colonization by plant defenses. *Mol. Plant-Microbe Interact.* 18, 169–178. doi: 10.1094/MPMI-18-0169

Jones, J. D., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286

Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M., and Lund, O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One* 9:e104984. doi: 10.1371/journal.pone.0104984

Klerks, M., Van Gent-Pelzer, M., Franz, E., Zijlstra, C., and Van Bruggen, A. (2007). Physiological and molecular responses of *Lactuca sativa* to colonization by *Salmonella enterica* serovar Dublin. *Appl. Environ. Microbiol.* 73, 4905–4914. doi: 10.1128/AEM.02522-06

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., et al. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 50, 1355–1361. doi: 10.1128/JCM.06094-11

Lee, J. H., Cho, M. H., and Lee, J. (2011). 3-Indolylacetonitrile decreases Escherichia coli O157: H7 biofilm formation and Pseudomonas aeruginosa virulence. *Environmental microbiology* 13:62–73.

Li, D., and Uyttendaele, M. (2018). Potential of human norovirus surrogates and *Salmonella enterica* contamination of pre-harvest basil (*Ocimum basilicum*) via leaf surface and plant substrate. *Front. Microbiol.* 9:1728. doi: 10.3389/fmicb.2018.01728

Ma, L., Zhang, G., and Doyle, M. P. (2011). Green fluorescent protein labeling of *Listeria, Salmonella,* and *Escherichia coli* O157: H7 for safety-related studies. *PLoS One* 6:e18083. doi: 10.1371/journal.pone.0018083

Maruzani, R., Sutton, G., Nocerino, P., and Marvasi, M. (2019). Exopolymeric substances (EPS) from *Salmonella enterica:* polymers, proteins and their interactions with plants and abiotic surfaces. *J. Microbiol.* 57, 1–8. doi: 10.1007/s12275-019-8353-y

Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Curr. Opin. Plant Biol.* 15, 349–357. doi: 10.1016/j.pbi.2012.05.006

Morel, J.-B., and Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death Differ.* 4, 671–683. doi: 10.1038/sj.cdd.4400309

Pakalniskiene, J., Falkenhorst, G., Lisby, M., Madsen, S., Olsen, K., Nielsen, E., et al. (2009). A foodborne outbreak of enterotoxigenic *E. coli* and *Salmonella* Anatum infection after a high-school dinner in Denmark, November 2006. *Epidemiol. Infect.* 137, 396–401. doi: 10.1017/S0950268808000484

Pezzoli, L., Elson, R., Little, C. L., Yip, H., Fisher, I., Yishai, R., et al. (2008). Packed with *Salmonella*—investigation of an international outbreak of *Salmonella* Senftenberg infection linked to contamination of prepacked basil in 2007. *Foodborne Pathog. Dis.* 5, 661–668. doi: 10.1089/fpd.2008.0103

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res.* 29:e45. doi: 10.1093/nar/29.9.e45

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., et al. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* 17, 7–15. doi: 10.3201/eid1701.P11101

Schikora, A., Virlogeux-Payant, I., Bueso, E., Garcia, A. V., Nilau, T., Charrier, A., et al. (2011). Conservation of *Salmonella* infection mechanisms in plants and animals. *PLoS One* 6:e24112. doi: 10.1371/journal.pone.0024112

Vestby, L. K., Møretrø, T., Langsrud, S., Heir, E., and Nesse, L. L. (2009). Biofilm forming abilities of *Salmonella* are correlated with persistence in fish meal-and feed factories. *BMC Vet. Res.* 5:20. doi: 10.1186/1746-6148-5-20

Wang, L., Cai, X., Wu, S., Bomjan, R., Nakayasu, E. S., Händler, K., et al. (2017). InvS coordinates expression of PrgH and FimZ and is required for invasion of epithelial cells by *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* 199:e00824-16. doi: 10.1128/JB.00824-16

White, A., Gibson, D., Kim, W., Kay, W., and Surette, M. (2006). Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. *J. Bacteriol.* 188, 3219–3227. doi: 10.1128/JB.188.9.3219-3227.2006

White, A. P., Weljie, A. M., Apel, D., Zhang, P., Shaykhutdinov, R., Vogel, H. J., et al. (2010). A global metabolic shift is linked to *Salmonella* multicellular development. *PLoS One* 5:e11814. doi: 10.1371/journal.pone.0011814

Yang, Y., Khoo, W. J., Zheng, Q., Chung, H.-J., and Yuk, H.-G. (2014). Growth temperature alters *Salmonella* Enteritidis heat/acid resistance, membrane lipid composition and stress/virulence related gene expression. *Int. J. Food Microbiol.* 172, 102–109. doi: 10.1016/j.ijfoodmicro.2013.12.006

Yaron, S., and Römling, U. (2014). Biofilm formation by enteric pathogens and its role in plant colonization and persistence. *Microb. Biotechnol.* 7, 496–516. doi: 10.1111/1751-7915.12186

Yoon, S.-H., Ha, S.-M., Lim, J., Kwon, S., and Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110, 1281–1286. doi: 10.1007/s10482-017-0844-4

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261

Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E. J., Jones, J. D., Felix, G., et al. (2004). Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature* 428, 764–767. doi: 10.1038/nature02485

Zwe, Y. H., and Yuk, H. G. (2017). Food quality and safety in Singapore: microbiology aspects. *Food Qual. Saf.* 1, 101–105. doi: 10.1093/fqsafe/fyx016

frontiers
in Microbiology

A Genomic Perspective Across Earth's Microbiomes Reveals That Genome Size in Archaea and Bacteria Is Linked to Ecosystem Type and Trophic Strategy

Alejandro Rodríguez-Gijón[1][†], Julia K. Nuy[1][†], Maliheh Mehrshad[2], Moritz Buck[2], Frederik Schulz[3], Tanja Woyke[3] and Sarahi L. Garcia[1]*

[1] Department of Ecology, Environment, and Plant Sciences, Science for Life Laboratory, Stockholm University, Stockholm, Sweden, [2] Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden, [3] DOE Joint Genome Institute, Berkeley, CA, United States

Our view of genome size in Archaea and Bacteria has remained skewed as the data has been dominated by genomes of microorganisms that have been cultivated under laboratory settings. However, the continuous effort to catalog Earth's microbiomes, specifically propelled by recent extensive work on uncultivated microorganisms, provides an opportunity to revise our perspective on genome size distribution. We present a meta-analysis that includes 26,101 representative genomes from 3 published genomic databases; metagenomic assembled genomes (MAGs) from GEMs and stratfreshDB, and isolates from GTDB. Aquatic and host-associated microbial genomes present on average the smallest estimated genome sizes (3.1 and 3.0 Mbp, respectively). These are followed by terrestrial microbial genomes (average 3.7 Mbp), and genomes from isolated microorganisms (average 4.3 Mbp). On the one hand, aquatic and host-associated ecosystems present smaller genomes sizes in genera of phyla with genome sizes above 3 Mbp. On the other hand, estimated genome size in phyla with genomes under 3 Mbp showed no difference between ecosystems. Moreover, we observed that when using 95% average nucleotide identity (ANI) as an estimator for genetic units, only 3% of MAGs cluster together with genomes from isolated microorganisms. Although there are potential methodological limitations when assembling and binning MAGs, we found that in genome clusters containing both environmental MAGs and isolate genomes, MAGs were estimated only an average 3.7% smaller than isolate genomes. Even when assembly and binning methods introduce biases, estimated genome size of MAGs and isolates are very similar. Finally, to better understand the ecological drivers of genome size, we discuss on the known and the overlooked factors that influence genome size in different ecosystems, phylogenetic groups, and trophic strategies.

Keywords: microbial ecology, genome size, bacteria, archaea, genomics

INTRODUCTION

As microbiologists, how do we define what is a small or a big genome? Perhaps, researchers working on model organisms such as *Escherichia coli* with a genome size of ~5 Mbp (Abram et al., 2021) would define "big" or "small" differently to researchers working on soil-dwelling bacteria with a genome size of 16 Mbp (Garcia et al., 2014). On the lower genome size scale, whereas genome sizes of bacterial endosymbionts of insects may have genomes merely larger than 100 kbp (Moran and Bennett, 2014), the abundant *Prochlorococcus* range between 1.6 and 1.9 Mbp for high-light and low-light ecotypes (Berube et al., 2018). In summary, it is known that genome sizes of Archaea and Bacteria range between 100 kbp and 16 Mbp, but the genome size distribution in nature is still undefined. Therefore, the aim of this review is to provide an overview of the distribution of genome sizes in different ecosystems.

We leveraged recently published databases of archaeal and bacterial metagenome assembled genomes (MAGs) (Nayfach et al., 2020; Buck et al., 2021a) together with isolate genomes to revisit and acquire an updated understanding of the estimated genome size distribution across different ecosystems. In this review, we also discuss the ecological drivers that potentially influence genome sizes. In summary, we found that 76.3% of representative archaeal and bacterial genomes recovered through genome-resolved metagenomics present estimated genome sizes below 4 Mbp. Furthermore, all MAGs from five archaeal phyla (Micrarchaeota, Ianarchaeota, Undinarchaeota, Nanohaloarchaeota, and Hadarchaeota) and two bacterial phyla (Coprothermobacterota and Dictyoglomota) were recovered exclusively from aquatic ecosystems and have genome sizes below 2 Mbp (**Figures 1A,B**).

APPROXIMATION OF GENETIC UNITS USING 95% AVERAGE NUCLEOTIDE IDENTITY AND ITS CAVEATS

Species are widely considered congruent genetic and ecological units for sexual eukaryotes (Mallet, 2008; Shapiro and Polz, 2015). However, there is no consensus regarding the concept of species for Archaea and Bacteria. Instead, 95% average nucleotide identity (ANI) has been a widely recognized as a genetic boundary to operationally estimate genetic units or "microbial species" (Konstantinidis and Tiedje, 2005; Varghese et al., 2015; Garcia et al., 2018; Jain et al., 2018). Several genomic and metagenomic studies have verified the existence of sequence discrete genetic units with 95% ANI as boundary (Olm et al., 2020; Rodriguez et al., 2021).

We used the 95% ANI boundary in published datasets (Nayfach et al., 2020; Parks et al., 2020; Buck et al., 2021a) to review and renew our view of the distribution of archaeal and bacterial estimated genome size. To minimize representation biases (Gweon et al., 2017), we included one representative per 95% ANI cluster. The approximation of these clusters will be called mOTUs (metagenomic Operational Taxonomic Units) (Buck et al., 2021b; Garcia et al., 2021). We used only one MAG per mOTU, making the estimated genome size dependent on the MAG assembly and quality. It is known that MAG assembly and binning might discriminate against ribosomal RNAs, transfer RNAs, mobile element functions and genes of unknown function (Nelson et al., 2020; Meziti et al., 2021), and also that completeness could be underestimated for streamlined genomes (Garcia et al., 2015). Despite potential biases introduced by these methodological limitations, we still offer valuable insights on genome size distribution across different environments.

EXTANT GENOME SIZE DISTRIBUTION IN THE ENVIRONMENT

In this review, we have included ~64,500 environmental MAGs available via two recently published datasets, stratfreshDB and GEMs. StratfreshDB offers ~12,000 MAGs (>40% completeness) from 41 stratified lakes and ponds assembled with Megahit (v1.1.13) and binned with Metabat (v2.12.1) (Buck et al., 2021a). GEMs offers ~52,000 MAGs (>50% completeness) from > 10,000 metagenomes collected from diverse habitats on Earth (Nayfach et al., 2020). GEMs dataset was assembled using metaSPAdes and binned with Metabat (v0.32.5). After dereplication using fastANI and mOTUlizer (Jain et al., 2018; Buck et al., 2021b), our meta-analysis includes 17 834 mOTUs, with one representative MAG each (completeness > 50% and contamination < 5%, assessed with CheckM v1.1.3) (Parks et al., 2015). We complemented these environmental MAGs by adding 8 267 representative genomes (>90% complete) of isolates from GTDB (Parks et al., 2020; **Figure 1A**). These genomes are marked in the GTDB databases (release 95) to originate from culture collections. After clustering at 95% ANI threshold, 540 mOTUs contained representatives from both environmental MAGs and isolate genomes (**Figure 1C**). Previous surveys based on 16S rRNA have found that the uncultured microbial fraction could constitute up to 81% of the total microbial diversity on Earth (Lloyd et al., 2018). However, it is known that 16S rRNA underestimate prokaryotic diversity (Rodriguez et al., 2018). Overall, our review shows that 3% of the reconstructed environmental mOTUs are represented among cultured microbes.

Furthermore, using completeness estimates from CheckM, we compared the estimated genome size distribution of all MAGs vs. genomes from isolates. The estimated genome size was calculated by dividing the MAG's assembly size by CheckM completeness (ranging from 0 to 1). Representative genomes from isolates have an average genome size of 4.3 Mbp which is significantly larger than that of MAGs (t-test $p < 0.0001$), both when comparing Archaea and Bacteria combined and separately. To compare estimated genome sizes between MAGs, ecosystem type was used according to the GEMs database. Although the ecosystem classification presented here is coarse and might contain countless niches, it still allowed us to see trends for genome sizes. Estimated genome sizes of aquatic MAGs have an average of 3.1 Mbp, host-associated MAGs average 3.0 Mbp, and terrestrial MAGs average 3.7 Mbp (**Figure 1A**). For the 540 mOTUs that contained both environmental MAGs and isolate
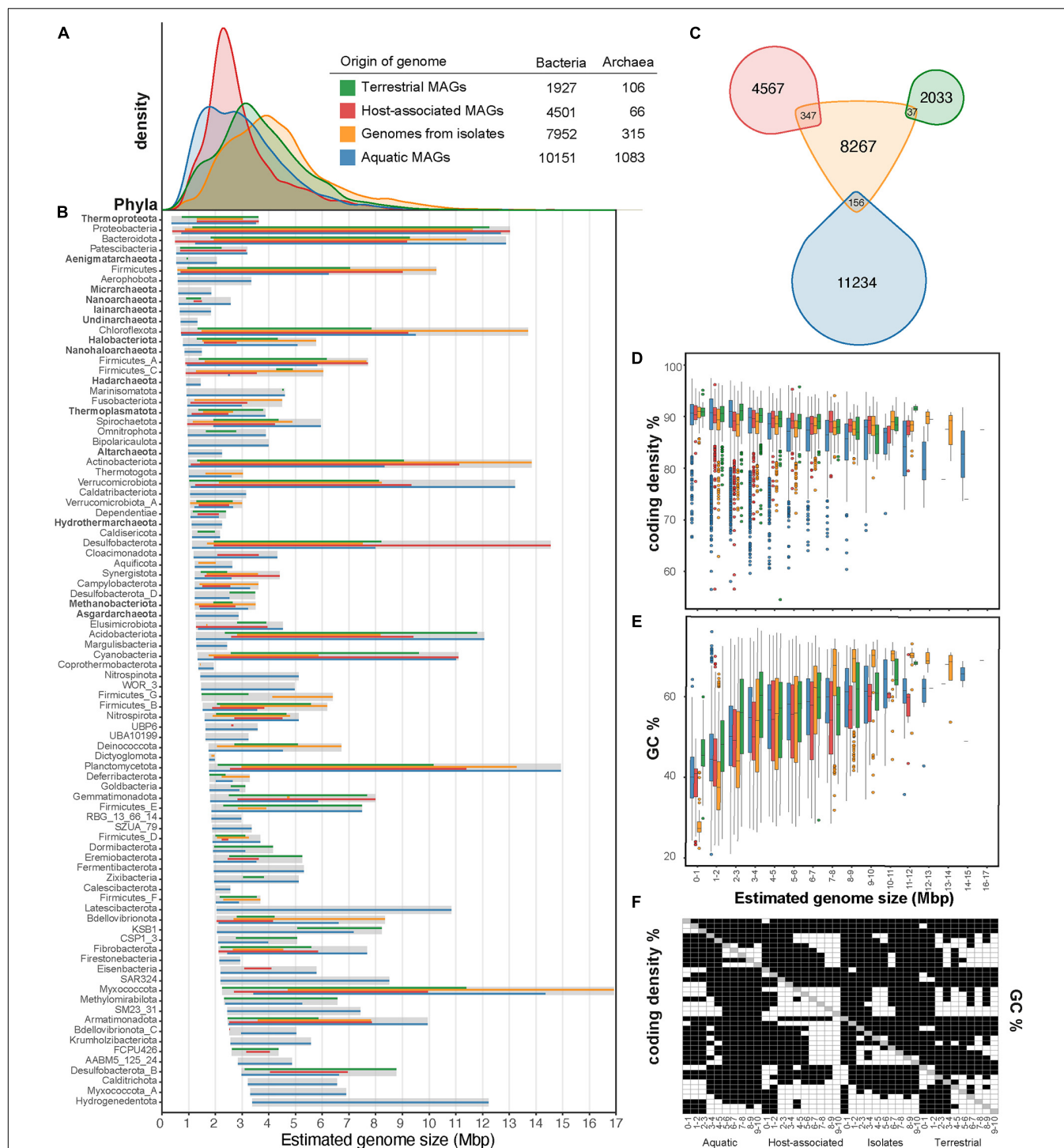
FIGURE 1 | Overview of the genome size distribution across Earth's microbiomes. Genome size distribution of Archaea and Bacteria (A) from different environmental sources and across different archaeal and bacterial phyla (B) are shown for a total of 26,101 representative genomes. Isolate genomes were gathered from GTDB (release95) and environmental MAGs were gathered from GEMs (Nayfach et al., 2020) and stratfreshDB (Buck et al., 2021a). We use one representative genome per mOTU (defined by 95% ANI) from the union of GEMs catalog and stratfreshDB in the plots. From the GTDB database, we selected one representative isolate genome per species cluster that was circumscribed based on the ANI (≥95%) and alignment fraction [(AF) > 65%] between genomes (Parks et al., 2020). To construct the figures, we plotted the min-max estimated genome sizes, which were calculated based on the genome assembly size and completeness estimation provided. Venn diagram of the intersection between the representative environmental MAGs and the representative isolate genomes (C). The intersection was calculated using FastANI (Jain et al., 2018) and was determined with a threshold of 95%. The coding density (D) and GC content (%) (E) are shown for the archaeal and bacterial MAGs across different ecosystem categories and isolates. Pair-wise t-test was performed in all variables of (D,E) and shown in (F), where white is significant ($p < 0.05$) and black is not significant ($p > 0.05$). In (B), we only included phyla with more than five genomes.

genomes (**Figure 1C**), we found that MAGs were estimated on average 3.7% smaller than isolate genomes (**Supplementary Figure 1**). In other words, even when assembly and binning methods introduce biases, estimated genome size of MAGs and isolates are very similar. Overall, this suggests that the bias in metagenome assembly and binning would not account for the genome size difference observed between all isolate representatives and ecosystem MAGs, neither for the differences among ecosystem MAGs.

A reason for the difference in genome size between isolates and genomes reconstructed from metagenomes might be related to the fact that traditional isolation techniques select for rare microorganisms (Shade et al., 2012) and do not capture the entire ecosystem's diversity (**Figure 1C**). For example, it is known that classical cultivation techniques with rich media bias the cultivation toward copiotrophic and fast-growing microorganisms (Swan et al., 2013). Cultivation biases our view of nature because it selects against slow growing microorganisms (Imachi et al., 2020), host dependency (Cross et al., 2019), and dormancy (Hoehler and Jorgensen, 2013) among others. In nature there are many microorganisms with very limited metabolic capacity (Figueroa-Gonzalez et al., 2020) that is linked with dependencies and smaller genomes sizes (Morris et al., 2012). Microorganisms in nature have coevolved with other microorganisms and might have specific requirements that are difficult to mimic in batch-culture standard-media isolation techniques (Garcia, 2016). Although there have been many advances on cultivation techniques (Dedysh, 2011; Carini et al., 2013; Henson et al., 2016; Imachi et al., 2020), more innovations to culture the uncultivated microbial majority (Lewis et al., 2020) will enable us to bring more natural abundant representatives to culture.

Placing archaeal and bacterial genome sizes in phylogenetic trees using GTDB-tk (**Figures 2A,B**) shows that the distribution of representative genomes and their estimated sizes varies widely between different phyla and within phyla. MAGs assigned to eight phyla in the domain Archaea were reconstructed exclusively from aquatic ecosystems, whereas eight other archaea phyla were reconstructed from multiple ecosystems. There was no significant difference between the genome sizes of aquatic archaea phyla or those from non-specific ecosystem (**Figure 2C**). However, estimated genome sizes in bacterial phyla were significantly larger than those in archaeal phyla. Moreover, genera from phyla with genome sizes below 3 Mbp, such as Halobacteriota, Thermoproteota, and Patescibacteria, do not show genome size variation in different ecosystems (**Figures 2D,E,I**). Nevertheless, genera from these smaller genome sizes phyla are significantly smaller than genera with more genome size variation in any ecosystem category (**Figures 2K–N**). For phyla with genome sizes above 3 Mbp, the genome sizes in aquatic or host-associated genera are significantly smaller than those in terrestrial or non-specific ecosystems (**Figures 2F–J**). We observe that while the microorganisms' ecosystem can certainly be linked to genome size, phyla where genome sizes are mostly below 3 Mbp show no variation in estimated genome size across ecosystems.

Clustering microorganisms together by the three ecosystem categories is not optimal since each contains innumerable niches.

In each niche, there will be different selective pressures on the genome size. An example is clearly shown in a study (Nayfach and Pollard, 2015) in which it is observed that Archaea and Bacteria sampled from different parts of the human body have differences in genome sizes. Low metadata resolution and clustering of all genomes into three main ecosystem types might be a reason why we see a range of genome sizes in the genera of different ecosystems (**Figure 2**). With more precise metadata and higher sampling resolution of microhabitats, it might be possible to identify the ecological drivers of genome sizes in the different niches in fine-scale.

IMPACT OF ECOSYSTEM AND TROPHIC STRATEGY ON GENOME SIZE

Terrestrial ecosystems harbor immense microbial diversity (Delgado-Baquerizo et al., 2018). Yet, the most up-to-date data compilation provided here shows only 2033 MAGs from terrestrial ecosystems (**Figure 1C**) with an average genome size of 3.7 Mb (**Figure 1A**). The sub-ecosystems considered in this view are soil and deep subsurface, among others (**Supplementary Figure 2**). While the terrestrial microorganism's genome size is the biggest of the three ecosystem categories in this review, they are smaller than expected based on previous metagenomic predictions, which placed the genome size of soil bacteria at 4.74 Mb (Raes et al., 2007). Trends for larger genome sizes in soil have been hypothesized to be related to scarcity and high diversity of nutrients, and a fluctuating environment combined with little penalty for the slow growth rate (Konstantinidis and Tiedje, 2004; Cobo-Simon and Tamames, 2017; Chen et al., 2021). Although terrestrial environments are physically structured, they are generally characterized by two to three orders of magnitude greater variations (in temperature and currents) than marine environments (Steele et al., 2019). *In silico* studies predict that large genome sizes could result from higher environmental variability (Bentkowski et al., 2015). A recent example showed that isolates of terrestrial Cyanobacteria have genomes on the larger size scale (6.0–8.0. Mb) that are enriched in genes involved in regulatory, transport and motility functions (Chen et al., 2021). These functional categories enable them to thrive in a fluctuating environment with high nutrient diversity. Despite these general trends showing larger genome sizes in terrestrial ecosystems, it is worth noting that the diversity captured in the GEMs survey is probably a small fraction of the total terrestrial microbial diversity. It is, for example, also known that streamlined microorganisms such as Patescibacteria (Ortiz et al., 2021) and "*Candidatus* Udaeobacter copiosus" (Verrucomicrobiota) are abundant in soils (Brewer et al., 2016). Therefore, we predict that the view on genome size distribution and microbial diversity in terrestrial ecosystems will become more complete with more sequencing, assembly, binning and novel isolation efforts.

In host-associated microbiomes, genetic drift, deletion biases and low populations sizes drive the reduction of genomes (Li et al., 2021). In these environments, the differing levels of intimacy with their host can influence the evolution of the genome of microorganisms. For example, within the
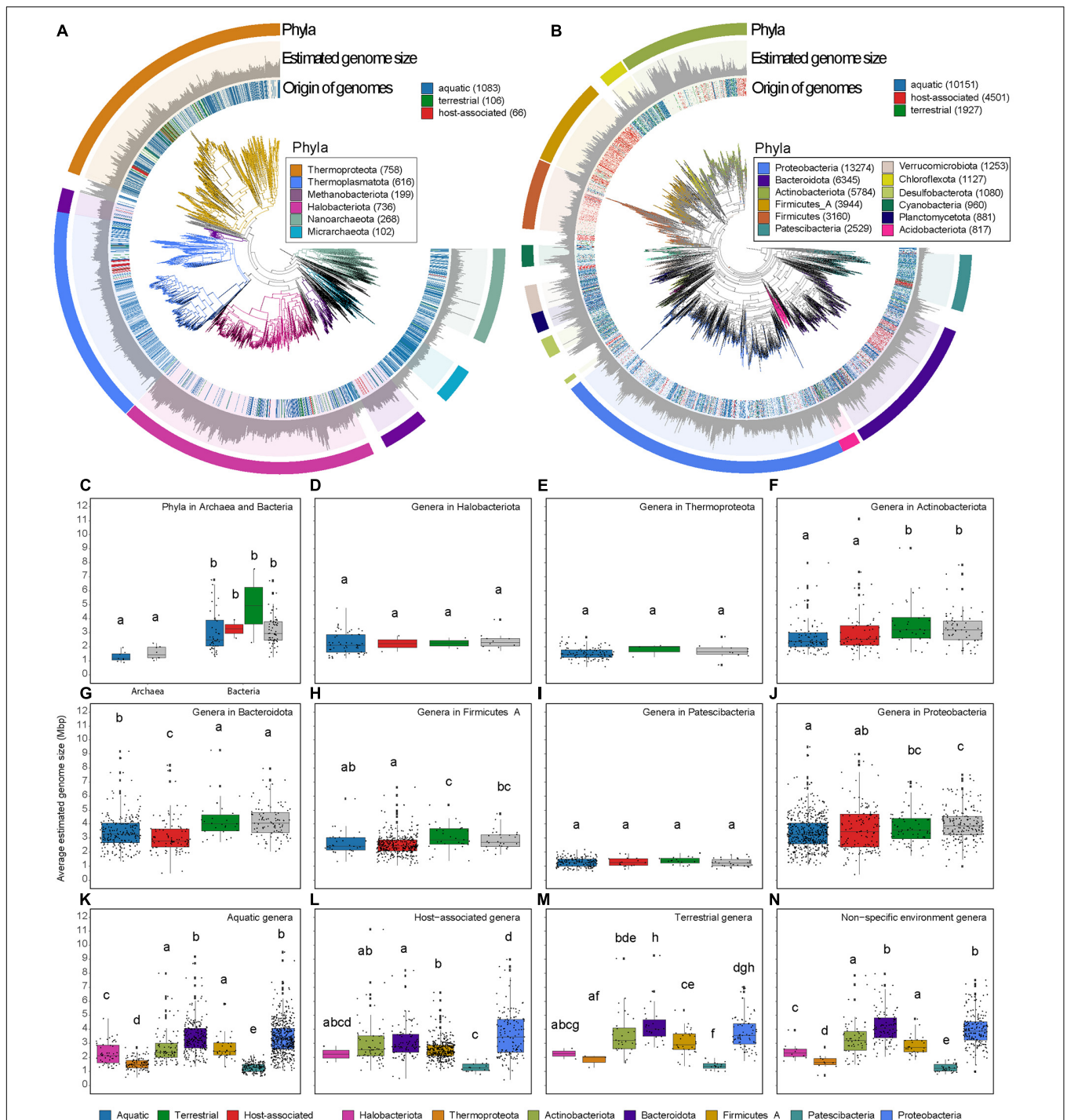
FIGURE 2 | Phylogenetic trees of archaeal **(A)** and bacterial **(B)** representative genomes show variation in genome size between and within phyla. The trees were constructed using GTDB-tk (v 1.5.0) using *de novo* workflow using aligned concatenated set of 122 and 120 single copy marker proteins for Archaea and Bacteria, respectively (Chaumeil et al., 2020). Moreover, in this mode, GTDB-tk adds 1,672 and 30,238 backbone genomes for Archaea and Bacteria, respectively. Tree is visualized in anvi'o (Eren et al., 2021). Estimated genome size is presented in scale from 0 to 6 Mpb or 0 to14 Mbp for Archaea or Bacteria. In the tree, the origin of the environmental genomes is labeled: aquatic, terrestrial and host-associated (same MAGs as **Figure 1**). Highlighted phyla with more representative genomes are color-coded. Boxplots show the average estimated genome size per phyla within archaeal and bacterial **(C)** domains. The average estimated size per genus within Halobacteriota **(D)**, Thermoproteota **(E)**, Actinobacteriota **(F)**, Bacteroidota **(G)**, Firmicutes A **(H)**, Patescibacteria **(I)**, Proteobacteria **(J)**. The presence of phyla and genera is colored in gray if they contain MAGs from different ecosystem category (non-specific ecosystem). The average estimated size per genus extracted from aquatic ecosystems **(K)**, host-associated ecosystems **(L)**, terrestrial ecosystems **(M)**, or non-specific ecosystems **(N)**. Letters in boxplot panels are the result of non-parametric tests, Wilcoxon **(C)** or Kruskal-Wallis **(D–N)**. Different letters show significant differences $p < 0.05$ (all statistical test with multiple testing were corrected with Benjamini-Hochberg).

Chlamydiaceae family, some lineages have evolved intracellular associations with eukaryotes (Toft and Andersson, 2010; Collingro et al., 2020). These intracellular Chlamydiaceae have lost many genes that were likely present in their common ancestor that lived in the environment (Dharamshi et al., 2020). Moreover, host-associated bacterial genomes show a variation in size depending on the type of host (plant, animal, etc.) and the type of association they have with the host, such as endosymbiotic, ectobiotic, or epibiotic (**Supplementary Table 1**). Generally, microorganisms associated with Arthropoda (Tamas et al., 2002), humans (McLean et al., 2020) and other mammals show smaller genomes sizes, whereas protist- and plant-associated bacteria present larger genomes (Levy et al., 2017; **Supplementary Figure 2**). In fact, *in silico* studies of Alphaproteobacteria show massive genome expansions diversifying plant-associated Rhizobiales and extreme gene losses in the ancestor of the intracellular lineages Rickettsia, Wolbachia, Bartonella, and Brucella that are animal- and human-associated (Boussau et al., 2004). Although host-associated microorganisms are widely known for their reduced genomes, the characteristics of host-associated MAGs show coding densities of ∼91% for genomes below 2 Mbp (**Figure 1D**).

Small genomes exhibit either strong dependency on other community members or have specific nutrient requirements. Two diverging views on genome reduction have emerged to explain mechanisms of gene loss. On the one hand, genetic drift is more pronounced in species that have a small effective population size, such as host-associated endosymbiotic microorganisms. These microorganisms might thrive because hosts provide energy or nutrients. On the other hand, streamlining is the process of gene loss through selection and it is mainly observed in free-living microorganisms with high effective population sizes (Giovannoni et al., 2014). Some of the most numerically abundant and streamlined microorganisms known to date, such as Pelagibacter (class Alphaproteobacteria) (Giovannoni et al., 2014), *Prochlorococcus* (phylum Cyanobacteria) (Rocap et al., 2003) Thermoproteota (Aylward and Santoro, 2020) and Patescibacteria (Tian et al., 2020), are commonly found in aquatic niches. Paradoxically, even though these microorganisms are free-living, their small genomes increase their nutritional connectivity to other individuals (Giovannoni et al., 2014). Free-living aquatic microorganisms have been used as exemplary streamlining cases in which many have gone through community adaptive selections and gene loss (Morris et al., 2012). Genome reduction can be so intense that microorganisms lose the capacity to biosynthesize essential metabolites and, thus, become auxotrophs. To overcome required nutritional needs, microorganisms thrive in functional cohorts (Mondav et al., 2020). As opposed to prototrophic lifestyle, the auxotrophic lifestyle is reflected by smaller genome sizes (Grote et al., 2012; Garcia et al., 2015; Brewer et al., 2016; Kang et al., 2017; **Supplementary Table 1**). An opportunity for future studies includes research on auxotrophy prevalences across the entire spectrum of metabolites (amino acids, nucleotides, fatty acids, vitamins, etc.) in different microbial communities and how those auxotrophies are linked with genome size.

In this review, the largest fraction of MAGs is recovered from aquatic environments. The two main sub-ecosystems in our survey are freshwater with MAGs estimated average genome size of 3.2 Mbp significantly different ($p < 0.0001$) from marine genome size distribution with average estimated genome size of 2.9 Mbp. When comparing freshwater and marine environments, the most obvious difference is salinity followed by nutrient concentration. Further exploring the impact of differing levels of salinity on genome size is an interesting research prospect. Additionally, we compared the union of representative freshwater MAGs from both databases (StratfreshDB and GEMs) (**Supplementary Figure 3**). The difference of mean estimated genome size between the representatives from freshwater GEMs and StratfreshDB is 0.52 Mbp. However, this is because each database captures genetic units that were not found in the other database.

In general, aquatic environments are vertically structured by gradients of light penetration, temperature, oxygen, and nutrient (**Supplementary Table 1**). Moreover, microorganisms might experience a microscale spatial and nutrient structure due to the presence of heterogeneous particles. These aquatic structures are drivers of the genetic repertoire of aquatic microorganisms. Metagenomic sequencing reported the increase of genome sizes for Archaea and Bacteria with increasing depths (Mende et al., 2017). Temperature may be as important; for example, a study based on twenty-one Thermoproteota and Euryarchaeota fosmids (Euryarchaetoa currently reclassified into Methanobacteriota, Halobacteriota, and Nanohaloarchaeota) showed high rates of gene gains through HGT to adapt to cold and deep marine environments (Brochier-Armanet et al., 2011). It has been observed that light is a relevant driver of genome size in aquatic environments as it decreases with depth. Photosynthetic bacteria such as *Prochlorococcus* spp. are well-differentiated into a high-light adapted ecotype with smaller genome sizes (average 1.6 Mbp) and a low-light-adapted ecotype with a slightly bigger genome size (average 1.9 Mbp) (Berube et al., 2018; **Supplementary Table 1**). Limitation of nutrients such as nitrogen (Elser et al., 2007) might also be one of the central factors determining genomic properties (Grzymski and Dussaq, 2012). Nitrogen fixation is a complex process that requires a great amount of genes (Franche et al., 2008) and most nitrogen-fixing marine cyanobacteria have the largest genomes in its phylum (Bergman et al., 2013).

Diversity and quantity of nutrients might be two understudied factors that drive ecology and genome size evolution. A recent example shows that polysaccharide xylan triggers microcolonies, whereas monosaccharide xylose promotes solitary growth in Caulobacter (D'Souza et al., 2021). This is a striking example of how nutrient complexity can foster diverse niches for well-studied cells such as Caulobacter with genome size 4 Mbp. To fully understand the link between genome size and nutritional requirements of diverse environmental microorganisms, we need to systematically explore the ∼90% of molecules/metabolites still unknown (Wienhausen et al., 2017; Hawkes et al., 2018; Patriarca et al., 2020). The wide nutrient complexity in the environment might prompt microorganisms to shape their genome. Their genomic content and metabolic potential defines whether they

are capable to feed on the available nutrients, forcing them to develop dependencies with other community members in order to acquire energy and metabolic precursors. Metagenomics combined with metabolomics will provide an understanding of the link between genome size evolution of microorganisms and their nutritional and trophic strategy.

CONCLUSION

This review offers a broad overview of genome size distribution across three different ecosystem categories, showing that MAGs recovered from aquatic and host-associated ecosystems present smaller estimated genome sizes than those recovered from terrestrial ecosystems. Moreover, genomes obtained from environmental samples present a smaller estimated genome size than obtained by cultivation approaches. We find that the distribution of genome sizes across the phylogenetic tree of Archaea and Bacteria can be linked to the ecosystem type from which the microorganisms' genomes have been extracted (aquatic, host-associated or terrestrial). Finally, we review the ecological factors that may cause the varying sizes of genomes in different ecosystems. In comparison with the aquatic and host-associated ecosystems, terrestrial ecosystems might harbor microorganisms with bigger estimated genome sizes mainly due to higher fluctuations in this ecosystem. Host-associations might shape genomes sizes differentially based on the type of host and level of intimacy between the microorganisms and the host. Genomes in aquatic ecosystems might be shaped by vertical stratification of abiotic factors such as nutrient distribution, light penetration, and temperature. Moreover, different trophic strategies such as auxotrophies might be connected to smaller genome sizes. We expect that as the microbial ecology field keeps moving forward with sequencing, bioinformatics, chemical analysis, and novel cultivation techniques, we will get a deeper resolution on physicochemical, metabolic, spatial, and biological drivers of archaeal and bacterial genome sizes.

REFERENCES

Abram, K., Udaondo, Z., Bleker, C., Wanchai, V., Wassenaar, T. M., Robeson, M. S. II., et al. (2021). Mash-based analyses of *Escherichia coli* genomes reveal 14 distinct phylogroups. *Commun. Biol.* 4:117. doi: 10.1038/s42003-020-01626-5

Aylward, F. O., and Santoro, A. E. (2020). Heterotrophic *Thaumarchaea* with small genomes are widespread in the dark ocean. *mSystems* 5, e00415–e00420. doi: 10.1128/mSystems.00415-20

Bentkowski, P., Van Oosterhout, C., and Mock, T. (2015). A model of genome size evolution for prokaryotes in stable and fluctuating environments. *Genome Biol. Evol.* 7, 2344–2351. doi: 10.1093/gbe/evv148

Bergman, B., Sandh, G., Lin, S., Larsson, J., and Carpenter, E. J. (2013). *Trichodesmium*–a widespread marine cyanobacterium with unusual nitrogen fixation properties. *FEMS Microbiol. Rev.* 37, 286–302. doi: 10.1111/j.1574-6976.2012.00352.x

Berube, P. M., Biller, S. J., Hackl, T., Hogle, S. L., Satinsky, B. M., Becker, J. W., et al. (2018). Single cell genomes of *Prochlorococcus*. *Synechococcus*, and sympatric microbes from diverse marine environments. *Sci. Data* 5:180154. doi: 10.1038/sdata.2018.154

Boussau, B., Karlberg, E. O., Frank, A. C., Legault, B. A., and Andersson, S. G. (2004). Computational inference of scenarios for alpha-proteobacterial genome evolution. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9722–9727. doi: 10.1073/pnas. 0400975101

Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A., and Fierer, N. (2016). Genome reduction in an abundant and ubiquitous soil bacterium 'Candidatus Udaeobacter copiosus'. *Nat. Microbiol.* 2:16198. doi: 10.1038/nmicrobiol.2016. 198

Brochier-Armanet, C., Deschamps, P., Lopez-Garcia, P., Zivanovic, Y., Rodriguez-Valera, F., and Moreira, D. (2011). Complete-fosmid and fosmid-end sequences reveal frequent horizontal gene transfers in marine uncultured planktonic archaea. *ISME J.* 5, 1291–1302. doi: 10.1038/ismej.2011.16

Buck, M., Garcia, S. L., Fernandez, L., Martin, G., Martinez-Rodriguez, G. A., Saarenheimo, J., et al. (2021a). Comprehensive dataset of shotgun metagenomes from oxygen stratified freshwater lakes and ponds. *Sci. Data* 8:131. doi: 10.1038/s41597-021-00910-1

Buck, M., Mehrshad, M., and Bertilsson, S. (2021b). mOTUpan: a robust Bayesian approach to leverage metagenome assembled genomes for core-genome estimation. *bioRxiv* [Preprint]. doi: 10.1101/2021.06.25.449606

Carini, P., Steindler, L., Beszteri, S., and Giovannoni, S. J. (2013). Nutrient requirements for growth of the extreme oligotroph 'Candidatus Pelagibacter ubique' HTCC1062 on a defined medium. *ISME J.* 7, 592–602. doi: 10.1038/Ismej.2012.122

AUTHOR CONTRIBUTIONS

SG, AR-G, and JN conceptualized the literature and data review idea. JN, MB, and FS gathered the data. AR-G and JN performed data analysis. SG, AR-G, and MM drafted the first manuscript. All authors did literature searches, contributed to the writing, and editing of the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb. 2021.761869/full#supplementary-material

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2020). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848

Chen, M. Y., Teng, W. K., Zhao, L., Hu, C. X., Zhou, Y. K., Han, B. P., et al. (2021). Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation. *ISME J.* 15, 211–227. doi: 10.1038/s41396-020-00775-z

Cobo-Simon, M., and Tamames, J. (2017). Relating genomic characteristics to environmental preferences and ubiquity in different microbial taxa. *BMC Genomics* 18:499. doi: 10.1186/s12864-017-3888-y

Collingro, A., Kostlbacher, S., and Horn, M. (2020). *Chlamydiae* in the Environment. *Trends Microbiol.* 28, 877–888. doi: 10.1016/j.tim.2020.05.020

Cross, K. L., Campbell, J. H., Balachandran, M., Campbell, A. G., Cooper, S. J., Griffen, A., et al. (2019). Targeted isolation and cultivation of uncultivated bacteria by reverse genomics. *Nat. Biotechnol.* 37, 1314–1321. doi: 10.1038/s41587-019-0260-6

Dedysh, S. N. (2011). Cultivating uncultured bacteria from northern wetlands: knowledge gained and remaining gaps. *Front. Microbiol.* 2:184. doi: 10.3389/fmicb.2011.00184

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-Gonzalez, A., Eldridge, D. J., Bardgett, R. D., et al. (2018). A global atlas of the dominant bacteria found in soil. *Science* 359, 320–325. doi: 10.1126/science.aap9516

Dharamshi, J. E., Tamarit, D., Eme, L., Stairs, C. W., Martijn, J., Homa, F., et al. (2020). Marine sediments illuminate chlamydiae diversity and evolution. *Curr. Biol.* 30, 1032–1048.e7. doi: 10.1016/j.cub.2020.02.016

D'Souza, G. G., Povolo, V. R., Keegstra, J. M., Stocker, R., and Ackermann, M. (2021). Nutrient complexity triggers transitions between solitary and colonial growth in bacterial populations. *ISME J.* 15, 2614–2626. doi: 10.1038/s41396-021-00953-7

Elser, J. J., Bracken, M. E., Cleland, E. E., Gruner, D. S., Harpole, W. S., Hillebrand, H., et al. (2007). Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol. Lett.* 10, 1135–1142. doi: 10.1111/j.1461-0248.2007.01113.x

Eren, A. M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S. E., Schechter, M. S., et al. (2021). Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* 6, 3–6. doi: 10.1038/s41564-020-00834-3

Figueroa-Gonzalez, P. A., Bornemann, T. L. V., Adam, P. S., Plewka, J., Revesz, F., von Hagen, C. A., et al. (2020). *Saccharibacteria* as Organic carbon sinks in hydrocarbon-fueled communities. *Front. Microbiol.* 11:587782. doi: 10.3389/fmicb.2020.587782

Franche, C., Lindström, K., and Elmerich, C. (2008). Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant Soil* 321, 35–59. doi: 10.1007/s11104-008-9833-8

Garcia, R., Gemperlein, K., and Muller, R. (2014). *Minicystis rosea* gen. nov., sp. nov., a polyunsaturated fatty acid-rich and steroid-producing soil myxobacterium. *Int. J. Syst. Evol. Microbiol.* 64, 3733–3742. doi: 10.1099/ijs.0.068270-0

Garcia, S. L. (2016). Mixed cultures as model communities: hunting for ubiquitous microorganisms, their partners, and interactions. *Aquat. Microb. Ecol.* 77, 79–85. doi: 103354/ame01796

Garcia, S. L., Buck, M., McMahon, K. D., Grossart, H. P., Eiler, A., and Warnecke, F. (2015). Auxotrophy and intrapopulation complementary in the "interactome' of a cultivated freshwater model community. *Mol. Ecol.* 24, 4449–4459. doi: 10.1111/mec.13319

Garcia, S. L., Mehrshad, M., Buck, M., Tsuji, J. M., Neufeld, J. D., McMahon, K. D., et al. (2021). Freshwater chlorobia exhibit metabolic specialization among cosmopolitan and endemic populations. *mSystems* 6, e01196–e01220. doi: 10.1128/mSystems.01196-20

Garcia, S. L., Stevens, S. L. R., Crary, B., Martinez-Garcia, M., Stepanauskas, R., Woyke, T., et al. (2018). Contrasting patterns of genome-level diversity across distinct co-occurring bacterial populations. *ISME J.* 12, 742–755. doi: 10.1038/s41396-017-0001-0

Giovannoni, S. J., Cameron Thrash, J., and Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *ISME J.* 8, 1553–1565. doi: 10.1038/ismej.2014.60

Grote, J., Thrash, J. C., Huggett, M. J., Landry, Z. C., Carini, P., Giovannoni, S. J., et al. (2012). Streamlining and core genome conservation among highly divergent members of the SAR11 clade. *Mbio* 3, e00252–e00312. doi: 10.1128/mBio.00252-12

Grzymski, J. J., and Dussaq, A. M. (2012). The significance of nitrogen cost minimization in proteomes of marine microorganisms. *ISME J.* 6, 71–80. doi: 10.1038/ismej.2011.72

Gweon, H. S., Bailey, M. J., and Read, D. S. (2017). Assessment of the bimodality in the distribution of bacterial genome sizes. *ISME J.* 11, 821–824. doi: 10.1038/ismej.2016.142

Hawkes, J. A., Patriarca, C., Sjöberg, P. J. R., Tranvik, L. J., and Bergquist, J. (2018). Extreme isomeric complexity of dissolved organic matter found across aquatic environments. *Limnol. Oceanogr. Lett.* 3, 21–30. doi: 10.1002/lol2.10064

Henson, M. W., Pitre, D. M., Weckhorst, J. L., Lanclos, V. C., Webber, A. T., Thrash, J. C., et al. (2016). Artificial seawater media facilitate cultivating members of the microbial majority from the gulf of Mexico. *mSphere* 1, e28–e116. doi: 10.1128/mSphere.00028-16

Hoehler, T. M., and Jorgensen, B. B. (2013). Microbial life under extreme energy limitation. *Nat. Rev. Microbiol.* 11, 83–94. doi: 10.1038/nrmicro2939

Imachi, H., Nobu, M. K., Nakahara, N., Morono, Y., Ogawara, M., Takaki, Y., et al. (2020). Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature* 577, 519–525. doi: 10.1038/s41586-019-1916-6

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9

Kang, I., Kim, S., Islam, M. R., and Cho, J.-C. (2017). The first complete genome sequences of the acI lineage, the most abundant freshwater Actinobacteria, obtained by whole-genome-amplification of dilution-to-extinction cultures. *Sci. Rep.* 7:42252. doi: 10.1038/sre42252

Konstantinidis, K. T., and Tiedje, J. M. (2004). Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3160–3165. doi: 10.1073/pnas.0308653100

Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2567–2572. doi: 10.1073/Pnas.0409727102

Levy, A., Salas Gonzalez, I., Mittelviefhaus, M., Clingenpeel, S., Herrera Paredes, S., Miao, J., et al. (2017). Genomic features of bacterial adaptation to plants. *Nat. Genet.* 50, 138–150. doi: 10.1038/s41588-017-0012-9

Lewis, W. H., Tahon, G., Geesink, P., Sousa, D. Z., and Ettema, T. J. G. (2020). Innovations to culturing the uncultured microbial majority. *Nat. Rev. Microbiol.* 19, 225–240. doi: 10.1038/s41579-020-00458-8

Li, L., Liu, Z., Zhou, Z., Zhang, M., Meng, D., Liu, X., et al. (2021). Comparative genomics provides insights into the genetic diversity and evolution of the DPANN superphylum. *mSystems* 6:e00602211. doi: 10.1128/mSystems.00602-21

Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J. Q., and Crosby, L. (2018). Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* 3, e55–e118. doi: 10.1128/mSystems.00055-18

Mallet, J. (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 363, 2971–2986. doi: 10.1098/rstb.2008.0081

McLean, J. S., Bor, B., Kerns, K. A., Liu, Q., To, T. T., Solden, L., et al. (2020). Acquisition and adaptation of ultra-small parasitic reduced genome bacteria to *Mammalian* hosts. *Cell Rep.* 32:107939. doi: 10.1016/j.celrep.2020.107939

Mende, D. R., Bryant, J. A., Aylward, F. O., Eppley, J. M., Nielsen, T., Karl, D. M., et al. (2017). Environmental drivers of a microbial genomic transition zone in the ocean's interior. *Nat. Microbiol.* 2, 1367–1373. doi: 10.1038/s41564-017-0008-3

Meziti, A., Rodriguez, R. L., Hatt, J. K., Pena-Gonzalez, A., Levy, K., and Konstantinidis, K. T. (2021). The reliability of metagenome-assembled genomes (mags) in representing natural populations: insights from comparing mags against isolate genomes derived from the same fecal sample. *Appl. Environ. Microbiol.* 87, e02593–e02620. doi: 10.1128/AEM.02593-20

Mondav, R., Bertilsson, S., Buck, M., Langenheder, S., Lindstrom, E. S., and Garcia, S. L. (2020). Streamlined and abundant bacterioplankton thrive in functional cohorts. *mSystems* 5, e00316–e00420. doi: 10.1128/mSystems.00316-20

Moran, N. A., and Bennett, G. M. (2014). The tiniest tiny genomes. *Annu. Rev. Microbiol.* 68, 195–215. doi: 10.1146/annurev-micro-091213-112901

Morris, J. J., Lenski, R. E., and Zinser, E. R. (2012). The black queen hypothesis: evolution of dependencies through adaptive gene loss. *Mbio* 3, e00036–e00112. doi: 10.1128/mBio.00036-12

Nayfach, S., and Pollard, K. S. (2015). Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. *Genome Biol.* 16:51. doi: 10.1186/s13059-015-0611-7

Nayfach, S., Roux, S., Seshadri, R., Udwary, D., Varghese, N., Schulz, F., et al. (2020). A genomic catalog of Earth's microbiomes. *Nat. Biotechnol.* 39, 499–509. doi: 10.1038/s41587-020-0718-6

Nelson, W. C., Tully, B. J., and Mobberley, J. M. (2020). Biases in genome reconstruction from metagenomic data. *PeerJ.* 8:e101191.

Olm, M. R., Crits-Christoph, A., Diamond, S., Lavy, A., Matheus Carnevali, P. B., and Banfield, J. F. (2020). Consistent metagenome-derived metrics verify and delineate bacterial species boundaries. *mSystems* 5, e00731–e00819. doi: 10.1128/mSystems.00731-19

Ortiz, M., Leung, P. M., Shelley, G., Jirapanjawat, T., Nauer, P. A., Van Goethem, M. W., et al. (2021). Multiple energy sources and metabolic strategies sustain microbial diversity in Antarctic desert soils. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2025322118. doi: 10.1073/pnas.2025322118

Parks, D. H., Chuvochina, M., Chaumeil, P. A., Rinke, C., Mussig, A. J., and Hugenholtz, P. (2020). A complete domain-to-species taxonomy for Bacteria and *Archaea. Nat. Biotechnol.* 38, 1079–1086. doi: 10.1038/s41587-020-0501-8

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Patriarca, C., Sedano-Núñez, V. T., Garcia, S. L., Bergquist, J., Bertilsson, S., Sjöberg, P. J. R., et al. (2020). Character and environmental lability of cyanobacteria-derived dissolved organic matter. *Limnol. Oceanogr.* 66, 496–509. doi: 10.1002/lno.11619

Raes, J., Korbel, J., Lercher, M., von Mering, C., and Bork, P. (2007). Prediction of effective genome size in metagenomic samples. *Genome Biol.* 8:R10. doi: 10.1186/gb-2007-8-1-r10

Rocap, G., Larimer, F. W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N. A., et al. (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* 424, 1042–1047. doi: 10.1038/nature01947

Rodriguez, R. L., Castro, J. C., Kyrpides, N. C., Cole, J. R., Tiedje, J. M., and Konstantinidis, K. T. (2018). How much do rRNA gene surveys underestimate extant bacterial diversity? *Appl. Environ. Microbiol.* 84, e00014–e00118. doi: 10.1128/AEM.00014-18

Rodriguez, R. L., Jain, C., Conrad, R. E., Aluru, S., and Konstantinidis, K. T. (2021). Reply to: "re-evaluating the evidence for a universal genetic boundary among microbial species". *Nat. Commun.* 12:4060. doi: 10.1038/s41467-021-24129-1

Shade, A., Hogan, C. S., Klimowicz, A. K., Linske, M., McManus, P. S., and Handelsman, J. (2012). Culturing captures members of the soil rare biosphere. *Environ. Microbiol.* 14, 2247–2252. doi: 10.1111/j.1462-2920.2012.02817.x

Shapiro, B. J., and Polz, M. F. (2015). Microbial speciation. *Cold Spring Harb. Perspect. Biol.* 7:a0181431.

Steele, J. H., Brink, K. H., and Scott, B. E. (2019). Comparison of marine and terrestrial ecosystems: suggestions of an evolutionary perspective influenced by environmental variation. *ICES J. Mar. Sci.* 76, 50–59. doi: 10.1093/icesjms/fsy149

Swan, B. K., Tupper, B., Sczyrba, A., Lauro, F. M., Martinez-Garcia, M., Gonzalez, J. M., et al. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11463–11468. doi: 10.1073/pnas.1304246110

Tamas, I., Klasson, L., Canback, B., Naslund, A. K., Eriksson, A. S., Wernegreen, J. J., et al. (2002). 50 million years of genomic stasis in endosymbiotic bacteria. *Science* 296, 2376–2379.

Tian, R., Ning, D., He, Z., Zhang, P., Spencer, S. J., Gao, S., et al. (2020). Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. *Microbiome* 8:51. doi: 10.1186/s40168-020-00825-w

Toft, C., and Andersson, S. G. (2010). Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat. Rev. Genet.* 11, 465–475. doi: 10.1038/nrg2798

Varghese, N. J., Mukherjee, S., Ivanova, N., Konstantinidis, K. T., Mavrommatis, K., Kyrpides, N. C., et al. (2015). Microbial species delineation using whole genome sequences. *Nucleic Acids Res.* 43, 6761–6771. doi: 10.1093/nar/gkv657

Wienhausen, G., Noriega-Ortega, B. E., Niggemann, J., Dittmar, T., and Simon, M. (2017). The Exometabolome of two model strains of the roseobacter group: a marketplace of microbial metabolites. *Front. Microbiol.* 8:1985. doi: 10.3389/fmicb.2017.01985

Mycorrhizal Symbionts and Associated Bacteria: Potent Allies to Improve Plant Phosphorus Availability and Food Security

Cristiana Sbrana[1], Monica Agnolucci[2], Luciano Avio[2], Luca Giovannini[2], Michela Palla[2], Alessandra Turrini[2*] and Manuela Giovannetti[2*]

[1] National Research Council (CNR), Institute of Agricultural Biology and Biotechnology, Pisa, Italy, [2] Department of Agriculture, Food and Environment, University of Pisa, Pisa, Italy

INTRODUCTION

Phosphorus (P) is an essential nutrient for food production and a main component of nitrogen-phosphorus-potassium fertilizers used for the cultivation of food crops worldwide. The most significant source of P is phosphate rock, which is mined globally (223 million tons in 2020) in only a few countries, including China (90), Morocco and Western Sahara (37), USA (24) and Russia (13) (U.S. Geological Survey, 2021). As a consequence, the European Union and most world countries are almost entirely dependent on imports of phosphate rock, making them vulnerable to disruptions in its supply and undermining their food security. Moreover, phosphate rock is a finite, non-renewable resource, whose reserves are estimated to be depleted within 100–200 years (Cordell and White, 2014).

In high-intensity agriculture, starting in the mid-20th century with the green revolution, the increased use of P fertilizers has been linked to diverse environmental effects, encompassing eutrophication, the loss of landscape quality, greenhouse gas emissions, excessive fresh-water consumption and cadmium accumulation in food plants (Chen and Graedel, 2016). Yet, there are no substitutes for P in food production systems, as P is essential to life. Indeed, P is a structural component of DNA and RNA (in the form of orthophosphate ions PO_4^{3-}), plays a key role in energy transfer as an element of ATP, and is a building component of biomolecules involved in important biological processes, such as photosynthesis, phospholipid biosynthesis and respiration. The extensive use of P fertilizers has led to very high concentrations of P in soils, but the soluble fraction (PO_4^{3-}), which is the only form directly available for plant nutrition, represents a very small fraction because more than 80% of soil P is insoluble and therefore unavailable for plant uptake. Such dominant P form is represented by organic and mineral P, due to its immobilization in organic matter and precipitation with other soil minerals, i.e., iron (Fe) and aluminum (Al) in acid soils and calcium (Ca) in alkaline soils. Thus, the issue of improving a more efficient use of P is becoming increasingly topical, and diverse potential measures have been considered, in order to reduce P fertilizers in agriculture, while maintaining crop yields and minimizing their environmental impact. Such measures range from optimizing land use by crop rotations, preventing erosion, maintaining soil quality and organic matter, improving fertilizer recommendations and crop genotypes, extending plant root P uptake capacity by means of beneficial microorganisms, primarily mycorrhizal symbionts (Schröder et al., 2011). This opinion article focuses on the current knowledge and prospects for the role of mycorrhizal fungi and their associated bacteria in P availability to food plants.

FIGURE 1 | Schematic drawing representing phosphate pathway from soil to plant via arbuscular mycorrhizal fungi and mycorrhizosphere bacteria.
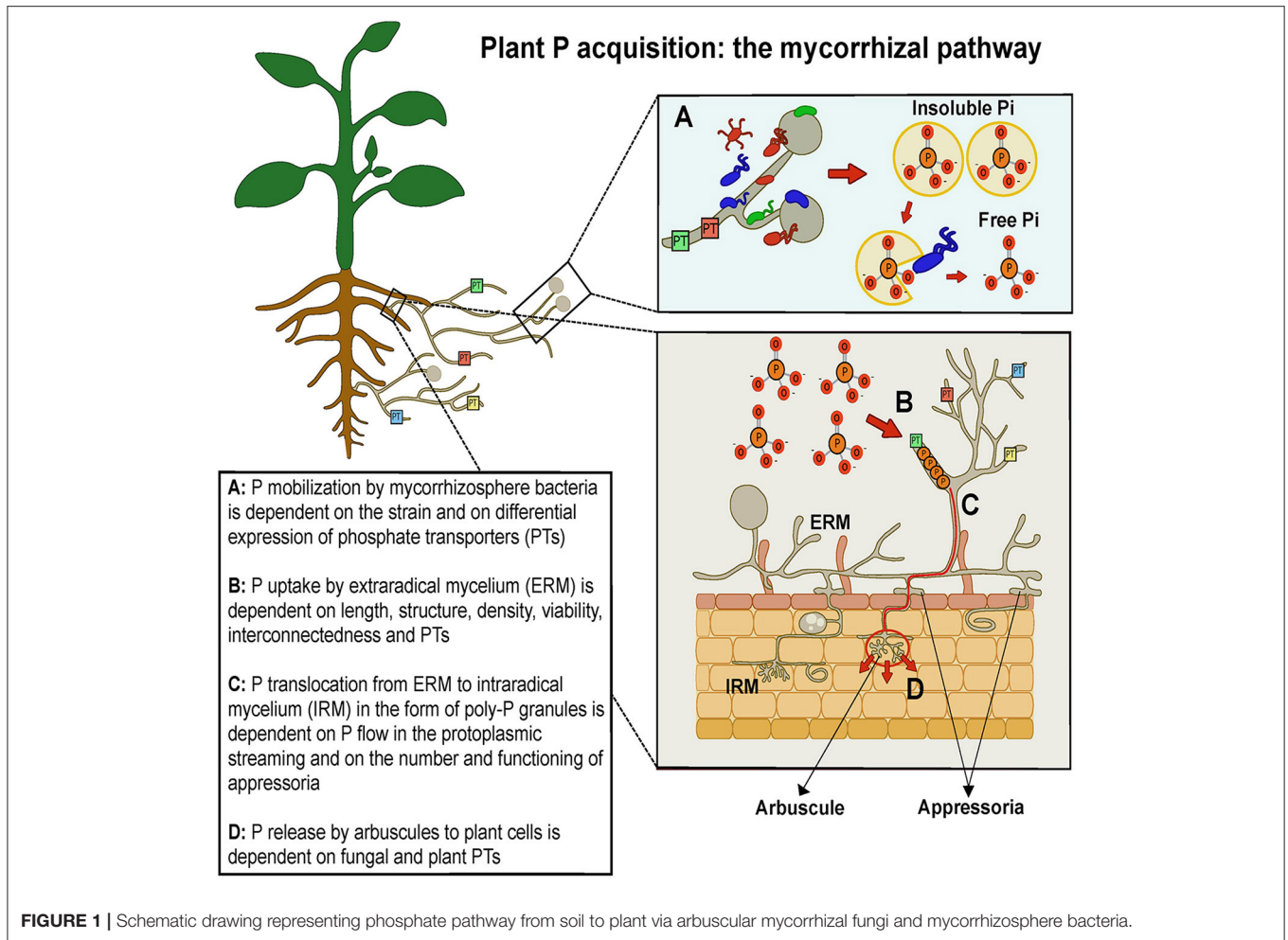
ARBUSCULAR MYCORRHIZAL SYMBIOSES AND SOIL P UPTAKE

Arbuscular mycorrhizal (AM) fungi (AMF), belonging to Mucoromycota, sub-phylum Glomeromycotina (Spatafora et al., 2016), establish beneficial symbiotic associations with the roots of most land plants, including major staple food crops. AMF improve plant nutrition and tolerance to biotic and abiotic stresses and are key organisms in soil nutrient cycles and in the maintenance of biological soil fertility (Smith and Read, 2008). Moreover, AMF affect plant secondary metabolism, enhancing the biosynthesis of health-promoting phytochemicals, such as polyphenols and carotenoids (Agnolucci et al., 2020). In this mutualistic symbiosis AMF obtain photosynthesis-derived carbon (up to 20%) in exchange for soil mineral nutrients, whose uptake and translocation is facilitated by a fine network of extraradical hyphae functioning as an auxiliary absorbing system (**Figure 1**). The direct uptake of P by roots is rapid and results in P depletion zones because P mobility in soil is low; however, the extraradical mycelium (ERM) of AMF spread well beyond root P

depletion zones (Smith and Read, 2008). Thus, ERM represents an important variable affecting fungal absorbing surface and foraging ability, and therefore the rate of plant P uptake from the soil. Using *in vitro* monoxenic cultures, ERM was unequivocally demonstrated to be able to hydrolyse organic phosphate and release acid phosphatase, facilitating P mobilization (Koide and Kabir, 2000; Sato et al., 2019). In different AMF isolates, ERM density ranges between 2.7 and 20.5 m g^{-1} of soil, with a mean growth rate of 0.74–1.1 m d^{-1} (Giovannetti et al., 2001; Mikkelsen et al., 2008). Actually, ERM length was found to be positively correlated with plant shoot biomass and P content in different plants, including clover and maize (Jakobsen et al., 1992; Sawers et al., 2017), while the production of large numbers of appressoria—the structures connecting ERM to intraradical mycelium (IRM)—showed significant positive correlations with shoot and plant P content (Pepe et al., 2020). However, great intraspecific differences were detected among AMF species and isolates in ERM extent, biomass and viability, affecting plant growth responses (Munkvold et al., 2004; Pepe et al., 2017, 2018).

FUNGAL P TRANSPORTERS, P TRANSLOCATION AND TRANSFER TO PLANT CELLS

Beyond ERM structural and functional traits, P uptake is facilitated by fungal P transporters (PTs), whose occurrence and differential expression affect plant P absorption from the soil solution and nutrition. P is absorbed from the soil in the form of PO_4^{3-}, a negatively charged ion requiring metabolic energy and high-affinity PTs for its transport across the fungal plasma membrane (Ferrol et al., 2019). A pioneering study detected the first gene encoding a trans-membrane phosphate transporter (GvPT), which was predominantly expressed in ERM and supposed to play a role in P uptake from the soil by the AM symbiont *Glomus versiforme* (Harrison and van Buuren, 1995). Other PT genes, *GintPT* and *GmosPT*, expressed in *Glomus intradices* and *Glomus mosseae* extraradical and intraradical hyphae, respectively, were found to be regulated in the response to P concentrations in ERM environment (Maldonado-Mendoza et al., 2001; Benedetto et al., 2005; Fiorilli et al., 2013). Recent studies highlighted that four different *Rhizophagus irregularis* PTs, expressed both in ERM and IRM, were modulated by P availability in microcosm systems (Walder et al., 2016; Calabrese et al., 2019). Interestingly, a PT identified in *Gigaspora margarita* was found to be involved also in the P signaling pathway, functioning as a P transceptor (Xie et al., 2016).

P uptake by ERM is followed by P storage in vacuoles as polyphosphate (poly-P) granules, which are translocated through protoplasmic streaming and/or a tubular vacuole system from soil- to root-based hyphae, before being absorbed by plant root cells (Ezawa and Saito, 2018). P flow in ERM is very rapid, given the high rates of hyphal protoplasmic streaming, $3–9\,\mu m\ s^{-1}$ (Logi et al., 1998; Giovannetti et al., 2000) and generates P fluxes from extraradical to intraradical hyphae ranging from 1.7 to $4.2 \times 10^{-8}\ mol\ cm^{-2}\ s^{-1}$, depending on AMF species and isolates (Pepe et al., 2020). In intraradical hyphae poly-P is supposed to be hydrolysed by fungal polyphosphatases and then transferred to the plant cell, although the biochemical and molecular mechanisms of such processes are largely unknown (Ferrol et al., 2019).

The transfer of P released by fungal symbionts to plant cells is mediated by specific plant PTs and membrane-integral proton ATPases. The first plant PTs were detected in mycorrhizal potato, rice and *Medicago truncatula*, for example *StPT3* was expressed in potato root cells colonized by *G. intradices* (Rausch et al., 2001), while *MtPT4* mediated the uptake of P released by the AM symbionts *G. versiforme* and *Gigaspora gigantea* (Harrison et al., 2002). The latter PT was shown to be essential for the acquisition of P released by the fungus and the establishment of the symbiosis (Javot et al., 2007). Indeed, in *Astragalus sinicus*, P transporters *AsPT1* and *AsPT4* were required for the formation of AM symbiosis, as their suppression reduced root colonization and arbuscule development (Xie et al., 2013), while in tomato specific transcripts occurred only in arbuscule-containing cells (Gomez-Ariza et al., 2009). Interestingly, a high plant P status was found to represent a major regulator of PHT1 genes expression

(Nagy et al., 2009). Other mycorrhiza-specific PTs were characterized in different plant species, such as rice, petunia, sorghum and flax, as well as H^+-ATPase genes, whose expression leads to the generation of the proton gradient needed for P uptake (Ferrol et al., 2019). It is important to note that in maize plants colonized by *R. irregularis*, transcripts encoded by *ZmPHT1* genes were correlated with ERM length and P uptake (Sawers et al., 2017), two variables found to be related in previous studies (Jakobsen et al., 1992).

AMF-ASSOCIATED BACTERIA AND POSSIBLE ROLE IN P UPTAKE

Plant P nutrition is boosted also by another component of AMF symbioses, represented by the bacterial communities living closely associated with spores, sporocarps and extraradical mycelium, in the region defined mycorrhizosphere (Barea et al., 2002). The composition of AMF-associated microbiota may depend on both fungal taxon and host plant species (Roesti et al., 2005; Long et al., 2008; Agnolucci et al., 2015). Such bacteria, when isolated in pure culture, showed not only plant growth promoting (PGP) functions and mycorrhiza helper (MH) activities, but also the ability to solubilize P from mineral phosphates and mineralize P from phytates (Giovannini et al., 2020a). Indeed, 26 *Burkholderia* spp. strains were able to solubilize phosphate and to become strongly attached to *R. irregularis* mycelium (Taktek et al., 2015), while 12 out of 128 bacterial strains isolated from *Rhizoglomus irregulare* mycelium showed phosphate-solubilizing activity (Sharma et al., 2020) and 70% of bacteria isolated from *R. irregularis* spores were able to mineralize P (Battini et al., 2016). The mechanism underlying such activities was ascribed to the production of organic acids by AMF and associated microbiota (Andrino et al., 2021). A recent study reported that phosphate solubilising bacteria migrated along extraradical AMF mycelium toward a phytate source, which they were able to mineralise (Jiang et al., 2021). Thus, P-mobilizing bacteria can increase P availability for AMF, therefore playing a key role in AMF P acquisition from soil and plant P nutrition facilitation (**Figure 1**). Actually, a pioneering work posed the question as to whether selected bacterial strains isolated from AMF spores and showing *in vitro* P-mobilizing activities might be able to show the same specific traits also *in planta*, affecting P nutrition in mycorrhizal and control maize plants. The use of radioactive P allowed the detection of *Streptomyces* sp. W94 and *Streptomyces* sp. W77 as the strains producing the largest increases in the uptake and translocation of ^{33}P and the highest enhancement of hyphal length-specific ^{33}P uptake, respectively (Battini et al., 2017).

CONCLUDING REMARKS AND FUTURE PROSPECTS

The fundamental role of AM symbionts in plant P nutrition has been largely investigated and documented. Though, important questions remain to be answered in order to manipulate

the symbiosis and implement AMF inocula in sustainable agroecosystems. First of all, extensive genetic works should be performed on a large number of AMF, given their high genetic and functional diversity, not only at the interspecific but also at the intraspecific level (Koch et al., 2004; Avio et al., 2006; Wyss et al., 2016). Indeed, studies performed on the model species *Rhizophagus irregularis* showed that diverse AMF genotypes differentially interacted with edaphic traits and affected host plant responses (Koch et al., 2006; Angelard et al., 2010; Venegas et al., 2021). It is not yet known whether such differences in plant growth and P uptake are related to the differential ability to absorb P from the soil and translocate it to the host plant. Accordingly, the aim of future works should be to understand the genetic basis of the relationship between ERM structural and functional traits, investigating in particular the differential abundance and expression of P transporter genes in genetically different AMF isolates (Giovannini et al., 2020a).

Systematic studies on the chemical and biochemical events underlying P mobilization by different AMF isolates and mycorrhizosphere bacteria could detect the most efficient combinations showing synergistic activities related to organic acid secretion and phosphatase/phytase enzyme production. Though, another key question remains, as to whether P mobilization from insoluble to soluble forms operated by mycorrhizosphere bacteria can enhance the expression of P transporter genes in ERM hyphae, P uptake and plant P nutrition. In order to utilize the most efficient AMF and associated bacteria consortia as inocula in agriculture, parallel studies are needed on their ability to survive in the new soil environment, to colonize plant roots and to compete with native bacteria and AMF for P scavenging and uptake (Rodriguez and Sanders, 2015; Giovannini et al., 2020b). Such studies could lead to the selection of innovative biofertilizers to be used as inoculants for increasing P acquisition by crop plants, in new and more sustainable food production systems, in the years to come.

AUTHOR CONTRIBUTIONS

CS, MA, AT, and MG conceived the topic of the paper and wrote the original draft. LG, MP, and LA participated in the preparation and review of the manuscript. LG and MP provided editing assistance. All authors have read and agreed to the published version of the manuscript.

FUNDING

REFERENCES

Agnolucci, M., Avio, L., Palla, M., Sbrana, C., Turrini, A., and Giovannetti, M. (2020). Health-promoting properties of plant products: the role of mycorrhizal fungi and associated bacteria. *Agronomy* 10:1864. doi: 10.3390/agronomy10121864

Agnolucci, M., Battini, F., Cristani, C., and Giovannetti, M. (2015). Diverse bacterial communities are recruited on spores of different arbuscular mycorrhizal fungal isolates. *Biol. Fertil. Soils* 51, 379–389. doi: 10.1007/s00374-014-0989-5

Andrino, A., Guggenberger, G., Kernchen, S., Mikutta, R., Sauheitl, L., and Boy, J. (2021). Production of organic acids by arbuscular mycorrhizal fungi and their contribution in the mobilization of phosphorus bound to iron oxides. *Front. Plant Sci.* 12:661842. doi: 10.3389/fpls.2021.661842

Angelard, C., Colard, A., Niculita-Hirzel, H., Croll, D., and Sanders, I. R. (2010). Segregation in a mycorrhizal fungus alters rice growth and symbiosis-specific gene transcription. *Curr. Biol.* 20, 1216–1221. doi: 10.1016/j.cub.2010.05.031

Avio, L., Pellegrino, E., Bonari, E., and Giovannetti, M. (2006). Functional diversity of arbuscular mycorrhizal fungal isolates in relation to extraradical mycelial networks. *New Phytologist* 172, 347–357. doi: 10.1111/j.1469-8137.2006.01839.x

Barea, J. M., Azcon, R., and Azcon-Aguilar, C. (2002). Mycorrhizosphere interactions to improve plant fitness and soil quality. *Antonie Van Leeuwenhoek* 81, 343–351. doi: 10.1023/A:1020588701325

Battini, F., Cristani, C., Giovannetti, M., and Agnolucci, M. (2016). Multi-functionality and diversity of culturable bacterial communities strictly associated with spores of the plant beneficial symbiont *Rhizophagus intraradices*. *Microbiol. Res.* 183, 68–79. doi: 10.1016/j.micres.2015.11.012

Battini, F., Grønlund, M., Agnolucci, M., Giovannetti, M., and Jakobsen, I. (2017). Facilitation of phosphorus uptake in maize plants by mycorrhizosphere bacteria. *Sci. Rep.* 7:4686. doi: 10.1038/s41598-017-04959-0

Benedetto, A., Magurno, F., Bonfante, P., and Lanfranco, L. (2005). Expression profiles of a phosphate transporter gene (*GmosPT*) from the endomycorrhizal fungus *Glomus mosseae*. *Mycorrhiza* 15, 620–627. doi: 10.1007/s00572-005-0006-9

Calabrese, S., Cusant, L., Sarazin, A., Niehl, A., Erban, A., Brulé, D., et al. (2019). Imbalanced regulation of fungal nutrient transports according to phosphate availability in a symbiocosm formed by poplar, sorghum, and *Rhizophagus irregularis*. *Front. Plant Sci.* 10:1617. doi: 10.3389/fpls.2019.01617

Chen, M., and Graedel, T. E. (2016). A half-century of global phosphorus flows, stocks, production, consumption, recycling, and environmental impacts. *Glob. Environ. Chang.* 36, 39–152. doi: 10.1016/j.gloenvcha.2015.12.005

Cordell, D., and White, S. (2014). Life's bottleneck: sustaining the world's phosphorus for a food secure future. *Annu. Rev. Environ. Resour.* 39, 161–188. doi: 10.1146/annurev-environ-010213-113300

Ezawa, T., and Saito, K. (2018). How do arbuscular mycorrhizal fungi handle phosphate? New insight into fine-tuning of phosphate metabolism. *New Phytol.* 220, 1116–1121. doi: 10.1111/nph.15187

Ferrol, N., Azcón-Aguilar, C., and Pérez-Tienda, J. (2019). Arbuscular mycorrhizas as key players in sustainable plant phosphorus acquisition: an overview on the mechanisms involved. *Plant Sci.* 280, 441–447. doi: 10.1016/j.plantsci.2018.11.011

Fiorilli, V., Lanfranco, L., and Bonfante, P. (2013). The expression of GintPT, the phosphate transporter of *Rhizophagus irregularis*, depends on the symbiotic status and phosphate availability. *Planta* 237, 1267–1277. doi: 10.1007/s00425-013-1842-z

Giovannetti, M., Fortuna, P., Citernesi, A. S., Morini, S., and Nuti, M. P. (2001). The occurrence of anastomosis formation and nuclear exchange in intact arbuscular mycorrhizal networks. *New Phytol.* 151, 717–724. doi: 10.1046/j.0028-646x.2001.00216.x

Giovannetti, M., Sbrana, C., and Logi, C. (2000). Microchambers and video-enhanced light microscopy for monitoring cellular events in living hyphae of arbuscular mycorrhizal fungi. *Plant Soil* 226, 153–159. doi: 10.1023/A:1026415419193

Giovannini, L., Palla, M., Agnolucci, M., Avio, L., Sbrana, C., Turrini, A., et al. (2020a). Arbuscular mycorrhizal fungi and associated microbiota as plant biostimulants: research strategies for the selection of the best performing inocula. *Agronomy* 10:106. doi: 10.3390/agronomy10010106

Giovannini, L., Sbrana, C., Avio, L., and Turrini, A. (2020b). Diversity of a phosphate transporter gene among species and isolates of

arbuscular mycorrhizal fungi. *FEMS Microbiol. Lett.* 367:fnaa024. doi: 10.1093/femsle/fnaa024

Gomez-Ariza, J., Balestrini, R., Novero, M., and Bonfante, P. (2009). Cell-specific gene expression of phosphate transporters in mycorrhizal tomato roots. *Biol. Fertil.of Soils* 45, 845–853. doi: 10.1007/s00374-009-0399-2

Harrison, M. J., Dewbre, G. R., and Liu, J. (2002). A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *Plant Cell* 14, 2413–2429. doi: 10.1105/tpc.004861

Harrison, M. J., and van Buuren, M. L. (1995). A phosphate transporter from the mycorrhizal fungus *Glomus versiforme*. *Nature* 378, 626–629. doi: 10.1038/378626a0

Jakobsen, I., Abbott, L. K., and Robson, A. D. (1992). External hyphae of vesicular–arbuscular mycorrhizal fungi associated with *Trifolium subterraneum* L. 1. Spread of hyphae and phosphorus inflow into roots. *New Phytol.* 120, 371–379. doi: 10.1111/j.1469-8137.1992.tb01077.x

Javot, H., Penmetsa, R. V., Terzaghi, N., Cook, D. R., and Harrison, M. J. (2007). A *Medicago truncatula* phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1720–1725. doi: 10.1073/pnas.0608136104

Jiang, F., Zhang, L., Zhou, J., George, T. S., and Feng, G. (2021). Arbuscular mycorrhizal fungi enhance mineralisation of organic phosphorus by carrying bacteria along their extradical hyphae. *New Phytol.* 230, 304–315. doi: 10.1111/nph.17081

Koch, A. M., Croll, D., and Sanders, I. R. (2006). Genetic variability in a population of arbuscular mycorrhizal fungi causes variation in plant growth. *Ecol. Lett.* 9, 103–110. doi: 10.1111/j.1461-0248.2005.00853.x

Koch, A. M., Kuhn, G., Fontanillas, P., Fumagalli, L., Goudet, J., and Sanders, I. R. (2004). High genetic variability and low local diversity in a population of arbuscular mycorrhizal fungi. *PNAS* 101, 2369–2374. doi: 10.1073/pnas.0306441101

Koide, R. T., and Kabir, Z. (2000). Extraradical hyphae of the mycorrhizal fungus *Glomus intraradices* can hydrolyse organic phosphate. *New Phytol.* 148. 511–517. doi: 10.1046/j.1469-8137.2000.00776.x

Logi, C., Sbrana, C., and Giovanetti, M. (1998). Cellular events in survival of individual arbuscular mycorrhizal symbionts growing in the absence of the host. *Appl. Environ. Microbiol.* 64, 3473–3479. doi: 10.1128/AEM.64.9.3473-3479.1998

Long, L., Zhu, H., Yao, Q., and Ai, Y. (2008). Analysis of bacterial communities associated with spores of *Gigaspora margarita* and *Gigaspora rosea*. *Plant Soil* 310, 1–9. doi: 10.1007/s11104-008-9611-7

Maldonado-Mendoza, I. E., Dewbre, G. R., and Harrison, M. J. (2001). A phosphate transporter gene from the extraradical mycelium of an arbuscular mycorrhizal fungus *Glomus intraradices* is regulated in response to phosphate in the environment. *Mol. Plant Microbe Interact.* 14, 1140–1148. doi: 10.1094/MPMI.2001.14.10.1140

Mikkelsen, B. L., Rosendahl, S., and Jakobsen, I. (2008). Underground resource allocation between individual networks of mycorrhizal fungi. *New Phytol.* 180, 890–898. doi: 10.1111/j.1469-8137.2008.02623.x

Munkvold, L., Kjøller, R., Vestberg, M., Rosendahl, S., and Jakobsen, I. (2004). High functional diversity within species of arbuscular mycorrhizal fungi. *New Phytol.* 164, 357–364. doi: 10.1111/j.1469-8137.2004.01169.x

Nagy, R., Drissner, D., Amrhein, N., Jakobsen, I., and Bucher, M. (2009). Mycorrhizal phosphate uptake pathway in tomato is phosphorus-repressible and transcriptionally regulated. *New Phytol.* 181, 950–959. doi: 10.1111/j.1469-8137.2008.02721.x

Pepe, A., Giovannetti, M., and Sbrana, C. (2018). Lifespan and functionality of mycorrhizal fungal mycelium are uncoupled from host plant lifespan. *Sci. Rep.* 8:10235. doi: 10.1038/s41598-018-28354-5

Pepe, A., Giovannetti, M., and Sbrana, C. (2020). Appressoria and phosphorus fluxes in mycorrhizal plants: connections between soil- and plant-based hyphae. *Mycorrhiza* 30, 589–600. doi: 10.1007/s00572-020-00972-w

Pepe, A., Sbrana, C., Ferrol, N., and Giovannetti, M. (2017). An *in vivo* whole-plant experimental system for the analysis of gene expression in extraradical mycorrhizal mycelium. *Mycorrhiza* 7, 659–668. doi: 10.1007/s00572-017-0779-7

Rausch, C., Daram, P., Brunner, S., Jansa, J., Laloi, M., Leggewie, G., et al. (2001). A phosphate transporter expressed in arbuscule-containing cells in potato. *Nature* 414, 462–466. doi: 10.1038/35106601

Rodriguez, A., and Sanders, I. R. (2015). The role of community and population ecology in applying mycorrhizal fungi for improved food security. *ISME J.* 9, 1053–1061. doi: 10.1038/ismej.2014.207

Roesti, D., Ineichen, K., Braissant, O., Redecker, D., Wiemken, A., and Aragno, M. (2005). Bacteria associated with spores of the arbuscular mycorrhizal fungi *Glomus geosporum* and *Glomus constrictum*. *Appl. Environ. Microbiol.* 71, 6673–6679. doi: 10.1128/AEM.71.11.6673-6679.2005

Sato, T., Hachiya, S., Inamura, N., Ezawa, T., Cheng, W., and Tawaraya, K. (2019). Secretion of acid phosphatase from extraradical hyphae of the arbuscular mycorrhizal fungus *Rhizophagus clarus* is regulated in response to phosphate availability. *Mycorrhiza* 29, 599–605. doi: 10.1007/s00572-019-00923-0

Sawers, R. J. H., Svane, S. F., Quan, C., Grønlund, M., Wozniak, B., Gebreselassie, M. N., et al. (2017). Phosphorus acquisition efficiency in arbuscular mycorrhizal maize is correlated with the abundance of root-external hyphae and the accumulation of transcripts encoding PHT1 phosphate transporters. *New Phytol.* 214, 632–643. doi: 10.1111/nph.14403

Schröder, J. J., Smit, A. L., Cordell, D., and Rosemarin, A. (2011). Improved phosphorus use efficiency in agriculture: a key requirement for its sustainable use. *Chemosphere* 84, 822–831. doi: 10.1016/j.chemosphere.2011.01.065

Sharma, S., Compant, S., Ballhausen, M. B., Ruppel, S., and Franken, P. (2020). The interaction between *Rhizoglomus irregulare* and hyphae attached phosphate solubilizing bacteria increases plant biomass of *Solanum lycopersicum*. *Microbiol. Res.* 240:126556. doi: 10.1016/j.micres.2020.126556

Smith, S. E., and Read, D. J. (2008). *Mycorrhizal Symbiosis*. San Diego: Academic Press.

Spatafora, J. W., Chang, Y., Benny, G. L., Lazarus, K., Smith, M. E., Berbee, M. L., et al. (2016). A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* 108, 1028–1046. doi: 10.3852/16-042

Taktek, S., Trépanier, M., Servin, P. M., St-Arnaud, M., Piché, Y., Fortin, J. A., et al. (2015). Trapping of phosphate solubilizing bacteria on hyphae of the arbuscular mycorrhizal fungus *Rhizophagus irregularis* DAOM 197198. *Soil Biol. Biochem.* 90, 1–9. doi: 10.1016/j.soilbio,.2015.07.016

U.S. Geological Survey (2021). *National Minerals Commodity Summary*. Available online at: https://www.usgs.gov/centers/nmic/phosphate-rock-statistics-and-information

Venegas, R. A. P., Lee, S. J., Thuita, M., Mlay, D. P., Masso, C., Vanlauwe, B., et al. (2021). The phosphate inhibition paradigm: host and fungal genotypes determine arbuscular mycorrhizal fungal colonization and responsiveness to inoculation in cassava with increasing phosphorus supply. *Front. Plant. Sci.* 12:693037. doi: 10.3389/fpls.2021.693037

Walder, F., Boller, T., Wiemken, A., and Courty, P. E. (2016). Regulation of plants' phosphate uptake in common mycorrhizal networks: Role of intraradical fungal phosphate transporters. *Plant Signal. Behav.* 11:e1131372. doi: 10.1080/15592324.2015.1131372

Wyss, T., Masclaux, F. G., Rosikiewicz, P., Pagni, M., and Sanders, I. R. (2016). Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus *Rhizophagus irregularis*. *ISME J.* 10, 2514–2526. doi: 10.1038/ismej.2016.29

Xie, X., Huang, W., Liu, F., Tang, N., Liu, Y., Lin, H., et al. (2013). Functional analysis of the novel mycorrhiza-specific phosphate transporter AsPT1 and PHT1 family from *Astragalus sinicus* during the arbuscular mycorrhizal symbiosis. *New Phytol.* 198, 836–852. doi: 10.1111/nph.12188

Xie, X., Lin, H., Peng, X., Xu, C., Sun, Z., Jiang, K., et al. (2016). Arbuscular mycorrhizal symbiosis requires a phosphate transceptor in the *Gigaspora margarita* fungal symbiont. *Mol. Plant* 9, 1583–1608. doi: 10.1016/j.molp.2016.08.011

Check for updates

Correlation Between Antimicrobial Resistance, Virulence Determinants and Biofilm Formation Ability Among Extraintestinal Pathogenic *Escherichia coli* Strains Isolated in Catalonia, Spain

Victoria Ballén, Yaiza Gabasa, Carlos Ratia, Melany Sánchez and Sara Soto*

ISGlobal, Hospital Clínic—Universitat de Barcelona, Barcelona, Spain

Escherichia coli is a well-characterized bacterium highly prevalent in the human intestinal tract and the cause of many important infections. The aim of this study was to characterize 376 extraintestinal pathogenic *E. coli* strains collected from four hospitals in Catalonia (Spain) between 2016 and 2017 in terms of antimicrobial resistance, siderophore production, phylogroup classification, and the presence of selected virulence and antimicrobial resistance genes. In addition, the association between these characteristics and the ability to form biofilms was also analyzed. The strains studied were classified into four groups according to their biofilm formation ability: non-biofilm formers (15.7%), weak (23.1%), moderate (35.6%), and strong biofilm formers (25.6%). The strains were highly resistant to ciprofloxacin (48.7%), trimethoprim-sulfamethoxazole (47.9%), and ampicillin (38%), showing a correlation between higher resistance to ciprofloxacin and lower biofilm production. Seventy-three strains (19.4%) were ESBL-producers. However, no relationship between the presence of ESBL and biofilm formation was found. The virulence factor genes *fimH* (92%), *pgaA* (84.6%), and *irp1* (77.1%) were the most prevalent in all the studied strains. A statistically significant correlation was found between biofilm formation and the presence of *iroN*, *papA*, *fimH*, *sfa*, *cnf*, *hlyA*, *iutA*, and colibactin-encoding genes *clbA*, *clbB*, *clbN*, and *clbQ*. Interestingly, a high prevalence of colibactin-encoding genes (19.9%) was observed. Colibactin is a virulence factor, which interferes with the eukaryotic cell cycle and has been associated with colorectal cancer in humans. Most colibactin-encoding *E. coli* isolates belonged to phylogroup B2, exhibited low antimicrobial resistance but moderate or high biofilm-forming ability, and were significantly associated with most of the virulence factor genes tested. Additionally, the analysis of their clonal relatedness by PFGE showed 48 different clusters, indicating a high clonal diversity among the colibactin-positive strains. Several studies have correlated the pathogenicity of *E. coli* and the presence of virulence factor genes; however, colibactin and its relationship

to biofilm formation have been scarcely investigated. The increasing prevalence of colibactin in *E. coli* and other Enterobacteriaceae and the recently described correlation with biofilm formation, makes colibactin a promising therapeutic target to prevent biofilm formation and its associated adverse effects.

INTRODUCTION

Escherichia coli is a well-characterized bacterium which plays an essential role in the human microbiome. Nevertheless, some strains are responsible for intestinal and extraintestinal infections. Extraintestinal pathogenic *E. coli* (ExPEC) strains are commonly implicated in a variety of infectious diseases occurring in either the community or healthcare settings worldwide, resulting in high economic and social costs (Beloin et al., 2008; Blount, 2015; Manges et al., 2019).

Additionally, *E. coli* can form aggregates and attach to solid surfaces, forming complex structures called biofilms. Bacteria in biofilms secrete various components such as extracellular polymeric substances (EPS), which protect the bacterial community from host immunity and the effects of antibiotics, complicating the infection (Sharma et al., 2016).

In addition to the ability to form biofilms, the ExPEC group has many virulence factor genes (VFGs) encoding adhesins, toxins, siderophores, capsules, and invasins, which are often located into pathogenicity islands (PAIs), plasmids, and other mobile genetic elements (Sarowska et al., 2019). Some of these VFGs can promote biofilm formation. Among the well-characterized VFGs correlated with biofilm formation, type I fimbriae, curli fimbriae, and flagella are the most studied. The *fimA* gene encodes the major subunit of type I fimbriae, which are known to be involved in the first step of biofilm formation (Beloin et al., 2004). Curli fimbriae are encoded by curli-specific genes (*csg*) and are involved in adhesion to surfaces and invasion of eukaryotic host cells (Van Gerven et al., 2018). Flagella play an important role not only in cell motility but also in surface adhesion (Friedlander et al., 2015). Some studies suggest that deletion of some of these genes leads to a reduction in the ability to form biofilm (Guttenplan and Kearns, 2013; Smith et al., 2017).

One of the most recently studied VFGs is colibactin, a secondary metabolite encoded in the genomic island *pks*, which interferes with the eukaryotic cell cycle and has been linked to colorectal cancer in humans (Dziubańska-Kusibab et al., 2020). Interestingly, this PAI is commonly observed among *E. coli* strains belonging to phylogroup B2, including ExPEC, and has been found in isolates from the intestinal microbiota, septicemia, neonatal meningitis, and urinary tract infections (UTIs) (Faïs et al., 2018; Dziubańska-Kusibab et al., 2020).

The aim of this study was to characterize 376 ExPEC strains in terms of antimicrobial resistance, biofilm formation, siderophore production, phylogroup classification, and presence of selected virulence and antimicrobial resistance genes. The association between these features and the ability to form biofilm was also analyzed. Understanding the VFGs that correlate with biofilm production is needed, as these could be considered targets for developing new antimicrobial therapies.

MATERIALS AND METHODS

Bacterial Strains

A total of 376 ExPEC strains were collected from four hospitals in Catalonia, Spain, between 2016 and 2017. Of these, 132 were isolated from blood, 60 from respiratory samples (13 sputum, 12 bronchoalveolar aspirates, 4 tracheal samples, 2 pleural fluid, and 29 non-classified respiratory samples) and 184 from urine (1 from a urinary catheter and 183 from midstream urine). Strains were identified by matrix-assisted laser desorption ionization–time-of-flight mass spectrometry (MALDI-TOF) (Bruker Daltonik GmbH, Bremen, Germany). The modified score values suggested by the manufacturer were used: a score $\geq$ 2.3 meant species identification; a score between 2.0 and 2.299 meant genus identification and probable species identification; a score between 1.7 and 1.9 meant probable genus identification; and a score < 1.69 meant non-reliable identification. Only strains with a score $\geq$ 2.3 classified as *E. coli* were included in the study. After that, the strains were stored in Skim Milk (Becton Dickinson) at $-80°$C.

Biofilm Formation and Quantification

To determine biofilm formation, we performed a protocol previously developed by our group (Cepas et al., 2020). Briefly, strains were cultured in Luria Bertani (LB) agar (Miller's LB AGAR, Condalab) for 18 – 24 h at 37°C. Then, the cell suspension was prepared in 10 mL LB broth and incubated for 18 – 24 h at 37°C with shaking (180 rpm). After incubation, each culture was diluted 1:100 in M63 medium [13.5 g/L KH$_2$PO$_4$, 2 g/L (NH$_4$)$_2$SO$_4$, 5.0 $\times$ 10^{-4} g/L FeSO$_4$, 1 mL 1 M MgSO$_4$.7H$_2$O], supplemented with 0.25% glucose and tested in 96-well flat-bottomed non-treated polystyrene microtiter plates with lids (Nunc$^®$ Edge 2.0, VWR International, Barcelona, Spain) at 30°C for 48 h under static conditions.

The supernatant was then removed and the biofilms were washed once with 1x PBS and dried at 65°C. The plates were stained with crystal violet 2% (CV) for 10 min, washed with 1x PBS, and dried at 65°C. The CV was resuspended with glacial acetic acid 33%, and the biomass was quantified by measuring the optical density at 580 nm using a microplate reader (EPOCH 2 microplate reader; BioTek, VT). The experiment was performed in three technical and biological replicates, and the results were interpreted according to the criteria of

Stepanović et al. (2007). Thus, the strains were classified as non-biofilm formers (OD $\leq$ 0.150), weak biofilm formers ($\geq$ 0.151 OD $\leq$ 0.300), moderate biofilm formers ($\geq$ 0.301 OD $\leq$ 0.60), or strong biofilm formers (OD $\geq$ 0.601). *Escherichia coli* ATCC 25922 was used as positive control, and M63 broth without bacterial inoculum was used as negative control.

Antimicrobial Susceptibility Testing

Antimicrobial resistance profiling was performed using the most representative antimicrobial agents from the different antibiotic families, which are of great clinical and epidemiological relevance. Kirby-Bauer disk diffusion or broth microdilution methods (in the case of colistin) were done according to Clinical and Laboratory Standards Institute (CLSI) guidelines (CLSI, 2020). *Escherichia coli* ATCC 25922 was used as a control strain. The antimicrobials tested by disk diffusion were: amoxicillin/clavulanate (20/10 μg), ampicillin (10 μg), aztreonam (30 μg), cefepime (30 μg), cefotaxime (30 μg), ceftazidime (30 μg), chloramphenicol (30 μg), ciprofloxacin (5 μg), fosfomycin (200 μg/50 μg of glucose-6-phosphate), gentamicin (10 μg), imipenem (10 μg), meropenem (10 μg), and trimethoprim-sulfamethoxazole (1.25/23.75 μg) (Becton Dickinson). Extended-spectrum beta-lactamases were screened by the ESBL test following the CLSI guidelines (CLSI, 2020). Isolates were classified as susceptible, resistant to 1 or 2 antimicrobial categories, multidrug-resistant (MDR) if resistant to at least one agent in $\geq$ 3 antimicrobial categories; or extensively drug-resistant (XDR), if resistant to at least one agent in all but two or fewer antimicrobial categories (Magiorakos et al., 2012).

Identification of Antimicrobial Resistance Genes

Polymerase chain reaction (PCR) assays were performed using the supernatant of a boiled cell suspension of each isolate as DNA template. β-lactamase-encoding genes bla_{SHV-1}, bla_{TEM-1} and the five major groups bla_{CTX-M}: $bla_{CTX-M-1}$, $bla_{CTX-M-2}$, $bla_{CTX-M-8}$, $bla_{CTX-M-9}$, and $bla_{CTX-M-25}$ were detected, as well as genes conferring resistance to sulfonamides ($sul1$ and $sul2$), quinolones [$qnrB$ and $aac(6')-lb-cr$] and colistin ($mcr-1$, $mcr-2$, $mcr-3$, $mcr-4$ and $mcr-5$). Previously characterized strains from our laboratory collection were used as positive controls of the different genes in the corresponding PCR experiments. Water was used as negative control. The PCR products from the strains were sequenced (Genewiz). The obtained sequences were compared with those of the corresponding genes available in the GenBank. Primer sequences (Condalab, Spain) used in the study of antimicrobial resistance genes are listed in **Table 1**.

Virulence Determinants Detection

Virulence factor genes encoding for adhesins (*fimH-1*, *sfa*, *papA*, *pgaA*), siderophores (*iron*, *iutA*, *irp-1*, *iucA*) and toxins (*hlyA*, *cnf-1*, *clbA clbB*, *clbN* and *clbQ*) were detected by PCR. To determine the presence of the complete *pks* genomic island, primers for the four most representative genes were used: *clbA* and *clbQ* as flanking primers, and *clbB* and *clbN* as internal primers (Johnson et al., 2008; Dubois et al., 2010;

Suresh et al., 2018). The primer sequences (Condalab, Spain) used for the detection of the different VFGs are listed in **Table 2**. Previously characterized strains carrying the different VFGs were used as positive controls. Water was used as negative control.

Siderophore Assay

Siderophores production was determined according to the protocol described by Schwyn and Neiland (Schwyn and Neilands, 1987). Briefly, bacterial strains were cultured on chrome azurol S (CAS) (VWR) and hexadecyltrimethylammonium bromide (HDTMA) plates (Fisher Scientific). If a bacterium excretes a siderophore that removes iron from the dye complex, the color of the agar changes from blue to orange. *Acinetobacter baumannii* ATCC19606 was used as positive control.

Phylogroup Assignment Method

Escherichia coli strains were classified into seven phylogroups (A, B1, B2, C, D, E, and F) according to the PCR method designed by Clermont et al. (2013). Strains belonging to our group and whose phylogroup was previously identified were used as controls.

Typing of Colibactin-Positive Strains by Pulsed-Field Gel Electrophoresis

Due to the high prevalence of colibactin among the studied strains, we decided to analyze their clonal relationship by pulsed-field gel electrophoresis (PFGE) of *Xba*I-digested DNA, following the protocol described by Durmaz et al. (2009). The profiles obtained were compared using the InfoQuestTM FPv.5.4 software (Bio-Rad Laboratories) and the unweighted pair group method with arithmetic mean to create dendrograms based on Dice's similarity coefficient. Isolates were clustered together if their similarity index was $\geq$ 85%.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows software, version 21.0. The Chi-square test was used to evaluate correlations among variables. p-values < 0.05 were considered statistically significant.

RESULTS

Biofilm Formation and Quantification

Escherichia coli strains were analyzed for biofilm formation using M63 broth. Fifty-nine strains (15.7%) were classified as non-biofilm-forming isolates, 87 (23.1%) were classified as weakly biofilm-forming, 134 (35.6%) as moderately biofilm-forming, and 96 (25.6%) as strongly biofilm-forming strains.

Furthermore, the ability to form biofilm was investigated according to the origin of the isolates (urine, blood or respiratory tract) as shown in **Figure 1**; however, no correlation between strain source and biofilm formation ability was found ($p > 0.05$).

Antimicrobial Susceptibility Testing

The percentages of antimicrobial resistance are shown in **Figure 2**. Overall, the highest percentages of antibiotic resistance

TABLE 1 | Primers to detect antimicrobial resistance genes.

Target gene	Primer sequence (5′ → 3′)	Melting Temperature (Tm °C)	Product size (bp)	References
$bla_{TEM\text{-}1}$ – F	TCGCCGCATACACTATTCTCAGAATGA	53	445	Monstein et al., 2007
$bla_{TEM\text{-}1}$ – R	ACGCTCACCGGCTCCAGATTTAT			
$bla_{SHV\text{-}1}$ – F	ATGCGTTATATTCGCCTGTG	49	747	Wiegand et al., 2007
$bla_{SHV\text{-}1}$ – R	TGCTTTGTTATTCGGGCCAA			
$bla_{CTX\text{-}M\text{-}1}$ – F	AAAAATCACTGCGCCAGTTC	52	415	Woodford et al., 2006
$bla_{CTX\text{-}M\text{-}1}$ – R	AGCTTATTCATCGCCACGTT			
$bla_{CTX\text{-}M\text{-}2}$ – F	CGACGCTACCCCTGCTATT	52	552	Woodford et al., 2006
$bla_{CTX\text{-}M\text{-}2}$ - R	CCAGCGTCAGATTTTTCAGG			
$bla_{CTX\text{-}M\text{-}9}$ – F	CAAAGAGAGTGCAACGGATG	52	205	Woodford et al., 2006
$bla_{CTX\text{-}M\text{-}9}$ – R	ATTGGAAAGCGTTCATCACC			
$bla_{CTX\text{-}M\text{-}8}$ – F	TCGCGTTAAGCGGATGATGC	52	666	Woodford et al., 2006
$bla_{CTX\text{-}M\text{-}8}$ – R	AACCCACGATGTGGGTAG			
$bla_{CTX\text{-}M\text{-}25}$ – F	GCACGATGACATTCGGG	52	327	Woodford et al., 2006
$bla_{CTX\text{-}M\text{-}25}$ - R	AACCCACGATGTGGGTAG			
$bla_{CTX\text{-}M\text{-}15/28}$ – F	ATAAAACCGGCAGCGGTG	55	483	Leflon-Guibout et al., 2004
$bla_{CTX\text{-}M\text{-}15/28}$ - R	GAATTTTGACGATCGGGG			
$bla_{CTX\text{-}M\text{-}14/27}$ – F	CGCTTTATGCGCAGACGA	57	785	Pai et al., 2001
$bla_{CTX\text{-}M\text{-}14/27}$ - R	GATTCTCGCCGCTGAAGC			
sul1 – F	CTTCGATGAGAGCCGGCGGC	63	436	Guerra et al., 2004
sul1 – R	GCAAGGCGGAAACCCGCGCC			
sul2 – F	TCAACATAACCTCGGACAGT	55	707	Guerra et al., 2004
sul2 – R	GATGAAGTCAGCTCCACCT			
aac(6′)-Ib – cr –F	TTGCGATGCTCTATGAGTGGCTA	60	482	Park et al., 2006
aac(6′)-Ib – cr – R	CTCGAATGCCTGGCGTGTTT			
qnrB – F	GATCGTGAAAGCCAGAAAGG	52	469	Robicsek et al., 2006
qnrB – R	ACGATGCCTGGTAGTTGTCC			
mcr-1 – F	ATGCCAGTTTCTTTCGCGTG	59	502	Lescat et al., 2018
mcr-1 – R	TCGGCAAATTGCGCTTTTGGC			
mcr-2 – F	GATGGCGGTCTATCCTGTAT	59	379	Lescat et al., 2018
mcr-2 – R	AAGGCTGACACCCCATGTCAT			
mcr-3 – F	ACCAGTAAATCTGGTGGCGT	59	296	Lescat et al., 2018
mcr-3 – R	AGGACAACCTCGTCATAGCA			
mcr-4 – F	TTGCAGACGCCCATGGAATA	59	207	Lescat et al., 2018
mcr-4 – R	GCCGCATGAGCTAGTATCGT			
mcr-5 – F	GGACGCGACTCCCTAACTTC	59	608	Lescat et al., 2018
mcr-5 – R	ACAACCAGTACGAGAGCACG			

were observed with ciprofloxacin (48.7%), trimethoprim-sulfamethoxazole (47.9%), and ampicillin (38%). Only two strains were resistant to colistin (0.5%), and one to fosfomycin (0.3%). None was resistant to carbapenems (imipenem or meropenem). To determine whether biofilm formation correlates with antimicrobial resistance, the different categories of biofilm formation were compared with resistance profiles as shown in **Figure 3**. Non-biofilm-forming strains showed higher percentages of resistance to ciprofloxacin (74.6%) compared to biofilm-forming strains (**Figure 3H**). Thus, the higher the resistance to ciprofloxacin, the lower the biofilm production ability ($p < 0.0001$).

Regarding the antimicrobial resistance classification, 91 strains (24.2%) were susceptible to all the antimicrobial categories tested, 142 (37.8%) were resistant to one or two antimicrobial categories, 143 (38%) isolates were MDR, and none was XDR. The percentages for biofilm formation by

antimicrobial resistance classification are shown in **Figure 4**. A direct relationship between antimicrobial susceptibility and ability to form biofilm was observed. However, no statistical relationship was found between biofilm formation ability and the antimicrobial resistance classification ($p = 0.053$).

Identification of Antimicrobial Resistance Genes

In total, 73 (19.4%) isolates were found to be ESBL-producing strains. The most common β-lactamase-encoding gene was $bla_{CTX\text{-}M\text{-}15}$ ($n = 43$, 58.9%), followed by $bla_{TEM\text{-}1}$ ($n = 33$, 45.2%), $bla_{CTX\text{-}M\text{-}14}$ ($n = 8$, 12.3%), $bla_{SHV\text{-}1}$ ($n = 7$, 9.6%), $bla_{CTX\text{-}M\text{-}27}$ ($n = 5$, 6.8%) and $bla_{CTX\text{-}M\text{-}28}$ ($n = 3$, 4.1%). None carried the $bla_{CTX\text{-}M\text{-}2}$, $bla_{CTX\text{-}M\text{-}8}$ or $bla_{CTX\text{-}M\text{-}25}$ genes, but 27 strains (37%) harbored more than one β-lactamase-encoding gene.

TABLE 2 | Primers to detect virulence factor genes.

Target gene	Primer sequence (5′ → 3′)	Melting temperature (Tm °C)	Product size (bp)	Function	References
iroN – F	AAGTCAAAGCAGGGGTTGCCCG	56	827	Siderophore uptake transmembrane transporter activity	Johnson et al., 2000
iroN – R	GACGCCGACATTAAGACGCAG				
irp1 – F	GGCGTCTCCTCCTTTGGTATT	60	1729	Gene encoding for an iron regulatory protein	Xu et al., 2000
irp1 – R	GTGATTCCCGCTGTTGATGTT				
iucA – F	AGTCTGCATCTTAACCTTCA	56	1100	Gene encoding for an aerobactin	Guerrieri et al., 2019
iucA – R	CTCGTTATGATCGTTCAGAT				
papA – F	ATGGCAGTGGTGTCTTTTGGTG	62	717	Fimbrial major pilin protein precursor	Johnson and Stell, 2000
papA – R	CGTCCCACCATACGTGCTCTTC				
fimH – F	CAGCGATGATTTCCAGTTTGTGTG	59	461	Type 1 fimbrin D-mannose specific adhesin precursor	Sáez-López et al., 2017
fimH – R	TGCGTACCAGCATTAGCAATGTCC				
sfa – F	CTCCGGAGAACTGGGTGCATCTTAC	65	410	S fimbriae	Houdouin et al., 2006
sfa – R	CGGAGGAGTAATTACAAACCTGGCA				
cnf-1 – F	AAGATGGAGTTTCCTATGCAGGAG	56	498	Cytotoxic necrotizing factor	Takahashi et al., 2006
cnf-1 – R	CATTCAGAGTCCTGCCCTCATTATT				
hlyA – F	AACAAGGATAAGCACTGTTCTGGC	59	1177	Hemolysin	Johnson and Stell, 2000
hlyA – R	ACCATATAAGCGGTCATTCCCGTCA				
iutA – F	GGCTGGACATCATGGGAACTGG	60	300	Ferric aerobactin receptor precursor	Johnson and Stell, 2000
iutA – R	CGTCGGGAACGGGTAGAATCG				
pgaA – F	GGCTTTGAAACTTCTTACTGC	60	209	Poly-beta-1,6-N-acetyl-D-glucosamine export protein	Shrestha et al., 2019
pgaA – R	CCTGTTTATCTTGCCCGGCC				
clbQ – F	CTTGTATAGTTACACAACTATTTC	54	821	Colibactin biosynthesis thioesterase ClbQ	Morgan et al., 2019
clbQ – R	TTATCCTGTTAGCTTTCGTTC				
clbA – F	CTAGATTATCCGTGGCGATTC	54	1002	Colibactin biosynthesis phosphopantetheinyl transferase ClbA	Morgan et al., 2019
clbA – R	CAGATACACAGATACCATTCA				
clbB – F	GATTTGGATACTGGCGATAACCG	54	579	Colibactin hybrid non-ribosomal peptide synthetase/type I polyketide synthase ClbB	Johnson et al., 2008
clbB – R	CCATTTCCCGTTTGAGCACAC				
clbN – F	GTTTTGCTCGCCAGATAGTCATTC	54	733	Colibactin non-ribosomal peptide synthetase ClbN	Johnson et al., 2008
clbN – R	CAGTTCGGGTATGTGTGGAAGG				

Regarding biofilm formation, 17 of the ESBLs-producing strains (23.3%) were non-biofilm-formers, 15 (20.5%) were weak biofilm-formers, 25 (34.2%) were moderate biofilm-formers, and 16 (21.9%) were strong biofilm-formers. However, no relationship was found between ESBL production and biofilm formation.

Of the 190 strains resistant to trimethoprim-sulfamethoxazole, 78 (41.1%), 131 (68.9%), and 47 strains (24.7%) harbored *sul1, sul2,* or both genes, respectively.

Among the 188 quinolone-resistant strains, 6 (3.2%) *qnrB* and 40 (21.3%) *aac (6')-Ib-cr* genes were detected. We observed a higher prevalence of the *aac (6')-Ib-cr* gene among the non-biofilm-forming strains, being the correlation statistically significant ($p = 0.008$). We detected two colistin-resistant strains, one of which carried the *mcr-1* gene.

Virulence Determinants Detection

The results about the prevalence of the different VFGs tested are shown in **Figure 5**. It was observed that the prevalence of siderophore-related genes was variable: *irp1* (77.1%), *iutA* (66.8%), *iucA* (52.7%), and *iroN* (45.5%). A significant correlation was found between biofilm formation ability and the presence of
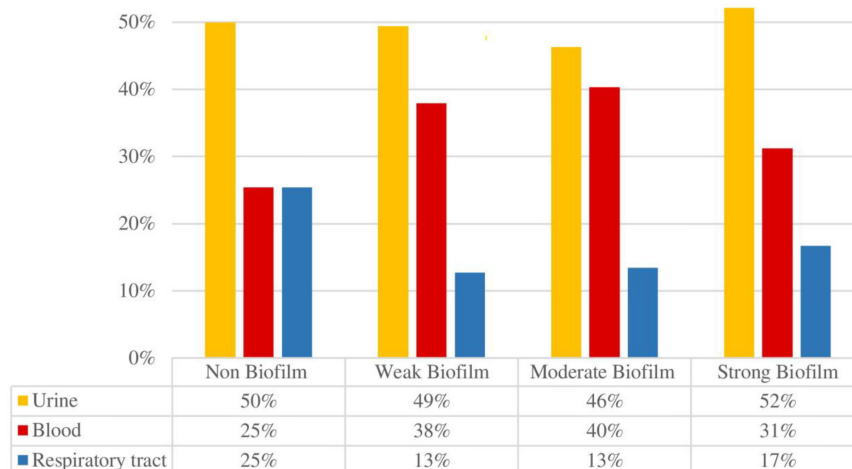
FIGURE 1 | Biofilm formation according to the origin of the isolates.
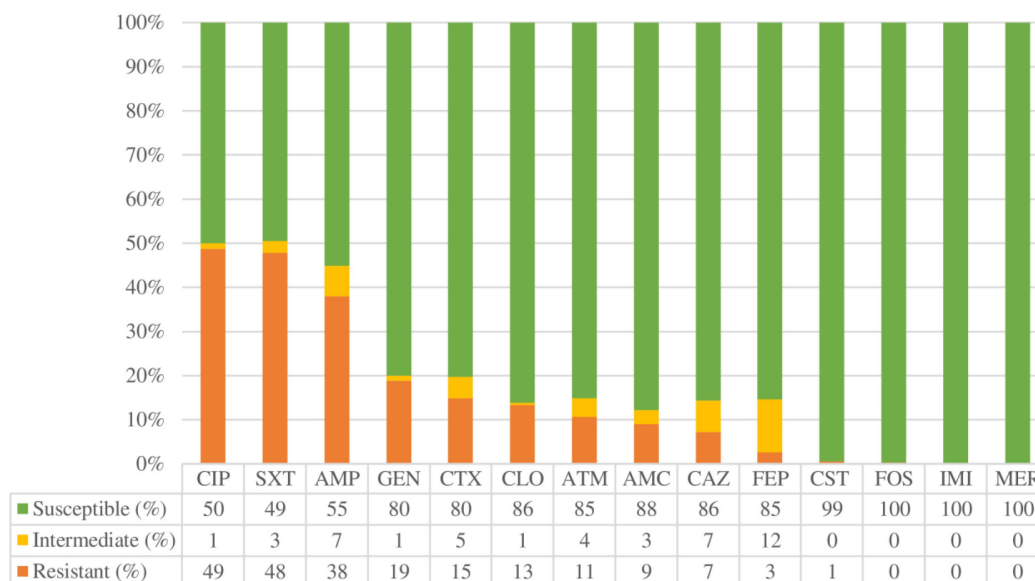
FIGURE 2 | Overall antimicrobial resistance rates expressed in percentage (%). AMC: Amoxicillin/clavulanic acid, AMP: Ampicillin, ATM: Aztreonam, CLO: Chloramphenicol, CAZ: Ceftazidime, CIP: Ciprofloxacin, CST: Colistin, CTX: Cefotaxime, FEP: Cefepime, FOS: Fosfomycin, GEN: Gentamicin, IMI: Imipenem, MER: Meropenem, SXT: Trimethoprim-sulfamethoxazole.

iroN ($p < 0.001$) or *iutA* ($p = 0.010$) genes. However, *iutA* gene was more prevalent among non-biofilm-forming strains.

The adhesin-encoding genes *fimH* and *pgaA* were the most frequent genes among the strains, with a prevalence of 92% and 84.6%, respectively. The prevalence of the other adhesin-encoding genes, such as *papA* and *sfa*, was 43.1% and 21.8%, respectively. A statistically significant correlation was found between biofilm formation and the presence of the *papA, fimH,* and *sfa* genes ($p < 0.001$).

Among the genes encoding toxins, *hlyA* and *cnf-1* showed a prevalence of 25% and 24.7%, respectively. The *clbA, clbB,* and *clbN* genes were present in 19.9% of the strains. *clbQ* gene

was observed in 20.5% of the isolates. It is of note that a statistically significant correlation was confirmed between the presence of toxin-encoding genes and the ability to form biofilm: *hlyA* ($p = 0.0002$), *cnf* ($p < 0.001$), and colibactin-encoding genes (*clbA, clbB, clbN* and *clbQ*) ($p < 0.001$).

Siderophore Assay (Chrome Azurol S Test)

The 376 isolates were screened on CAS agar plates, a useful method for identifying siderophores in *E. coli* isolates and other Gram-negative bacteria. 199 (52.9%) siderophore-producing
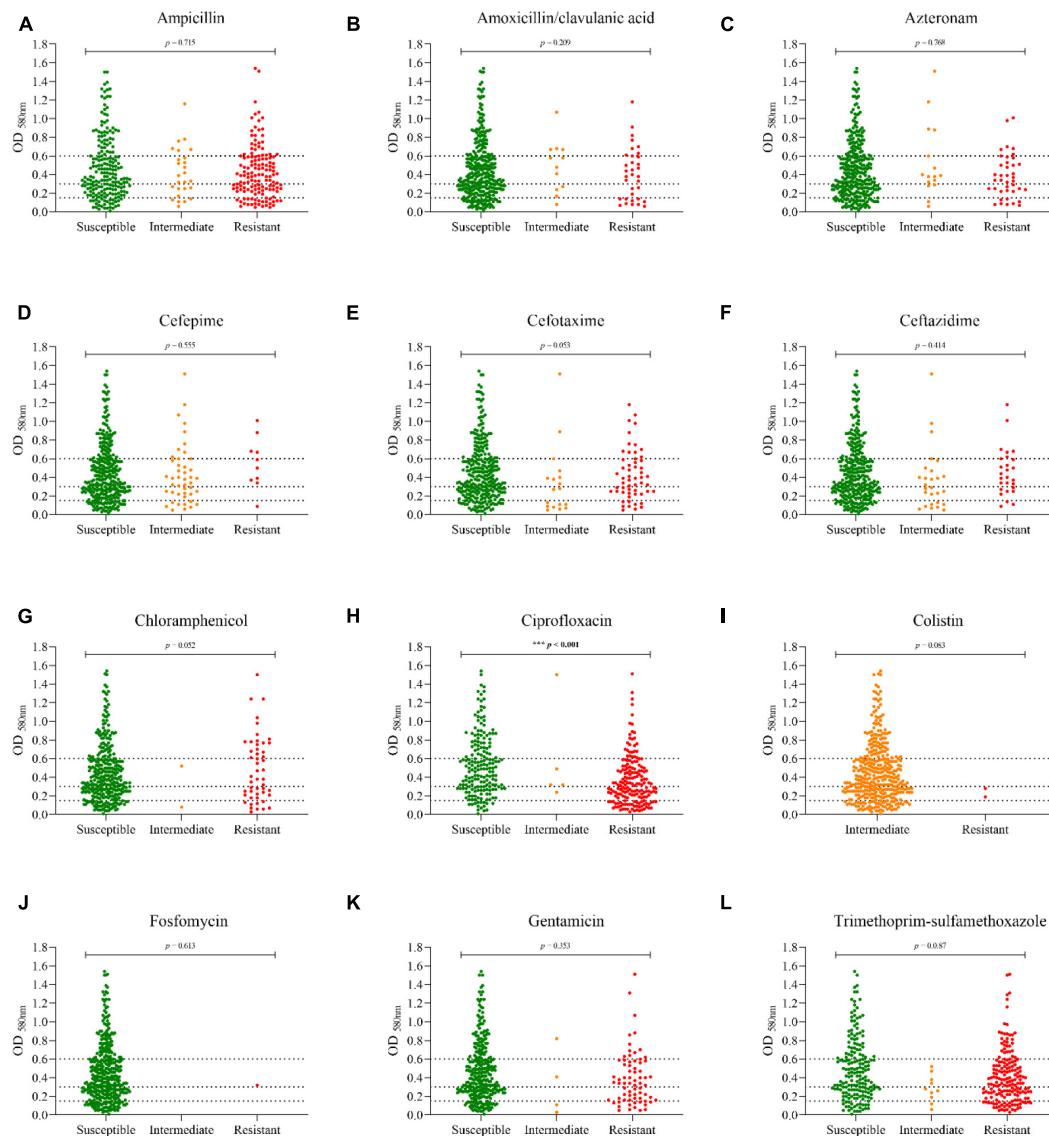
FIGURE 3 | Distribution of biofilm formation among the different antibiotic resistance profiles. **(A)** Ampicillin; **(B)** Amoxicillin/Clavulanic acid; **(C)** Aztreonam; **(D)** Cefepime; **(E)** Cefotaxime; **(F)** Ceftazidime; **(G)** Chloramphenicol; **(H)** Ciprofloxacin; **(I)** Colistin; **(J)** Fosfomycin; **(K)** Gentamicin; **(L)** Trimethoprim-sulfamethoxazole. The dotted lines indicate the cut-off points for each biofilm category [non-biofilm formers (OD $\leq$ 0.150), weak biofilm formers ($\geq$ 0.151 OD $\leq$ 0.300), moderate biofilm formers ($\geq$ 0.301 OD $\leq$ 0.60), or strong biofilm formers (OD $\geq$ 0.601)].

strains were found, as shown in **Figure 5**, but statistical analysis showed no relationship between CAS test positivity and biofilm formation ($p = 0.132$).

On the other hand, a statistically significant association between the positive CAS method and the presence of different siderophore-encoding genes *irp1*, *iucA*, *iutA* ($p < 0.001$) was found. The presence of the fimbriae H coding gene (*fimH*) was also correlated with the CAS test ($p = 0.025$). Contrarily, an association between the CAS test and the absence of the *sfa* gene ($p = 0.024$) was observed.

It is to note that strains resistant to antibiotics such as gentamicin ($p = 0.032$), ampicillin ($p = 0.001$), cefotaxime ($p = 0.001$), ceftazidime ($p = 0.002$), cefepime ($p = 0.001$) or

aztreonam ($p < 0.001$) showed high siderophore production (positive CAS test). Likewise, ESBL-producing strains were statistically associated with a positive CAS test ($p = 0.001$). Finally, strains belonged to phylogroup B2 produced more siderophores than strains belonging to the other phylogroups ($p < 0.001$).

Phylogroup Assignment Method

In our study, the phylogenetic groups considered more virulent B2 and D, accounted for 72.9% of the *E. coli* isolates [B2: $n = 235$ (62.5%); D: $n = 39$ (10.4%)]. The less virulent groups were found in varying percentages [A: $n = 18$ (4.8%); B1: $n = 31$ (8.2%);
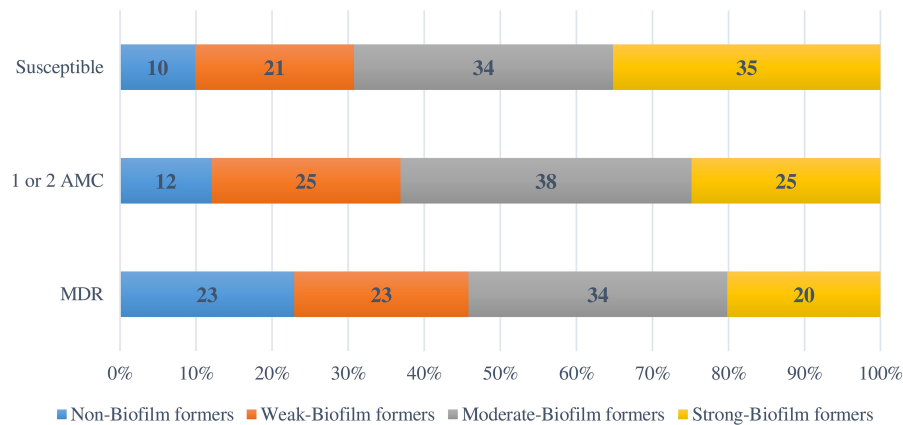
FIGURE 4 | Biofilm formation according to the antimicrobial resistance classification. AMC: Antimicrobial categories; MDR: Multidrug-resistant.
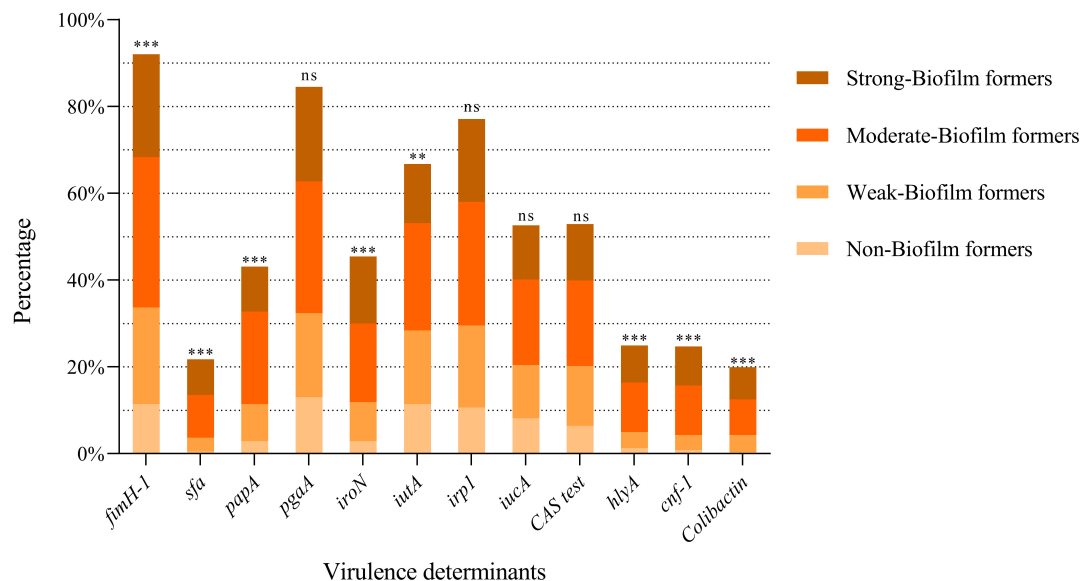
FIGURE 5 | Prevalence of the different VFGs tested according to the biofilm formation ability. Each bar shows the percentage of positivity of each gene within the total strains studied. Within each bar, the percentage according to the ability to form biofilms is shown. ***$p < 0.001$; **$p = 0.01$; ns: not significant ($p > 0.05$).

C: $n = 18$ (4.8%); E: $n = 4$ (1.1%); F: $n = 26$ (6.9%); unknown: $n = 5$ (1.3%)].

According to biofilm classification, a great variety of phylogroups was observed in the non-biofilm-forming strains group. In contrast, a direct relationship between the ability to form biofilm and phylogroup B2 was observed, being this relationship statistically significant ($p < 0.001$) (**Figure 6**).

Characterization of Colibactin-Encoding *Escherichia coli* Strains

The presence of *clbA, clbB, clbN,* and *clbQ* genes confirmed the full presence of the *pks*-island in 75 strains. Among these, 74 presented the ability to form biofilm. We compared whether the presence of colibactin-encoding genes was related to

biofilm formation ability and we found a statistically significant relationship between these two variables ($p < 0.001$).

As shown in **Table 3**, colibactin-producing strains showed a high prevalence of most of the VFGs tested (*iroN, irp1, papA, fimH, cnf, hlyA, pgaA,* and *sfa*) ($p < 0.05$), but had a low prevalence of the *iucA* and *iutA* genes. The relationship between susceptibility to most of the antimicrobials tested (amoxicillin/clavulanic acid, ampicillin, aztreonam, cefepime, cefotaxime, ceftazidime, chloramphenicol, ciprofloxacin, gentamicin, and trimethoprim-sulfamethoxazole) and the presence of colibactin genes was statistically significant ($p < 0.05$). Only two of the colibactin-encoding strains were ESBL-producers, showing an inverse correlation between colibactin genes and ESBL production ($p < 0.001$). Most colibactin-positive strains were isolated from urine ($n = 45$;

FIGURE 6 | Phylogroup assignment according to the biofilm formation. UNK: Unknown. **(A)** Non-Biofilm former strains. **(B)** Weak-Biofilm former strains. **(C)** Moderate-Biofilm former strains. **(D)** Strong-Biofilm former strains.
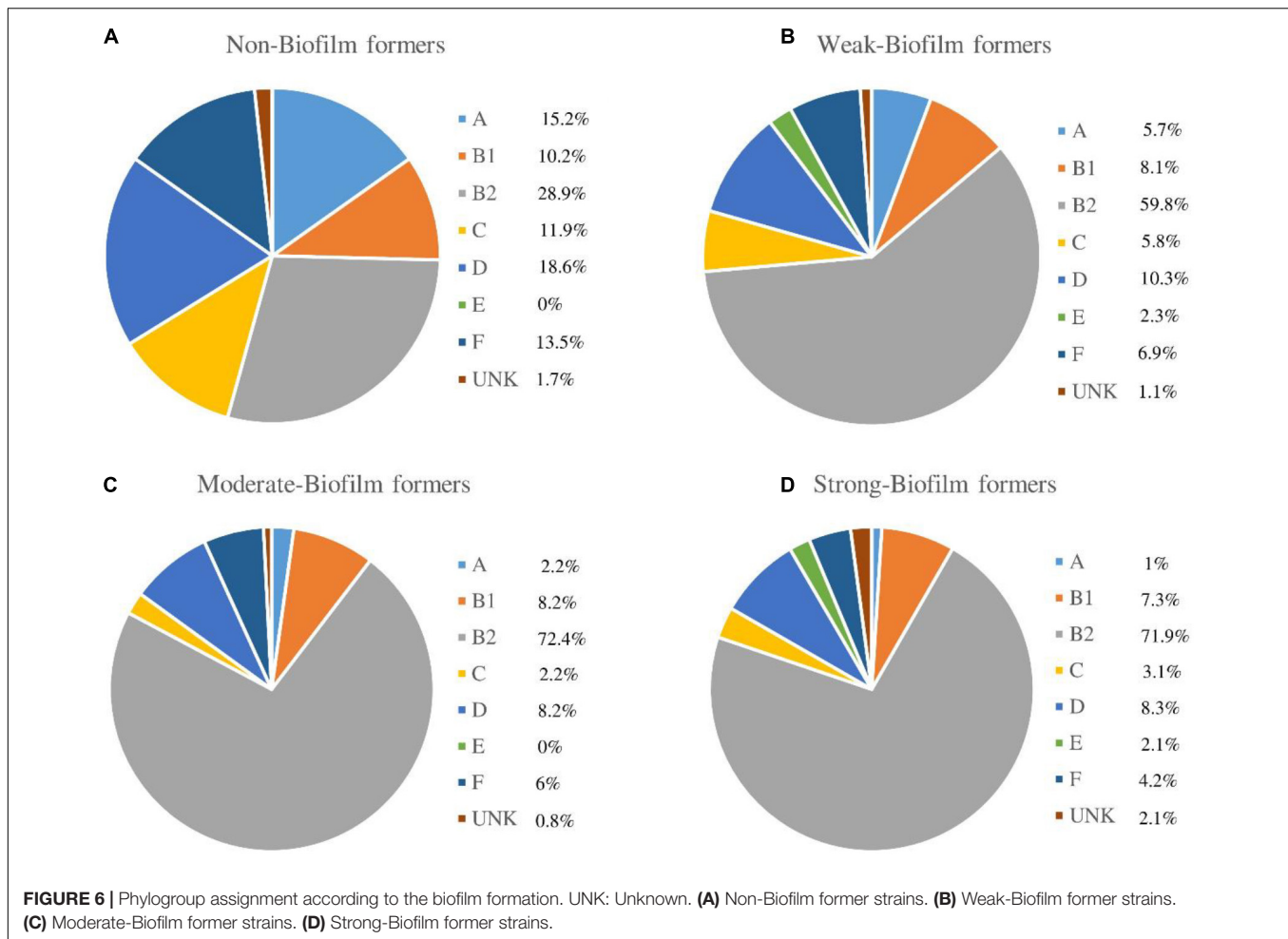
60%), followed by strains from blood (n = 22; 29.3%), and the respiratory tract (n = 8; 10.7%), but no correlation was found between the strain source and the presence of colibactin-encoding genes.

The majority of the colibactin-positive strains belonged to phylogroup B2, and the analysis of the clonal relatedness by PFGE of the colibactin-positive *E. coli* strains showed 48 different clusters with a Dice similarity index ≥ 85%. This indicates that the *E. coli* isolates harboring the colibactin toxin exhibit a high clonal diversity (**Figure 7**).

DISCUSSION

Extraintestinal pathogenic *E. coli* strains are a group of bacteria that can cause urinary tract, bloodstream and other non-intestinal infections, both in healthcare settings and in the community (Manges et al., 2019). Several studies have investigated the association between antimicrobial resistance and/or the presence of some VFGs and the ability to form biofilms (May and Okabe, 2011; Friedlander et al., 2015; Cepas et al., 2019). However, the virulence genes described in recent years need to be studied in more detail, as they may be related

to biofilm formation. In this study, we characterized 376 strains to investigate the possible relationship between antimicrobial resistance and/or the presence of selected virulence determinants and the ability to form biofilms. Our results showed that almost all the isolates studied (84.3%) were able to form biofilms, and this high rate might be related to some of the genes and features observed among the strains studied.

In terms of antimicrobial resistance, we found high rates of resistance to ciprofloxacin (48.7%), trimethoprim-sulfamethoxazole (47.9%), and ampicillin (38%), and fully susceptible strains to carbapenems (imipenem and meropenem). Comparing our results with the SMART study (Cantón et al., 2021), which investigated Gram-negative bacilli isolated from intra-abdominal, urinary, bloodstream and lower respiratory tract infection from 2016 to 2018 in 11 participating Spanish Hospitals, we found similar resistance rates for cephalosporines, carbapenems and colistin but their antimicrobial resistance rates to ciprofloxacin were lower (32.3%). Besides, they found a lower rate of ESBL-producing strains (8.6%). According to the data from the 2017 European Antimicrobial Resistance Surveillance Network report (European Centre for Disease Prevention and Control [ECDC], 2017), which is based on antimicrobial resistance data from invasive isolates, we

TABLE 3 | Characterization of colibactin-encoding *Escherichia coli* strains.

Strain	Origin	Phylogroup	Biofilm	Adhesines				Siderophores					Toxins				ARGs			
				fimH	*sfa*	*papA*	*pgaA*	*iroN*	*iutA*	*irp1*	*iucA*	CAS test	*hlyA*	*cnf1*	*clbA*	*clbQ*	*bla*ctx-m-15	*bla*ctx-m-14	*sul1*	*sul2*
HUB561	Urine	B2	Moderate	✓			✓		✓		✓	✓			✓	✓				▲
HUB562	Urine	B2	Moderate	✓	✓		✓	✓		✓		✓	✓	✓	✓	✓				
HUB564	Urine	B2	Weak	✓	✓		✓		✓		✓	✓	✓	✓	✓	✓				
HUB569	Urine	B2	Strong	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB0012	Blood	B2	Moderate	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HMT142	Blood	B2	Strong	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB0015	Blood	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HMT0019	Urine	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			▲	▲
HCB0029	Blood	B2	Moderate	✓		✓	✓		✓		✓	✓			✓	✓				
HCB0031	Blood	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓	✓				
HCB0035	Blood	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓				
HCB881	Blood	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB0038	Blood	B2	Weak	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	▲			
HCB0039	Blood	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			▲	▲
HCB0042	Urine	B2	Strong	✓			✓	✓	✓	✓	✓		✓	✓	✓	✓			▲	▲
HCB0044	Urine	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB0045	Urine	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				▲
HCB0047	Urine	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				▲
HCB883	Urine	B2	Strong	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB196	Blood	B2	Weak	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB0113	Respiratory	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB197	Respiratory	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB198	Respiratory	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB889	Urine	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB143	Urine	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB150	Urine	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB154	Urine	B2	Weak	✓		✓	✓		✓	✓	✓		✓	✓	✓	✓				
HUB576	Urine	B2	Strong	✓	✓		✓	✓		✓		✓	✓	✓	✓	✓				
HUB582	Blood	B2	Moderate	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HMT372	Urine	B2	Moderate	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓				▲
HMT378	Urine	B2	Moderate	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓				
HMT379	Blood	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HMT385	Urine	B2	Weak	✓		✓	✓		✓		✓				✓	✓				▲
HMT386	Urine	B2	Strong	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓				
HUB605	Blood	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HMT389	Urine	B2	Weak	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓				
HMT396	Urine	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HMT399	Urine	B2	Moderate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HUB610	Blood	B2	Weak	✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓				
HUB647	Urine	B2	Strong	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB207	Urine	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			▲	
HCB210	Blood	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HCB214	Blood	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HUB648	Urine	B2	Moderate	✓	✓		✓	✓	✓	✓	✓	✓	✓		✓	✓				
HCB218	Urine	B2	Moderate	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HUB649	Urine	B2	Strong	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HUB650	Urine	B2	Strong	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HUB652	Urine	B2	Strong	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HUB653	Urine	B2	Moderate	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
HUB655	Urine	B2	Strong	✓		✓	✓	✓	✓	✓	✓		✓	✓	✓	✓				

(Continued)

TABLE 3 | (Continued)

Strain	Origin	Phylogroup	Biofilm	Adhesines				Siderophores					Toxins				ARGs			
				fimH	*sfa*	*papA*	*pgaA*	*iroN*	*iutA*	*irp1*	*iucA*	CAS test	*hlyA*	*cnf1*	*clbA*	*clbQ*	*bla*CTX-M-15	*bla*CTX-M-14	*sul1*	*sul2*
HUB665	Urine	B2	Moderate	■	■	■	■		■	■		■	■	■	■	■				
HUB666	Urine	B2	Moderate	■	■	■	■		■	■	■		■	■	■	■				
HCB221	Respiratory	B2	Moderate	■	■		■		■	■		■	■	■	■	■				
HCB224	Blood	B2	Strong	■	■	■	■		■	■	■	■	■	■	■	■				
HCB226	Blood	B2	Strong	■		■	■		■	■	■	■	■	■	■	■				
HCB227	Blood	B2	Moderate	■	■		■		■	■		■		■	■	■				
HCB231	Urine	UN	Weak	■		■	■		■	■			■	■	■	■				
HCB232	Urine	B2	Non producer	■		■	■		■	■			■	■	■	■			▣	
HCB233	Urine	B2	Weak	■		■	■		■	■		■	■	■	■	■				
HMT400	Urine	B2	Weak	■		■	■		■	■		■	■	■	■	■				▣
HMT410	Urine	B2	Moderate	■		■	■		■	■		■	■	■	■	■		▣		
HMT415	Urine	B2	Strong	■		■	■		■	■		■	■	■	■	■				
HMT418	Blood	B2	Weak	■		■	■		■	■		■	■	■	■	■				
HMT426	Urine	B2	Moderate	■	■	■	■		■	■		■	■	■	■	■				▣
HUB809	Respiratory	B2	Strong	■		■	■		■	■	■	■	■	■	■	■				▣
HUB810	Respiratory	B2	Weak	■		■	■		■	■	■	■		■	■	■				
HCB477	Urine	B2	Moderate	■		■	■		■	■		■	■	■	■	■				
HCB476	Respiratory	B2	Moderate	■		■	■		■	■		■	■	■	■	■				
HCB475	Respiratory	B2	Weak	■		■	■		■	■		■	■	■	■	■				
HCB473	Urine	B2	Weak	■		■	■	■	■	■		■		■	■	■				
HCB469	Blood	B2	Moderate	■		■	■		■	■		■	■	■	■	■				
HCB463	Urine	B2	Strong	■		■	■		■	■		■	■	■	■	■				
HCB457	Urine	B2	Weak	■		■	■		■	■		■		■	■	■				
HCB458	Blood	B2	Moderate	■		■	■		■	■		■	■	■	■	■				
HCB455	Urine	B2	Moderate	■	■	■	■		■	■	■	■	■	■	■	■				

Colored squares denote the presence of the gene or positivity of the test.
ARGs: antimicrobial resistance genes. UN: Unknown.

found similar percentages of fully susceptible strains but a higher rate of MDR in the collection under study. They also reported the highest average resistance percentage for aminopenicillins (58.7%), followed by fluoroquinolones (25.7%). These differences could be probably due to most of the strains in the present study were isolated from urine, and some of the most common antibiotics prescribed in UTIs are ciprofloxacin and trimethoprim-sulfamethoxazole. Previous studies conducted by our group showed no statistically significant relationship between biofilm formation ability and ciprofloxacin resistance among *E. coli* strains, but this association was found in *Pseudomonas aeruginosa* (Cepas et al., 2019). However, after analyzing whether this association existed in our study, a statistically significant association was found between non-biofilm-forming strains and resistance to ciprofloxacin. These variations among the data could be due to the greater number of strains analyzed in the present research. Several studies reported that quinolone-resistant strains were less virulent than quinolone-susceptible strains, observing that fimbriae genes, associated with biofilm formation, were less prevalent among nalidixic acid-resistant *E. coli* isolates (Vila et al., 2002; Horcajada et al., 2005). Moreover,

the study by Fàbrega et al. (2014) showed that the acquisition of quinolone resistance was related to a decrease in biofilm formation in *Salmonella enterica* strains (Fàbrega et al., 2014) which may explain our results.

Regarding ESBL, these were highly prevalent in our study (19.4%) compared to the study conducted by Flament-Simon et al. (2020) in which only 13 of the strains (6.6%) isolated in Spain and France in 2016 produced ESBL enzymes (Flament-Simon et al., 2020). In agreement with our results, several epidemiological studies confirm that *bla*CTX-M-15 is one the most common ESBL among clinical isolates from Spain. However, *bla*CTX-M-14 and *bla*CTX-M-27 have also been found in this country (Dahbi et al., 2014; Merino et al., 2016, 2018). Interestingly, three strains carrying *bla*CTX-M-28 were found in our study, which is less prevalent than other *bla*CTX-M-1 enzymes (Livermore et al., 2007). When we analyzed the relation between biofilm formation and ESBL production, we did not find any correlation between these two variables. However, Shrestha et al. (2019) found a positive correlation between these two features (Shrestha et al., 2019). The differences between the two studies could be due to the higher percentage of beta-lactamases (50.9% vs. 19.4%) they
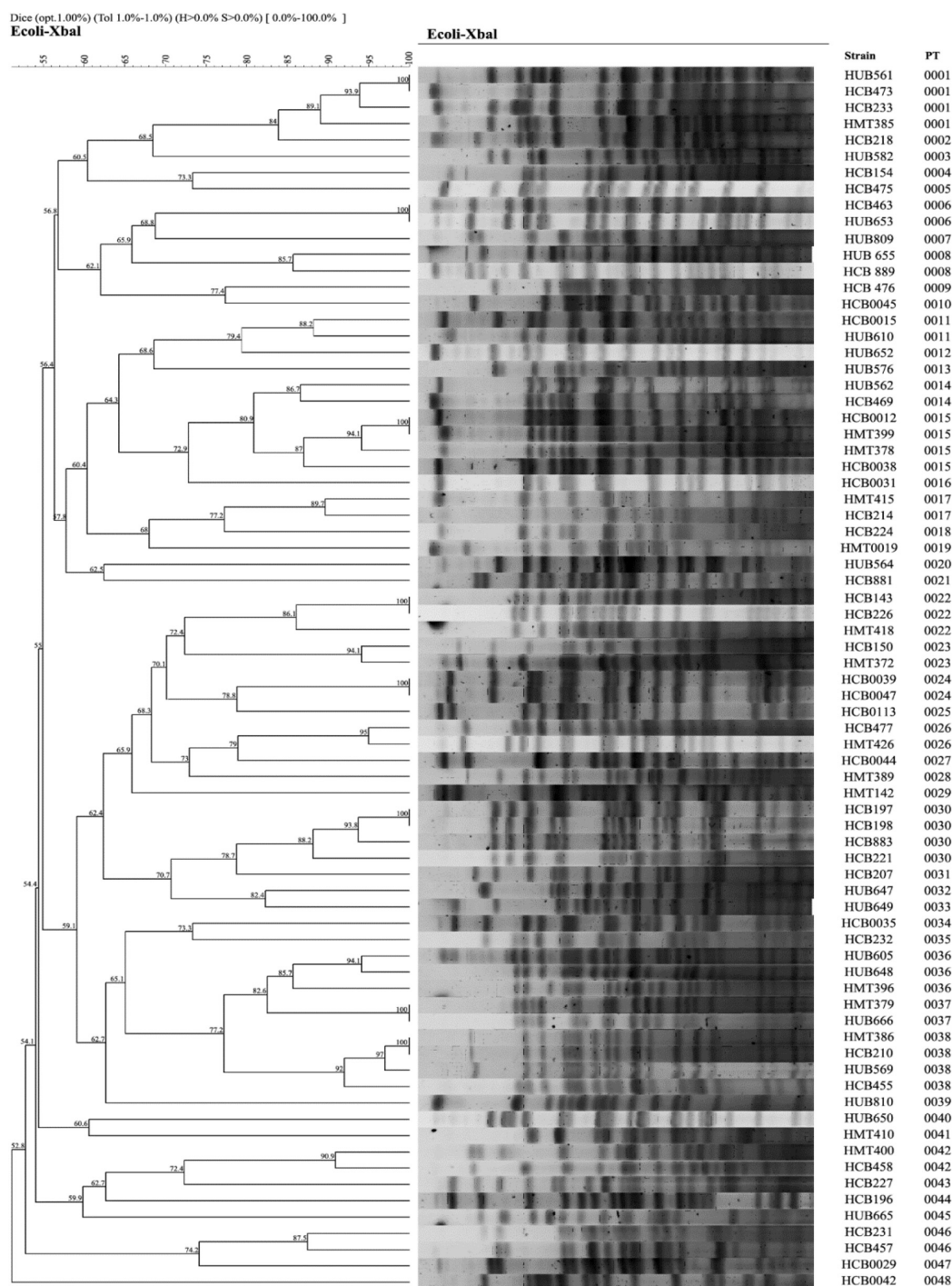
FIGURE 7 | Clonal relationship analysis of the colibactin strains assessed by PFGE using *Xba*I. PT: Pulsotype.

found. These results showing a worrying situation, since having a high percentage of ESBL among biofilm-forming strains makes very difficult their eradication.

As for virulence, we found 14 (3.7%) strains carrying all the VFGs tested. However, most were associated with low antimicrobial resistance rates. We observed a statistically significant correlation between the presence of different siderophores, adhesins or toxins and the ability to form biofilms. Siderophores are small molecules with high affinity for iron (Holden and Bachman, 2015). They are considered an important virulence factor in most Gram-negative bacteria. In *S. maltophilia*, iron plays a significant role in biofilm formation

and production of EPS (Kalidasan et al., 2018). In *P. aeruginosa*, siderophore-deficient clones showed reduced biofilm formation ability (Banin et al., 2005). In *E. coli*, studies by May and Okabe (2011) have shown that biofilm formation is favored in media with low iron concentrations (May and Okabe, 2011). In support of previous studies, we found a significant correlation between the presence of siderophore-encoding genes, such as *iroN*, and the ability to form biofilms. In addition, resistant strains produced more siderophores than their susceptible counterparts. Recently, it has been reported that the introduction of the mobile genetic element ICEKp from *Klebsiella pneumoniae* or a plasmid encoding YbtPQ (a siderophore importer encoded in the yersiniabactin cluster) reduced the susceptibility of *E. coli* to a wide range of antimicrobials (Farzand et al., 2021). Regarding adhesins, type 1 fimbriae are the best known adhesive organelles among Enterobacteriaceae (Abdelhalim et al., 2020) playing an important role in the initial steps of biofilm formation. As expected, we found a significant association between the presence of the *fimH* gene and biofilm formation. The adhesins *sfa* and *papA* also showed a significant association with biofilm formation. Although poly-β-1,6-N-acetyl-D-glucosamine polymer is associated with biofilm formation by mediating cell-to-cell adhesion and attachment to surfaces (Sharma et al., 2016), our results showed no statistical correlation between the presence of the *pgaA* gene and the capacity to form biofilms.

An important finding of the present work is the high presence of the toxin colibactin among the strains under study. Colibactin is a virulence determinant and a genotoxic enzyme synthesized by polyketide synthases and encoded by a 54-kb genomic island designated *pks* (Faïs et al., 2018). Colibactin may induce DNA damage in the host and has been correlated with colorectal cancer in humans (Wernke et al., 2020). Previous studies have confirmed the presence of *pks*-positive strains among ExPEC (Auvray et al., 2021). Interestingly, 75 of our isolates (19.9%) were colibactin-encoding strains, a higher percentage than this observed by Suresh et al. (2018). Analysis of the association between colibactin and other VFGs showed a significant correlation of this toxin with the siderophore-encoding genes *iroN* and *irp1;* the adhesin-encoding genes *papA, fimH, pgaA*, and *sfa*; and the toxin genes *cnf-1* and *hlyA*. Iron is an essential element for the survival of *E. coli*. Previous studies confirm that the *pks* island is involved in the synthesis of siderophores such as yersiniabactin, enterobactin, and salmochelins via phosphopantetheinyl transferase ClbA (Martin et al., 2013). In addition, it has been suggested that *E. coli* strains carrying hemolysin and colibactin have advantages in colorectal colonization and subsequent cancer (Yoshikawa et al., 2020). As in our case, it has been previously reported that phylogenetic group B2 is predominant among colibactin-encoding *E. coli* strains (Sarshar et al., 2017). However, the high clonal diversity of PFGE analysis among colibactin-positive *E. coli* isolates studied suggests that colibactin may be distributed in a wide variety of strains and clones. Worryingly, Putze et al. (2009) detected colibactin-related genes associated with an ICE-like element in several enterobacteria, not only in *E. coli*, paving the way for the spread of this gene

cluster among species (Putze et al., 2009). Furthermore, we observed low rates of resistance to antimicrobial agents among colibactin-positive isolates, which is consistent with previous studies (Sarshar et al., 2017; Morgan et al., 2019). Although colibactin-encoding genes may have a direct influence on other biosynthetic pathways, our findings support the hypothesis that the presence of colibactin-encoding genes may be related to biofilm formation. In this sense, several studies showed a high capacity for biofilm formation among the *pks*-carrying strains studied (Raisch et al., 2014; Suresh et al., 2018). Likewise, Dejea et al. (2014) conducted a study in two geographically distinct cohorts from the United States and Malaysia finding that 89% of cases of right-sided colorectal tumors presented biofilms in the biopsies samples, concluding that a significant association between the presence of biofilms and colorectal cancer exists (Dejea et al., 2014). In another study of patients with familial adenomatous polyposis, the bacterial biofilms found were composed predominantly by *E. coli* and *Bacteroides fragilis*. Genes encoding colibactin (*clbB*) were found at high levels in strains from these patients compared to strains isolated from healthy individuals (Dejea et al., 2018). Additionally, uropathogenic *E. coli* strains carrying the *pks* island have also been described. Chagneau et al. (2021) detected colibactin in 24.7% of urine samples from patients with community-acquired pyelonephritis, cystitis or asymptomatic bacteriuria (Chagneau et al., 2021). They also reported that colibactin was produced during UTIs and induced DNA damage in urothelial cells (Chagneau et al., 2021). Similarly, the study conducted by Morgan and collaborators in an *in vivo* model of ascending UTI in rats, showed that isolates carrying the *cnf-1*, *clbA* and *clbQ* genes induced severe UTIs within 48–72 h (Morgan et al., 2019). Thus, strains carrying colibactin-encoding genes can produce biofilms and cause severe disease not only in the intestine but also in the urinary tract, and probably in other less studied anatomic sites.

CONCLUSION

Ciprofloxacin resistance was associated with lower biofilm production ability among the *E. coli* strains under study. Furthermore, biofilm formation ability was correlated with the presence of some siderophores, adhesins, and toxins. The high percentage of strains harboring the colibactin-encoding genes and the significant correlation between this toxin and biofilm formation, suggest that colibactin could be a promising therapeutic target to prevent biofilm formation. Nevertheless, further studies are needed to confirm our hypothesis and to better understand the impact of colibactin production on human health.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

Abdelhalim, K. A., Uzel, A., and Unal, N. G. (2020). The role of major virulence factors and pathogenicity of adherent-invasive *Escherichia coli* in patients with Crohn's disease. *Prz. Gastroenterol.* 15, 279–288. doi: 10.5114/pg.2020.93235

Auvray, F., Perrat, A., Arimizu, Y., Chagneau, C. V., Bossuet-Greif, N., Massip, C., et al. (2021). Insights into the acquisition of the pks island and production of colibactin in the *Escherichia coli* population. *Microb. Genomics* 7:579. doi: 10.1099/mgen.0.000579

Banin, E., Vasil, M. L., and Greenberg, E. P. (2005). Iron and *Pseudomonas aeruginosa* biofilm formation. *Proc. Natl. Acad. Sci. USA* 102, 11076–11081. doi: 10.1073/pnas.0504266102

Beloin, C., Roux, A., and Ghigo, J.-M. (2008). *Escherichia coli* biofilms. *Curr. Top. Microbiol. Immunol.* 322, 249–289. doi: 10.1007/978-3-540-75418-3_12

Beloin, C., Valle, J., Latour-Lambert, P., Faure, P., Kzreminski, M., Balestrino, D., et al. (2004). Global impact of mature biofilm lifestyle on *Escherichia coli* K-12 gene expression. *Mol. Microbiol.* 51, 659–674. doi: 10.1046/j.1365-2958.2003.03865.x

Blount, Z. D. (2015). The unexhausted potential of E. coli. *Elife* 4, 1–12. doi: 10.7554/eLife.05826

Cantón, R., Loza, E., Arcay, R. M., Cercenado, E., Castillo, F. J., Cisterna, R., et al. (2021). Antimicrobial activity of ceftolozane-tazobactam against Enterobacterales and *Pseudomonas aeruginosa* recovered during the Study for Monitoring Antimicrobial Resistance Trends (SMART) program in Spain (2016-2018). *Rev. Esp. Quimioter.* 34, 228–237. doi: 10.37201/req/019.2021

Cepas, V., Ballén, V., Gabasa, Y., Ramírez, M., López, Y., and Soto, S. M. (2020). Transposon insertion in the *purL* gene induces biofilm depletion in *Escherichia coli* ATCC 25922. *Pathogens* 9, 1–22. doi: 10.3390/pathogens9090774

Cepas, V., López, Y., Muñoz, E., Rolo, D., Ardanuy, C., Martí, S., et al. (2019). Relationship between biofilm formation and antimicrobial resistance in Gram-negative bacteria. *Microb. Drug Resist.* 25, 72–79. doi: 10.1089/mdr.2018.0027

Chagneau, C. V., Massip, C., Bossuet-Greif, N., Fremez, C., Motta, J. P., Shima, A., et al. (2021). Uropathogenic *E. coli* induces DNA damage in the bladder. *PLoS Pathog.* 17:1–18. doi: 10.1371/JOURNAL.PPAT.1009310

Clermont, O., Christenson, J. K., Denamur, E., and Gordon, D. M. (2013). The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep.* 5, 58–65. doi: 10.1111/1758-2229.12019

CLSI (2020). *CLSI. Performance Standards for Antimicrobial Susceptibility Testing, 30th Edition. CLSI Supplement M100*. Wayne, PA: Clinical and Laboratory Standard Institute.

Dahbi, G., Mora, A., Mamani, R., Lopez, C., Alonso, M. P., Marzoa, J., et al. (2014). Molecular epidemiology and virulence of *Escherichia coli* O16: H5-ST131: Comparison with H30 and H30-Rx subclones of O25b: H4-ST131. *Int. J. Med. Microbiol.* 304, 1247–1257. doi: 10.1016/j.ijmm.2014.10.002

Dejea, C., Fathi, P., Craig, J. M., Boleij, A., Taddese, R., Geis, A. L., et al. (2018). Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* 359, 592–597. doi: 10.1126/science.aah3648.Patients

Dejea, C. M., Wick, E. C., Hechenbleikner, E. M., White, J. R., Mark Welch, J. L., Rossettid, B. J., et al. (2014). Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. USA* 111, 18321–18326. doi: 10.1073/pnas.1406199111

Dubois, D., Delmas, J., Cady, A., Robin, F., Sivignon, A., Oswald, E., et al. (2010). Cyclomodulins in urosepsis strains of *Escherichia coli*. *J. Clin. Microbiol.* 48, 2122–2129. doi: 10.1128/JCM.02365-09

Durmaz, R., Otlu, B., Koksal, F., Hosoglu, S., Ozturk, R., Ersoy, Y., et al. (2009). The optimization of a rapid pulsed-field gel electrophoresis protocol for the typing of *Acinetobacter baumannii*, *Escherichia coli* and Klebsiella spp. *Jpn. J. Infect. Dis.* 62, 372–377.

Dziubańska-Kusibab, P. J., Berger, H., Battistini, F., Bouwman, B. A. M., Iftekhar, A., Katainen, R., et al. (2020). Colibactin DNA-damage signature indicates mutational impact in colorectal cancer. *Nat. Med.* 26, 1063–1069. doi: 10.1038/s41591-020-0908-2

European Centre for Disease Prevention and Control (ECDC) (2017). *Antimicrobial resistance surveillance in Europe 2016. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)*. Solna Municipality: ECDC.

Fàbrega, A., Soto, S. M., Ballesté-Delpierre, C., Fernández-Orth, D., Jiménez de Anta, M. T., and Vila, J. (2014). Impact of quinolone-resistance acquisition on biofilm production and fitness in *Salmonella enterica*. *J. Antimicrob. Chemother.* 69, 1815–1824. doi: 10.1093/jac/dku078

Faïs, T., Delmas, J., Barnich, N., Bonnet, R., and Dalmasso, G. (2018). Colibactin: More than a new bacterial toxin. *Toxins* 10, 16–18. doi: 10.3390/toxins10040151

Farzand, R., Rajakumar, K., Barer, M. R., Freestone, P. P. E., Mukamolova, G. V., Oggioni, M. R., et al. (2021). A virulence associated siderophore importer reduces antimicrobial susceptibility of *Klebsiella pneumoniae*. *Front. Microbiol.* 12:1–9. doi: 10.3389/fmicb.2021.607512

Flament-Simon, S. C., Nicolas-Chanoine, M. H., García, V., Duprilot, M., Mayer, N., Alonso, M. P., et al. (2020). Clonal structure, virulence factor-encoding genes and antibiotic resistance of *Escherichia coli*, causing urinary tract infections and other extraintestinal infections in humans in spain and france during 2016. *Antibiotics* 9, 1–17. doi: 10.3390/antibiotics9040161

Friedlander, R. S., Vogel, N., and Aizenberg, J. (2015). Role of flagella in adhesion of *Escherichia coli* to abiotic surfaces. *Langmuir* 31, 6137–6144. doi: 10.1021/acs.langmuir.5b00815

Guerra, B., Junker, E., Miko, A., Helmuth, R., and Mendoza, M. C. (2004). Characterization and localization of drug resistance determinants in multidrug-resistant, integron-carrying *Salmonella enterica* Serotype Typhimurium strains. *Microb. Drug Resist.* 10, 83–91.

Guerrieri, C. G., Pereira, M. F., Galdino, A. C. M., Dos Santos, A. L. S., Elias, W. P., Schueuck, R. P., et al. (2019). Typical and atypical enteroaggregative *Escherichia coli* are both virulent in the *Galleria mellonella* model. *Front. Microbiol.* 10:1791. doi: 10.3389/fmicb.2019.01791

Guttenplan, S. B., and Kearns, D. B. (2013). Regulation of flagellar motility during biofilm formation. *FEMS Microbiol. Rev.* 37, 849–871. doi: 10.1111/1574-6976.12018

Holden, V. I., and Bachman, M. A. (2015). Diverging roles of bacterial siderophores during infection. *Metallomics* 7, 986–995. doi: 10.1039/c4mt00333k

Horcajada, J. P., Soto, S., Gajewski, A., Smithson, A., Jiménez De Anta, M. T., Mensa, J., et al. (2005). Quinolone-resistant uropathogenic *Escherichia coli* strains from phylogenetic group B2 have fewer virulence factors than their susceptible counterparts. *J. Clin. Microbiol.* 43, 2962–2964. doi: 10.1128/JCM.43.6.2962-2964.2005

Houdouin, V., Bonacorsi, S., Bidet, P., Bingen-Bidois, M., Barraud, D., and Bingen, E. (2006). Phylogenetic background and carriage of pathogenicity island-like domains in relation to antibiotic resistance profiles among *Escherichia coli* urosepsis isolates. *J. Antimicrob. Chemother.* 58, 748–751. doi: 10.1093/jac/dkl326

Johnson, J. R., and Stell, A. L. (2000). Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J. Infect. Dis.* 181, 261–272. doi: 10.1086/315217

Johnson, J. R., Johnston, B., Kuskowski, M. A., Nougayrede, J. P., and Oswald, E. (2008). Molecular epidemiology and phylogenetic distribution of the *Escherichia coli pks* genomic island. *J. Clin. Microbiol.* 46, 3906–3911. doi: 10.1128/JCM.00949-08

Johnson, J. R., Russo, T. A., Tarr, P. I., Carlino, U., Bilge, S. S., Vary, J. C., et al. (2000). Molecular epidemiological and phylogenetic associations of two novel putative virulence genes, *iha* and *iroN*$_{E. coli}$, among *Escherichia coli* isolates from patients with urosepsis. *Infect. Immun.* 68, 3040–3047. doi: 10.1128/IAI.68.5.3040-3047.2000

Kalidasan, V., Joseph, N., Kumar, S., Awang Hamat, R., and Neela, V. K. (2018). Iron and virulence in *Stenotrophomonas maltophilia*: All We Know So Far. *Front. Cell Infect. Microbiol.* 8:401. doi: 10.3389/fcimb.2018.00401

Leflon-Guibout, V., Jurand, C., Bonacorsi, S., Espinasse, F., Guelfi, M. C., Duportail, F., et al. (2004). Emergence and spread, of three clonally related virulent isolates of CTX-M-15-producing *Escherichia coli* with variable resistance to aminoglycosides and tetracycline in a French geriatric hospital. *Antimicrob. Agents Chemother.* 48, 3736–3742. doi: 10.1128/AAC.48.10.3736-3742.2004

Lescat, M., Poirel, L., and Nordmann, P. (2018). Rapid multiplex polymerase chain reaction for detection of *mcr-1* to *mcr-5* genes. *Diagn. Microbiol. Infect. Dis.* 92, 267–269. doi: 10.1016/j.diagmicrobio.2018.04.010

Livermore, D. M., Canton, R., Gniadkowski, M., Nordmann, P., Rossolini, G. M., Arlet, G., et al. (2007). CTX-M: Changing the face of ESBLs in Europe. *J. Antimicrob. Chemother.* 59, 165–174. doi: 10.1093/jac/dkl483

Magiorakos, A. P., Srinivasan, A., Carey, R. B., Carmeli, Y., Falagas, M. E., Giske, C. G., et al. (2012). Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: An international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* 18, 268–281. doi: 10.1111/j.1469-0691.2011.03570.x

Manges, A. R., Geum, H. M., Guo, A., Edens, T. J., Fibke, C. D., and Pitout, J. D. D. (2019). Global extraintestinal pathogenic *Escherichia coli* (ExPEC) lineages. *Clin. Microbiol. Rev.* 32, 1–25.

Martin, P., Marcq, I., Magistro, G., Penary, M., Garcie, C., Payros, D., et al. (2013). Interplay between siderophores and colibactin genotoxin biosynthetic pathways in *Escherichia coli*. *PLoS Pathog.* 9:3437. doi: 10.1371/JOURNAL.PPAT.1003437

May, T., and Okabe, S. (2011). Enterobactin is required for biofilm development in reduced-genome *Escherichia coli*. *Environ. Microbiol.* 13, 3149–3162. doi: 10.1111/j.1462-2920.2011.02607.x

Merino, I., Hernández-García, M., Turrientes, M. C., Pérez-Viso, B., López-Fresneña, N., Diaz-Agero, C., et al. (2018). Emergence of ESBL-producing *Escherichia coli* ST131-C1-M27 clade colonizing patients in Europe. *J. Antimicrob. Chemother.* 73, 2973–2980. doi: 10.1093/jac/dky296

Merino, I., Shaw, E., Horcajada, J. P., Cercenado, E., Mirelis, B., Pallarés, M. A., et al. (2016). CTX-M-15-H30Rx-ST131 subclone is one of the main causes of healthcare-associated ESBL-producing *Escherichia coli* bacteraemia of urinary origin in Spain. *J. Antimicrob. Chemother.* 71, 2125–2130. doi: 10.1093/jac/dkw133

Monstein, H. J., Ostholm-Balkhed, A., Nilsson, M. V., Nilsson, M., Dornbusch, K., and Nilsson, L. E. (2007). Multiplex PCR amplification assay for the detection of *bla*$_{SHV}$, *bla*$_{TEM}$ and *bla*$_{CTX-M}$ genes in *Enterobacteriaceae*. *Apmis* 115, 1400–1408. doi: 10.1111/j.1600-0463.2007.00722.x

Morgan, R. N., Saleh, S. E., Farrag, H. A., and Aboulwafa, M. M. (2019). Prevalence and pathologic effects of colibactin and cytotoxic necrotizing factor-1 (Cnf 1) in *Escherichia coli*: Experimental and bioinformatics analyses. *Gut. Pathog.* 11, 1–18. doi: 10.1186/s13099-019-0304-y

Pai, H., Choi, E. H., Lee, H. J., JungYun, H., and Jacoby, G. A. (2001). Identification of CTX-M-14 extended-spectrum β-lactamase in clinical isolates of *Shigella*

sonnei. *Escherichia coli*, and *Klebsiella pneumoniae* in Korea. *J. Clin. Microbiol.* 39, 3747–3749. doi: 10.1128/JCM.39.10.3747-3749.2001

Park, C. H., Robicsek, A., Jacoby, G. A., Sahm, D., and Hooper, D. C. (2006). Prevalence in the United States of *aac(6′)-Ib-cr* encoding a ciprofloxacin-modifying enzyme. *Antimicrob. Agents Chemother.* 50, 3953–3955. doi: 10.1128/AAC.00915-06

Putze, J., Hennequin, C., Nougayrède, J. P., Zhang, W., Homburg, S., Karch, H., et al. (2009). Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect. Immun.* 77, 4696–4703.

Raisch, J., Buc, E., Bonnet, M., Sauvanet, P., Vazeille, E., de Vallée, A., et al. (2014). Colon cancer-associated B2 *Escherichia coli* colonize gut mucosa and promote cell proliferation. *World J. Gastroenterol.* 20, 6560–6572. doi: 10.3748/wjg.v20.i21.6560

Robicsek, A., Strahilevitz, J., Jacoby, G. A., Macielag, M., Abbanat, D., Chi, H. P., et al. (2006). Fluoroquinolone-modifying enzyme: a new adaptation of a common aminoglycoside acetyltransferase. *Nat. Med.* 12, 83–88. doi: 10.1038/nm1347

Sáez-López, E., Bosch, J., Salvia, M. D., Fernández-Orth, D., Cepas, V., Ferrer-Navarro, M., et al. (2017). Outbreak caused by *Escherichia coli* O18:K1:H7 sequence type 95 in a neonatal intensive care unit in Barcelona, Spain. *Pediatr. Infect. Dis. J.* 36, 1079–1086. doi: 10.1097/INF.0000000000001652

Sarowska, J., Futoma-Koloch, B., Jama-Kmiecik, A., Frej-Madrzak, M., Ksiazczyk, M., Bugla-Ploskonska, G., et al. (2019). Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: Recent reports. *Gut. Pathog.* 11, 1–16. doi: 10.1186/s13099-019-0290-0

Sarshar, M., Scribano, D., Marazzato, M., Ambrosi, C., Aprea, M. R., Aleandri, M., et al. (2017). Genetic diversity, phylogroup distribution and virulence gene profile of *pks* positive *Escherichia coli* colonizing human intestinal polyps. *Microb. Pathog.* 112, 274–278. doi: 10.1016/j.micpath.2017.10.009

Schwyn, B., and Neilands, J. B. (1987). Universal chemical assay for the detection and determination of siderophores. *Anal. Biochem.* 160, 47–56. doi: 10.1016/0003-2697(87)90612-9

Sharma, G., Sharma, S., Sharma, P., Chandola, D., Dang, S., Gupta, S., et al. (2016). *Escherichia coli* biofilm: development and therapeutic strategies. *J. Appl. Microbiol.* 121, 309–319. doi: 10.1111/jam.13078

Shrestha, R., Khanal, S., Poudel, P., Khadayat, K., Ghaju, S., Bhandari, A., et al. (2019). Extended spectrum β-lactamase producing uropathogenic *Escherichia coli* and the correlation of biofilm with antibiotics resistance in Nepal. *Ann. Clin. Microbiol. Antimicrob.* 18, 1–6. doi: 10.1186/s12941-019-0340-y

Smith, D. R., Price, J. E., Burby, P. E., Blanco, L. P., Chamberlain, J., and Chapman, M. R. (2017). The production of curli amyloid fibers is deeply integrated into the biology of *Escherichia coli*. *Biomolecules* 7:75. doi: 10.3390/biom7040075

Stepanović, S., Vuković, D., Hola, V., Di Bonaventura, G., Djukić, S., Ćirković, I., et al. (2007). Quantification of biofilm in microtiter plates: Overview of testing conditions and practical recommendations for assessment of biofilm production by staphylococci. *Apmis* 115, 891–899. doi: 10.1111/j.1600-0463.2007.apm_630.x

Suresh, A., Ranjan, A., Jadhav, S., Hussain, A., Shaik, S., Alam, M., et al. (2018). Molecular genetic and functional analysis of *pks*-harboring, extra-intestinal pathogenic *Escherichia coli* from India. *Front. Microbiol.* 9:1–8. doi: 10.3389/fmicb.2018.02631

Takahashi, A., Kanamaru, S., Kurazono, H., Kunishima, Y., Tsukamoto, T., Ogawa, O., et al. (2006). *Escherichia coli* isolates associated with uncomplicated and complicated cystitis and asymptomatic bacteriuria possess similar phylogenies, virulence genes, and O-serogroup profiles. *J. Clin. Microbiol.* 44, 4589–4592. doi: 10.1128/JCM.02070-06

Van Gerven, N., Van der Verren, S. E., Reiter, D. M., and Remaut, H. (2018). The role of functional amyloids in bacterial virulence. *J. Mol. Biol.* 430, 3657–3684. doi: 10.1016/j.jmb.2018.07.010

Vila, J., Simon, K., Ruiz, J., Horcajada, J. P., Velasco, M., Barranco, M., et al. (2002). Are quinolone-resistant uropathogenic *Escherichia coli* less virulent? *J. Infect. Dis.* 186, 1039–1042.

Wernke, K. M., Xue, M., Tirla, A., Kim, C. S., Crawford, J. M., and Herzon, S. B. (2020). Structure and bioactivity of colibactin. *Bioorganic Med. Chem. Lett.* 30:127280. doi: 10.1016/j.bmcl.2020.127280

Wiegand, I., Geiss, H. K., Mack, D., Stu, E., Seifert, H., and Icrobiol, J. C. L. I. N. M. (2007). Detection of Extended-Spectrum Beta-Lactamases among *Enterobacteriaceae* by use of semiautomated microbiology systems and manual detection procedures. *J. Clin. Microbiol.* 45, 1167–1174. doi: 10.1128/ JCM.01988-06

Woodford, N., Fagan, E. J., and Ellington, M. J. (2006). Multiplex PCR for rapid detection of genes encoding CTX-M extended-spectrum β-lactamases [4]. *J. Antimicrob. Chemother.* 57, 154–155. doi: 10.1093/jac/dki412

Xu, J. G., Cheng, B., Wen, X., Cui, S., and Ye, C. (2000). High-pathogenicity island of *Yersinia* spp. in *Escherichia coli* strains isolated from diarrhea patients in China. *J. Clin. Microbiol.* 38, 4672–4675. doi: 10.1128/jcm.38.12.4672-4675. 2000

Yoshikawa, Y., Tsunematsu, Y., Matsuzaki, N., Hirayama, Y., Higashiguchi, F., Sato, M., et al. (2020). Characterization of colibactin-producing *Escherichia coli* isolated from Japanese patients with colorectal cancer. *Japanese J. Infect. Dis.* 73, 437–442.

Ferric Iron Reduction in Extreme Acidophiles

Luise Malik and Sabrina Hedrich*

Research Group Biohydrometallurgy and Microbiology, Institute of Biosciences, TU Bergakademie Freiberg, Freiberg, Germany

Biochemical processes are a key element of natural cycles occurring in the environment and enabling life on earth. With regard to microbially catalyzed iron transformation, research predominantly has focused on iron oxidation in acidophiles, whereas iron reduction played a minor role. Microbial conversion of ferric to ferrous iron has however become more relevant in recent years. While there are several reviews on neutrophilic iron reducers, this article summarizes the research on extreme acidophilic iron reducers. After the first reports of dissimilatory iron reduction by acidophilic, chemolithoautotrophic *Acidithiobacillus* strains and heterotrophic *Acidiphilium* species, many other prokaryotes were shown to reduce iron as part of their metabolism. Still, little is known about the exact mechanisms of iron reduction in extreme acidophiles. Initially, hypotheses and postulations for the occurring mechanisms relied on observations of growth behavior or predictions based on the genome. By comparing genomes of well-studied neutrophilic with acidophilic iron reducers (e.g., *Ferroglobus placidus* and *Sulfolobus* spp.), it became clear that the electron transport for iron reduction proceeds differently in acidophiles. Moreover, transcriptomic investigations indicated an enzymatically-mediated process in *Acidithiobacillus ferrooxidans* using respiratory chain components of the iron oxidation in reverse. Depending on the strain of *At. ferrooxidans*, further mechanisms were postulated, e.g., indirect iron reduction by hydrogen sulfide, which may form by disproportionation of elemental sulfur. Alternative scenarios include Hip, a high potential iron-sulfur protein, and further cytochromes. Apart from the anaerobic iron reduction mechanisms, sulfur-oxidizing acidithiobacilli have been shown to mediate iron reduction at low pH (< 1.3) under aerobic conditions. This presumably non-enzymatic process may be attributed to intermediates formed during sulfur/tetrathionate and/or hydrogen oxidation and has already been successfully applied for the reductive bioleaching of laterites. The aim of this review is to provide an up-to-date overview on ferric iron reduction by acidophiles. The importance of this process in anaerobic habitats will be demonstrated as well as its potential for application.

Keywords: extreme acidophiles, iron reduction, biochemistry, biohydrometallurgy, reductive bioleaching

LIFESTYLE OF ACIDOPHILIC MICROORGANISMS

Acidic environments such as acid mine drainage are the result of anthropogenic influences or have a natural origin, e.g., solfataric springs (Johnson, 2009). Depending on the occurring pH, these environments are habitats for moderate (pH 3–5) or extreme acidophilic (pH < 3) organisms. Representatives of all three domains have been identified under acidic conditions, also eukaryotes, mainly fungi and algae, some rotifers, ciliates and species of the genus *Euglena*, were determined in acidic waterbodies (Quatrini and Johnson, 2016; Stolz, 2017; Johnson and Quatrini, 2020). However, this article will focus on acidophilic prokaryotes, which are well adapted to the challenging conditions of extremely low pH habitats (Dopson, 2016; Golyshina et al., 2016).

Adaptation to Extreme Environments

Most of the challenges acidophiles have to face with respect to their metabolism and morphological adaptations, are dictated by the conditions in their extreme environment. Low pH values are characterized by high concentrations of hydronium ions caused by a combination of proton-producing biotic and abiotic reactions, for instance, the generation of sulfuric acid in mine-impacted and geothermal areas (Johnson, 2009). Acidophiles have evolved appropriate mechanisms to overcome this cell homeostasis-threatening issue, e.g., the presence of a reverse positive inner membrane potential (Matin, 1999). With respect to highly acidic habitats, high temperatures are also an important influence, as these two quite hostile conditions occasionally appear together, for example, in hot solfatara springs or volcanoes (Johnson, 2009). Besides high concentrations of protons, a variety of metal cations is present in most acidic environments. These are generated by mineral leaching processes, which cause mobilization of partially toxic metal and metalloid cations (Gadd, 2009). The cations present, especially iron and copper, can in turn generate reactive oxygen species, e.g., through the Fenton reaction, which increases the oxidative stress and causes damage of cellular components (Ferrer et al., 2016). However, the inorganic compounds released during leaching processes are also a source of energy for many microorganisms. This is of particular importance since most acidic environments have low levels of organic carbon sources which forced the development of widespread abilities to utilize autotrophic metabolisms (Johnson, 2009).

Metabolic Diversity of Extreme Acidophiles

Despite the fact that acidophiles are to some extent constrained by their environmental conditions, diverse metabolisms have evolved that are used by these organisms (**Figure 1**; Johnson, 1998). Acidophiles can exhibit both obligate heterotrophic or obligate/facultative autotrophic carbon metabolisms, with the latter using oxidation or reduction of elemental sulfur, iron or hydrogen to generate ATP (Johnson and Hallberg, 2008; Johnson, 2009). Some acidophiles use a wide range of

dissimilation pathways, while the obligate aerobic autotrophic *Leptospirillum* is only capable of oxidizing ferrous iron (Dopson, 2016). Depending on the utilized electron acceptor and donor, the acidophilic electron transport systems differ considerably. Especially in the case of sulfur metabolism, the variety of occurring oxidation states (−2 to + 6) allow several pathways by which reduced inorganic sulfur compounds (RISC) are partially or completely oxidized to sulfate. Nevertheless, a lot of effort was put in the elucidation of the underlying enzymatic processes, as reviewed elsewhere (Wang et al., 2019; Dahl, 2020). In addition to the well-studied sulfur metabolism, hydrogen oxidation under anaerobic or aerobic conditions is also a common feature of acidophiles (Hedrich and Johnson, 2013b; Kucera et al., 2020).

Regarding the iron oxidation mechanism in acidophiles, different strategies evolved simultaneously in several phyla, e.g., γ-*Proteobacteria*, *Firmicutes*, *Nitrospira*, *Crenarchaeota*, and *Euryarchaeota* (Ilbert and Bonnefoy, 2013). In particular, for *Acidithiobacillus ferrooxidans*, bioinformatic, genetic, proteomic, and transcriptomic investigations of the electron transfer chain and related genes led to a widely accepted iron oxidation model (Ingledew, 1982; Bonnefoy, 2010; Amouric et al., 2011; Kucera et al., 2013; Jiang et al., 2021). This is still work in progress, as the results obtained on an ongoing basis extend the previously postulated iron oxidation mechanism to some extent, for example through newly discovered gene clusters (Ai et al., 2018). Nevertheless, the acknowledged electron transport chain from iron to the terminal electron acceptor includes an outer membrane cytochrome *c* (Cyc2), transferring the electron from iron to rusticyanin, a copper protein situated in the periplasm (Giudici-Orticoni et al., 1999; Yarzábal et al., 2002). Depending on its destination, the electron is further transported to the NADH-1 complex *via* cytochrome *c* (CycA1), the cytochrome *bc*₁ complex and ubiquinone pool to produce reduction equivalents (Elbehti et al., 1999, 2000). Electron transport along the thermodynamic gradient to reduce oxygen to water occurs, if the electron is transferred from rusticyanin to cytochrome *c* (Cyc1) and further to the *aa*₃ cytochrome oxidase (Kai et al., 1992; Malarte et al., 2005). Accordingly, the components of the iron oxidation chain in *Acidithiobacillus ferrooxidans* form a super-complex extending from the outer membrane integrated Cyc2 through the periplasm to the cytochrome oxidase, *bc*₁ complex, and NADH-1 complex in the inner membrane (**Figure 2**; Ilbert and Bonnefoy, 2013).

Many acidophiles are facultative anaerobes, preferentially using oxygen as electron acceptor, but in its absence, they may alternatively reduce iron or rarely sulfur (Gyure et al., 1990; Coupland and Johnson, 2008). Dissimilatory ferric iron reduction under anaerobic conditions is widespread among acidophiles and can be coupled to the oxidation of hydrogen, sulfur, RISC or organic compounds (Coupland and Johnson, 2008; Johnson, 2010). The greater iron solubility at low pH and its abundance in many acidic environments are reasons why iron is used by acidophiles as both electron acceptor and donor (Johnson, 2009). Additionally, the much higher redox potential of the ferrous/ferric iron pair at low pH (+ 663 mV, pH 2, sulfate-rich solution) (Johnson et al., 2017) makes it a more feasible electron acceptor compared to neutral conditions (+ 200 mV,
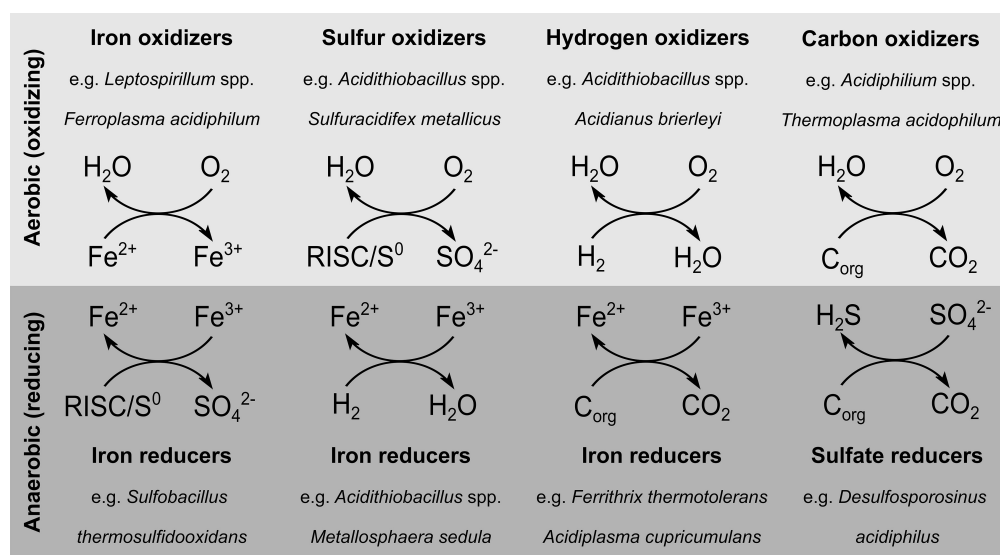
FIGURE 1 | Diversity of dissimilation pathways used by acidophiles. The underlying data can be found in the associated studies (Hedrich et al., 2011; Hedrich and Johnson, 2013b; Dopson, 2016). RISC, reduced inorganic sulfur compounds; C_{org}, organic carbon.

pH 7, bicarbonate-containing environments) (Ehrenreich and Widdel, 1994). Accordingly, microbial cycling of iron can occur between habitat niches of different dissolved oxygen levels, which contributes significantly to the natural iron cycle.

IRON REDUCTION BY ACIDOPHILES

Iron reduction as part of the natural iron cycle plays an important role on Earth (Kappler and Straub, 2005). Despite the well-studied mechanisms of iron reduction by neutrophilic microorganisms, the process in acidophilic microbes has not yet been sufficiently explored. Nevertheless, genome comparisons of known neutrophilic iron reducers with those of acidophiles revealed few similarities suggesting a distinct mechanism in acidophiles (Masaki et al., 2018). Depending on the considered microorganism, ferric iron reduction by acidophiles can take place under anaerobic, aerobic as well as micro-aerobic conditions. The most important milestones of ferric iron reduction research in acidophiles are summarized in **Table 1**. First investigations of acidithiobacilli described aerobic ferric iron reduction at low pH (Brock and Gustafson, 1976; Sand, 1989) and anaerobic ferrous iron formation (Brock and Gustafson, 1976). Micro-aerobic iron reduction was especially studied for the genus *Acidiphilium* (Johnson and Bridge, 2002). Regarding archaea, only a few selected genera have been studied profoundly with respect to their iron reduction mechanism, most notably the anaerobic and micro-aerobic studies of *Sulfolobus*, *Sulfuracidifex*, and *Saccharolobus* (latter no longer termed as extreme acidophile due to a pH optimum of 4.5) (Masaki et al., 2018). Moreover, investigations of the archeon "*Ferroplasma acidarmanus*" Fer1 revealed anaerobic ferric iron reduction rates which were comparable with those determined for *Acidiphilium* species (Dopson et al., 2007). Overall, ferric iron reduction

was shown for different physiologic groups of acidophilic microorganisms, covering heterotrophs and chemoautotrophs which feature temperature optima ranging from mesophilic to thermophilic adaptations (Johnson and Bridge, 2002). In particular, the ability of acidophiles to mediate reductive dissolution of ferric iron minerals is frequently explored, with emphasis on the applicability prior to elucidation of the exact biochemical mechanism (Hallberg et al., 2011a; Nancucheo et al., 2014; Norris et al., 2015; Marrero et al., 2017). However, certain research groups provided insight into the mechanisms of ferric iron reduction that occur in extreme acidophiles at different oxygen concentrations, which will be discussed in more detail.

Anaerobic Iron Reduction at Low pH

After the observation of iron-reducing processes by numerous neutrophilic prokaryotes (Bromfield, 1954; Ottow and Glathe, 1971; Lovley, 1991), anaerobic iron reduction by extreme acidophiles was first described in 1976, for the species *Acidithiobacillus ferrooxidans* (Brock and Gustafson, 1976). Under the given conditions, elemental sulfur-respiring *At. ferrooxidans* reduced iron exclusively in the absence of oxygen. Accordingly, it was hypothesized that ferric iron may serve as an alternative electron acceptor for acidophiles in anaerobic habitats. To date, iron reduction in the absence of oxygen has been demonstrated for a variety of moderate and extremely acidophilic microorganisms. The phylogenetically heterogenous group of iron-reducing acidophiles includes archaeal as well as bacterial species covering the phyla *Crenarchaeota*, *Euryarchaeota*, *Acidobacteria*, *Actinobacteria*, *Firmicutes*, and *Proteobacteria* (**Tables 2–5**).

The investigation of the ferric iron reduction mechanisms in acidophiles is still ongoing and only few organisms, e.g., *Acidithiobacillus ferrooxidans*, were examined in detail. Over the

FIGURE 2 | Components of the iron oxidation pathway in *Acidithiobacillus ferrooxidans*. The scheme is a modified version of that given by Ilbert and Bonnefoy (2013). PP, periplasm; CP, cytoplasm; OM, outer membrane; IM, inner membrane; Cyc, cytochrome *c*; Rus, rusticyanin; QP, quinol pool; NDHI, NADH dehydrogenase complex; *bc₁*, cytochrome *bc₁* complex (encoded by *petI/II*).
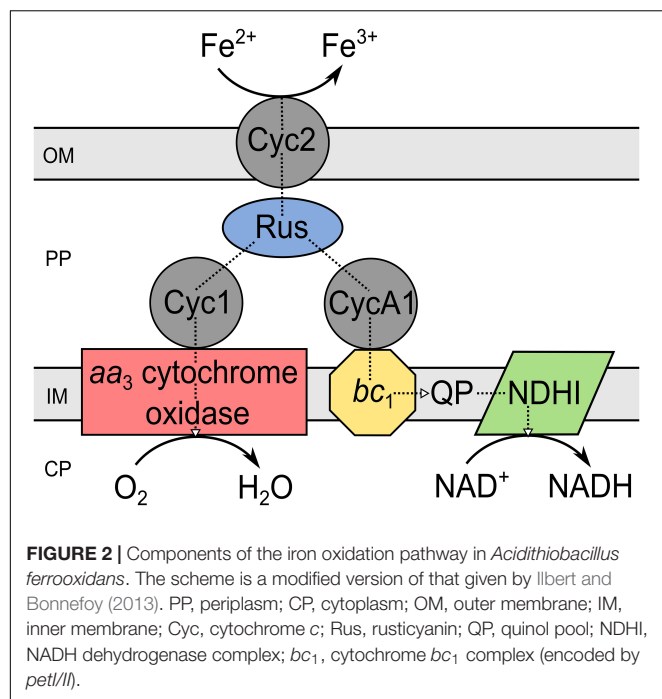
years, various hypothesis were given, for instance an involvement of a nitrate reductase or tetrathionate hydrolase (Johnson and McGinness, 1991; Sugio et al., 2009), which could not be confirmed by later studies (Kucera et al., 2012, 2016b; Osorio et al., 2013). In addition to the biochemical process, it was crucial to determine whether anaerobic iron reduction is accompanied by cell growth or merely a mechanism for maintaining cellular functions. Initially, no growth was detectable for anaerobic iron-reducing *At. ferrooxidans* (Sugio et al., 1988), which was later contributed to an insufficient carbon dioxide supply (Das et al., 1992). Soon, anaerobic dissimilatory iron reduction was reported for different *At. ferrooxidans* strains (Pronk et al., 1991, 1992; Das et al., 1992). In this context, a doubling time of 24 h was determined for ferric iron-respiring *At. ferrooxidans*[T] on elemental sulfur (Pronk et al., 1992) and the existence of two distinct electron transport systems for ferric iron reduction was suggested (Das et al., 1992). Additionally, the electron donor provided during pre-cultivation seemed to have a major impact on the ferric iron reduction ability of *At. ferrooxidans*. Cells grown on elemental sulfur for several generations were not able to reduce ferric iron when changed to anaerobic cultivation conditions, while the ones pre-cultivated with ferrous iron were (Kucera et al., 2012).

Uncovering Anaerobic Ferric Iron Reduction Mechanisms in Acidithiobacilli

All to date described extreme acidophilic iron-oxidizing acidithiobacilli have been identified as anaerobic ferric iron reducers (Pronk et al., 1992; Hallberg et al., 2010; Hedrich and Johnson, 2013a; Falagán and Johnson, 2016; Norris et al., 2020). Accordingly, this genus offers great potential for research into the biochemical mechanisms of iron reduction,

especially because *Acidithiobacillus ferrooxidans* has been studied in detail for its molecular processes as a flagship organism of bioleaching. Since the first investigation of anaerobic ferric iron reduction by *At. ferrooxidans* in the 1970s, a lot of effort has been put into the elucidation of the underlying enzymatic process. Sugio et al. (1985, 1987) purified and to some extent characterized a sulfur:ferric iron oxidoreductase from ferrous iron-grown *At. ferrooxidans* AP19-3. This periplasmic space enzyme catalyzed the oxidation of elemental sulfur to sulfite while reducing ferric to ferrous iron as one step of the sulfur metabolism and was therefore seen as an important part of the iron reduction chain (Sugio et al., 1985, 1987). In the same year, Corbett and Ingledew presented a model for the anaerobic iron respiration pathway of *At. ferrooxidans*[T] (**Figure 3**) utilizing complexes and electron transporters of the aerobic iron oxidation chain (section "Metabolic Diversity of Extreme Acidophiles"). Thereby, the electron transfer from the donor elemental sulfur to the terminal acceptor iron may involve the cytochrome *bc₁* complex, cytochrome *c*, rusticyanin, cytochrome oxidase, cytochrome *a₁*, and other transporters of the ferrous iron oxidation pathway located in the periplasmic space (Corbett and Ingledew, 1987). With this, the hypothesis arose that components of the iron oxidation chain could also be used in reverse for ferric iron respiration in acidophiles under anaerobic conditions, which was further investigated since then. Additional evidence was given by the observation, that *At. ferrooxidans* lacked the ability of anaerobic iron reduction after six consecutive aerobic cultivations on sulfur, which was contributed to the absence of some iron oxidation chain components in sulfur adapted cells (Kucera et al., 2012). Further, the investigation of anaerobic ferric iron respiration by *At. ferrooxidans* JCM 7811, later renamed as *At. ferriphilus* JCM 7811 (Kucera et al., 2020), with varying electron donors (hydrogen or elemental sulfur) resulted in the production of high amounts of a new c-type cytochrome (Ohmura et al., 2002). The purified cytochrome differed from other *At. ferrooxidans*-originating ones by an intermediate midpoint potential of about + 560 mV (pH 2.0) and was shown to act as an electron acceptor for ferric iron reduction (Ohmura et al., 2002). Consequently, the evidence was growing that cytochromes may play a major role in electron transfer during ferric iron reduction by sulfur-respiring acidophiles. The increased abundance of rusticyanin and cytochrome Cyc1 (both part of the ferrous iron oxidation pathway and encoded by the *rus* operon) in ferric iron-respiring *At. ferrooxidans* CCM 4253 further confirmed the contribution of *rus*-encoded, iron oxidation-related proteins in the iron reduction process (Kucera et al., 2012).

In 2013, Osorio et al. (2013) suggested two models for the anaerobic ferric iron reduction in *At. ferrooxidans*[T] (**Figure 3**), based on microarray and proteomic examination during aerobic and anaerobic growth. One supported the aforementioned reverse utilization of the iron oxidation chain, by proposing the involvement of the *c₄*-type cytochrome CycA2 and the *bc₁* complex. The second model introduced a new aspect, stating that an enzymatic disproportionation of sulfur during anaerobic sulfur oxidation may produce H_2S which could mediate ferric iron reduction at low pH (Osorio et al., 2013).

TABLE 1 | Most important discoveries during research of iron reduction in acidophiles.

Year	Investigation	References
1976	First description of ferric iron reduction by acidophilic bacteria and archaea: • *Acidithiobacillus thiooxidans* (aerobic) • *Acidithiobacillus ferrooxidans* (anaerobic) • *Sulfolobus acidocaldarius* (aerobic)	Brock and Gustafson, 1976
1987	Purification and characterization of a sulfur:ferric iron oxidoreductase from ferrous iron-grown *At. ferrooxidans*	Sugio et al., 1987
1987	First model of anaerobic iron respiration pathway of *At. ferrooxidans* utilizing transporters of the ferrous iron oxidation chain (cytochrome bc_1 complex, cytochrome c, rusticyanin, cytochrome oxidase, cytochrome a_1)	Corbett and Ingledew, 1987
1989	Aerobic iron reduction by resting cells of *At. ferrooxidans* at extremely low pH	Sand, 1989
1991	Anaerobic dissimilatory iron reduction by *At. ferrooxidans*	Pronk et al., 1991
1991	Aerobic dissimilatory iron reduction by heterotrophic *Acidiphilium* spp.	Johnson and McGinness, 1991
1998	Iron reduction by moderately thermophilic iron-oxidizing bacteria: • *Sulfobacillus thermosulfidooxidans* • *Sulfobacillus acidophilus* • *Acidimicrobium ferrooxidans* Iron cycling in static, aerobic cultures	Bridge and Johnson, 1998
2002	Distinction between a constitutive and oxygen-inducible iron reduction system in *Acidiphilium* spp.	Johnson and Bridge, 2002
2011	First demonstration of using iron reduction by acidophiles for mineral processing and metal extraction	Hallberg et al., 2011a
2013	Suggestion of a second, indirect iron reduction mechanism in *At. ferrooxidans* mediated by H_2S at low pH	Osorio et al., 2013
2015	Indication of an indirect, non-growth-related iron reduction mechanism in *At. thiooxidans* at aerobic, extreme acidic conditions	Marrero et al., 2015
2016	Hypothetic iron reduction pathway in *At. ferrooxidans* including a high potential iron-sulfur protein	Kucera et al., 2016a
2020	Model for anaerobic iron reduction coupled to hydrogen oxidation in *At. ferrooxidans*	Kucera et al., 2020
2021	Introduction of the term "latent iron reduction" for acidithiobacilli	Johnson et al., 2021

However, the ability of acidithiobacilli to grow anaerobically by coupling hydrogen oxidation (no H_2S production) to iron reduction (Hedrich and Johnson, 2013b) suggested that ferric iron reduction mechanisms other than the indirect one must exist (Osorio et al., 2013). Still, the existence of an indirect iron reduction mechanism by H_2S was not confirmed for *At. ferrooxidans* CCM 4253 due to no detection of H_2S with lead acetate and a repressed expression of the sulfur reductase-encoding *sre* operon during anaerobic growth (Kucera et al., 2016a). Besides, multiple ferric iron reduction mechanisms were again indicated during comparative examination of strain CCM 4253 with and without ferric iron reduction ability (Kucera et al., 2015). Proteomic results showed 150 repressed spots of important proteins involved in the iron oxidation (e.g., rusticyanin, cytochrome Cyc2) and sulfur metabolism for cells lacking the iron reduction ability (Kucera et al., 2015). The induction of rusticyanin and Cyc2 in anaerobic iron-sulfur-respiring *At. ferrooxidans* implied that electrons may be collected by rusticyanin and transported to the outer-membrane cytochrome which could act as terminal ferric iron reductase (Kucera et al., 2015, 2016b). Moreover, transcriptional investigation of strain CCM 4253 coupling iron reduction to anaerobic sulfur oxidation revealed an overexpression of the *hip* gene encoding a high potential iron-sulfur protein (Hip) and thereby a new hypothetic iron reduction pathway in *At. ferrooxidans* (Kucera et al., 2016a). During aerobic growth, this iron-sulfur protein enables electron transfer between the quinone pool and the final oxidase, which reduces oxygen (Quatrini et al., 2006). Accordingly, the predicted, Hip-including anaerobic iron reduction mechanism suggested a transfer of

electrons obtained through sulfur oxidation to take place from the quinone pool to the bc_1 complex and is further transported by CycA2 and Hip to an unknown iron reductase (Kucera et al., 2016a; **Figure 3**). It has to be taken in consideration, that the pre-cultures used during the aforementioned anaerobic ferric iron reduction examinations were aerobically ferrous iron-grown, which invited Norris et al. (2018) to perform proteomic studies with anaerobically adapted cultures. Thereby, two further cytochromes, Cyc2B and Cyc1B, were detected and the allover results supported a participation of *petII*-encoded proteins (bc_1 complex II and CycA2) in the anaerobic ferric iron reduction of sulfur-oxidizing *At. ferrooxidans*[T] (Norris et al., 2018). This in turn reinforces the alternative ferric iron reduction mechanism using a high potential iron-sulfur protein as predicted by Kucera et al. (2016a).

To date, all the given insights into the biochemical iron reduction mechanism coupled to sulfur oxidation indicate the occurrence of at least two enzymatic, species-dependent iron reduction mechanisms in acidithiobacilli (**Figure 3**). The existence of an indirect process mediated by H_2S needs further investigation, as the contradicting results of different strains do not imply a tendency. Nevertheless, there is a lot of evidence for both suggested electron transport chains from the inner to the outer membrane which may transfer electrons from the quinone pool to (i) the final reductase Cyc2 *via* bc_1 complex I, CycA1, and rusticyanin; or (ii) a still unknown reductase by bc_1 complex II, CycA2, and Hip (Corbett and Ingledew, 1987; Ohmura et al., 2002; Kucera et al., 2012, 2015, 2016a,b; Osorio et al., 2013; Norris et al., 2018). Analogous to ferric iron reduction during sulfur oxidation, the proposed mechanisms

TABLE 2 | Ferric iron-reducing extreme acidophilic archaea.

Organism	Fe ox.	S⁰ ox.	pH	Temp. [°C]	Conditions Fe red., ED	Growth	References
Acidianus brierleyi[T]	+	+	1.5–2.0 1.0–6.0	70 45–75	Anaerobic: H_2S (w/, w/o OC)	+	Plumb et al., 2007
"*Acidianus copahuensis*"	+	+	2.5–3.0 1.0–5.0	75 55–80	Anaerobic: S^0 or H_2	+	Giaveno et al., 2013
"*Acidianus manzaensis*"[T]	−	+	1.2–1.5 1.0–5.0	80 60–90	Anaerobic: S^0, H_2 or OC	+	Yoshida et al., 2006
Acidianus sulfidivorans[T]	+	+	0.8–1.4 0.35–3.0	74 45–83	Anaerobic: H_2S (w/, w/o OC)	+	Plumb et al., 2007
Acidiplasma aeolicum[T]	+	RISC	1.4–1.6 0–4.0	42–45 15–65	Anaerobic: OC	+	Golyshina et al., 2009
Acidiplasma cupricumulans[T]	+	−	1.0–1.2 > 0.4	53.6 22–63	Anaerobic: $K_2S_4O_6$ w/ OC	+	Hawkes et al., 2006
"*Ferroplasma acidarmanus*"[T]	+	ND	1.2 0.2–2.5	42 23–46	Anaerobic: OC	+	Dopson et al., 2004
Ferroplasma acidiphilum[T], DR1 and MT17	+	–RISC	1.7 1.3–2.2	35 20–45	Anaerobic: OC	+	Dopson et al., 2004; Johnson, 2010
"*Ferroplasma thermophilum*"	+	−	1.0 0.2–2.5	45 30–60	Anaerobic: OC	+	Zhou et al., 2008
Metallosphaera sedula[T]	+	+	1.0–4.5	75 50–80	Aerobic: H_2	ND	Auernik and Kelly, 2010
Saccharolobus caldissimus[T]	FeS₂	−	3.0 1.5–6.0	85 65–93	Anaerobic: OC	+	Sakai and Kurosawa, 2018
Saccharolobus shibatae[T]	FeS₂	−	3.0 1.5–6.0	81 55–86	Anaerobic: OC	+	Sakai and Kurosawa, 2018
Sulfolobus acidocaldarius[T]	+	+	2.0–3.0 1.0–5.9	75–80 55–80	Aerobic: S^0 or OC Microaerobic: OC Anaerobic: OC	− ND −	Brock and Gustafson, 1976; Masaki et al., 2018
Sulfuracidifex metallicus[T]	+	+	2.0–3.0 1.0–4.5	65 50–75	Anaerobic: S^0	−	Masaki et al., 2018
Sulfurisphaera ohwakuensis[T]	FeS₂	+	2.0 1.5–6.0	84 60–91	Anaerobic: OC	+	Tsuboi et al., 2018
Sulfurisphaera tokodaii[T]	+	+	2.5–3.0 1.5–6.0	80 60–96	Anaerobic: S^0 or OC Micro-aerobic: OC	+ ND	Masaki et al., 2018
Thermoplasma acidophilum[T]	ND	−	1.0–2.0 0.5–34.0	59 45–63	Anaerobic: S^0	ND	Zhang et al., 2018

pH, temperature, sulfur oxidation and iron oxidation data correspond to the optima and range, which are described in the respective characterization publications. The growth column indicates whether dissimilatory iron reduction occurs under the given conditions. Temp, temperature; ox, oxidation; red, reduction; ED, electron donor; ND, no data; OC, organic carbon.

of anaerobic iron reduction coupled to hydrogen oxidation are only partially understood and are largely based on proteomic and transcriptomic analyses. In 2020, Kucera et al. (2020) introduced a model for anaerobic hydrogen metabolism in *At. ferrooxidans* CCM 4253 which suggests an electron transfer from hydrogen to iron *via* [NiFe] hydrogenase, ubiquinone pool, bc_1 complex, cytochrome c_4, Cyc1 or Hip, rusticyanin, and Cyc2. Consequently, the mechanisms that enable ferric iron reduction in an anaerobic environment may involve several similar enzymes and proteins which are also part of the ferrous iron oxidation pathway in acidophiles, regardless of the electron donor used (**Figure 4**). Therefore, it is likely, that organisms lacking these important components in their electron transport chain of iron oxidation (e.g., *rus*-encoded proteins), may not be able to enzymatically reduce ferric iron. However, there is still great research potential concerning these iron reduction processes. Only well-understood biochemical mechanisms can be optimally adapted to applications and supported by the required environmental conditions. Moreover, a detailed insight into the underlying microbial processes is key for the assessment of their large-scale implementation as biohydrometallurgical processes for reducing ferric iron minerals, such as goethite.

Iron Reduction Under Micro-Aerobic Conditions

Besides iron-reducing processes under oxic or anoxic conditions, some acidophilic, heterotrophic microorganisms reduce ferric iron and/or mediate reductive dissolution of ferric iron minerals under micro-aerobic conditions. The type strains of *Acidibacter ferrireducens*, *Acidicapsa acidiphila,* and *Acidicapsa ferrireducens* catalyze the reductive dissolution of schwertmannite while metabolizing organic carbon sources (glucose, yeast extract) under oxygen-limitation (Falagán and Johnson, 2014;

TABLE 3 | Ferric iron-reducing extreme acidophilic autotrophic bacteria.

Organism	Fe ox.	S⁰ ox.	pH	Temp. [°C]	Conditions Fe red., ED	Growth	References
Acidiferrobacter thiooxydans[T]	+	+	2.0 > 1.2	38 < 47	Anaerobic: S^0	+	Hallberg et al., 2011b
Acidithiobacillus caldus[T]	–	+	2.0–2.5 1.0–3.5	45 32–52	Aerobic: S^0	–	Johnson et al., 2017
Acidithiobacillus ferrianus[T]	+	+	2.0	30	Anaerobic: S^0 or H_2	+	Norris et al., 2020
Acidithiobacillus ferridurans[T]	+	+	2.1 < 1.3	29	Anaerobic: S^0, $K_2S_4O_6$ or H_2 Aerobic: S^0	+ –	Hedrich and Johnson, 2013a; Johnson et al., 2017
Acidithiobacillus ferriphilus[T]	+	+	2.0	30	Anaerobic: RISC	+	Falagán and Johnson, 2016
Acidithiobacillus ferrivorans[T], Peru6	+	+	2.5 1.9–3.4	27–32 4–37	Anaerobic: S^0 Aerobic: S^0	+ –	Hallberg et al., 2010; Johnson et al., 2017
Acidithiobacillus ferrooxidans[T]	+	+	2.5 1.3–4.5	30–35 10–37	Anaerobic: S^0 Aerobic: S^0	– +	Pronk et al., 1992; Johnson et al., 2017
Acidithiobacillus thiooxidans[T]	–	+	2.0–3.0 0.5–5.5	28–30 10–37	Aerobic: S^0	–	Johnson et al., 2017

pH, temperature, sulfur oxidation and iron oxidation data correspond to the optima and range, which are described in the respective characterization publications. The growth column indicates whether dissimilatory iron reduction occurs under the given conditions. Temp, temperature; ox, oxidation; red, reduction; ED, electron donor.

TABLE 4 | Ferric iron-reducing extreme acidophilic mixotrophic bacteria.

Organism	Fe ox.	S⁰ ox.	pH	Temp. [°C]	Conditions Fe red., ED	Growth	References
Aciditerrimonas ferrireducens[T]	–	ND	3.0 2.0–4.5	50 35–58	Anaerobic: H_2 or OC	+	Itoh et al., 2011
Sulfobacillus acidophilus[T]	+	+	~2.0	45–50	Anaerobic: S^0 or OC	+	Johnson, 2010; Zhang et al., 2021
Sulfobacillus thermosulfidooxidans[T]	+	+	1.7–2.4 1.5–5.5	50–55 20–60	Anaerobic: H_2, $K_2S_4O_6$ or OC	+	Bridge and Johnson, 1998; Hedrich and Johnson, 2013b; Zhang et al., 2021

pH, temperature, sulfur oxidation and iron oxidation data correspond to the optima and range, which are described in the respective characterization publications. The growth column indicates whether dissimilatory iron reduction occurs under the given conditions. Temp, temperature; ox, oxidation; red, reduction; ED, electron donor; ND, no data; OC, organic carbon.

Falagán et al., 2017). Furthermore, heterotrophic, micro-aerobic *Acidicaldus organivorans*[T], *Acidobacterium capsulatum*[T], *Acidocella aromatica*[T], *Acidithrix ferrooxidans*[T], and *Sulfolobus acidocaldarius*[T] cultures showed iron-reducing properties (Johnson et al., 2006; Coupland and Johnson, 2008; Jones et al., 2013; Jones and Johnson, 2015; Masaki et al., 2018). In addition, several micro-aerophilic, acidophilic iron reducers perform this process under anaerobic conditions as well (Falagán and Johnson, 2014; Jones and Johnson, 2015), while others require low oxygen concentrations for this purpose (Falagán et al., 2017). In this context, discriminating between anaerobic, aerobic and micro-aerobic conditions is often difficult, as it is determined by the oxygen solubility under the given conditions. For an accurate classification, the oxygen content would have to be monitored during the entire experiment, which is not always feasible and involves more intricate experimental effort. Accordingly, some of the described aerobic iron reduction processes might actually occur under oxygen limitation, as has been discussed, for instance, for the thermophilic archaeon *Sulfolobus acidocaldarius* strain 79–13 (Brock and Gustafson, 1976). Therefore, a precise monitoring of the oxygen saturation

in the culture medium is essential to obtain information on e.g., inhibition of iron reduction mechanisms by certain oxygen concentrations. This is particularly interesting to understand iron cycling at natural transition zones between oxic and anoxic habitats (Emerson et al., 2010).

Oxygen-Dependent Iron Reduction Processes of the Acidophilic (Chemo-) Organoheterotrophic Genus *Acidiphilium*

Comparative investigation on the impact of dissolved oxygen on ferric iron reduction by the extreme acidophile *Acidiphilium* spp. was conducted (Johnson and Bridge, 2002). Based on their different levels of aerobic iron reduction capacity (Johnson and McGinness, 1991), ferric iron reduction by *Acidiphilium acidophilum*[T] and *Acidiphilium* sp. SJH was studied under various dissolved oxygen concentrations. Comparison of both, specific iron reduction rates and whole-cell protein profiles of these species, revealed oxygen-dependent differences within the genus *Acidiphilium* (Johnson and Bridge, 2002). *Acidiphilium* sp. SJH efficiently reduced ferric iron under aerobic and micro-aerobic conditions, demonstrated by nearly constant specific

TABLE 5 | Ferric iron-reducing extreme acidophilic heterotrophic bacteria.

Organism	Fe ox.	S⁰ ox.	pH	Temp. [°C]	Conditions Fe red., ED	Growth	References
Acidicaldus organivorans[T]	ND	+	2.5–3.0 1.75– > 3.0	50–55 40–65	Anaerobic: OC Micro-aerobic: OC	+ +	Johnson et al., 2006
Acidiferrimicrobium australe[T]	+	−	3.0 1.7–4.5	30 20–39	Anaerobic: OC	+	González et al., 2020
Acidimicrobium ferrooxidans[T]	+	−	2.0	45–50	Anaerobic: OC	+	Jones and Johnson, 2015
Acidiphilium acidophilum[T]	−	+	2.0 2.0–4.5	30	Micro-aerobic: OC	+	Johnson and Bridge, 2002
Acidiphilium sp. SJH	ND	ND	2.0	28	Aerobic: OC Micro-aerobic: OC Anaerobic: OC	+ + −	Johnson and Bridge, 2002
Acidithrix ferrooxidans[T]	+	−	 2.0–4.4	 10–30	Anaerobic: OC Micro-aerobic: OC	+ +	Jones and Johnson, 2015
Alicyclobacillus tolerans[T]	+	+	2.5–2.7 1.5–5.0	37–42 20–55	Anaerobic: OC	+	Karavaiko et al., 2005; Johnson, 2010
Ferrimicrobium acidiphilum[T]	+	−	2.0 > 1.4	35 < 37	Anaerobic: OC	+	Johnson et al., 2009
Ferrithrix thermotolerans[T]	+	−	1.8 > 1.6	43 < 50	Anaerobic: OC	+	Johnson et al., 2009
Sulfobacillus benefaciens[T]	+	+	1.5 0.8–2.2	38.5 30–47	Anaerobic: OC	+	Johnson et al., 2008
Sulfobacillus harzensis[T]	+	+	3.0 1.5–5.0	45 25–55	Anaerobic: OC	+	Zhang et al., 2021
Sulfobacillus sibiricus[T]	+	+	2.0 1.1–2.6	55 17–60	Anaerobic: OC	+	Zhang et al., 2021
Sulfobacillus thermotolerans[T]	+	+	2.0 1.2–2.4	40 20–60	Anaerobic: OC	+	Zhang et al., 2021

pH, temperature, sulfur oxidation and iron oxidation data correspond to the optima and range, which are described in the respective characterization publications. The growth column indicates whether dissimilatory iron reduction occurs under the given conditions. Temp, temperature; ox, oxidation; red, reduction; ED, electron donor; ND, no data; OC, organic carbon.

iron reduction rates between 40 and 80% dissolved oxygen. In contrast, effective iron reduction by *Acidiphilium acidophilum*[T] was only observed under micro-aerobic conditions (20–40% dissolved oxygen). Both strains did not reduce iron in the absence of oxygen. The detection of three additional proteins in *Acidiphilium acidophilum*[T] under micro-aerobic conditions indicated an oxygen-inducible iron reduction mechanism for this species. Since the dissolved oxygen content had no significant influence on the protein profile of *Acidiphilium* SJH, it was suggested to feature a constitutive iron reductase system. This tendency was also detected in further species of *Acidiphilium*, dividing the genus into two groups (Johnson and Bridge, 2002). However, *Acidiphilium cryptum* JF-5 reduced ferric iron independent of the given oxygen saturation even under anaerobic conditions at pH 3 (Küsel et al., 2002). Still, small amounts of oxygen are required to support growth of *Acidiphilium* species (Johnson and Bridge, 2002).

Aerobic Iron Reduction at Extremely Low pH

In oxygen-rich environments, the detection of ferric iron reduction is hindered by the auto- and microbial oxidation of the generated ferrous iron. Consequently, the study of iron reduction processes under these conditions is challenging. The higher stability of ferrous iron at low pH values enables the monitoring of ferric iron reduction even in the presence of oxygen. Already in the 1970s, aerobic ferrous iron formation was reported for sulfur-oxidizing *Acidithiobacillus thiooxidans* and *Acidithiobacillus ferrooxidans* at low pH (Brock and Gustafson, 1976; Sand, 1989). Further, iron oxidation of *At. ferrooxidans* was non-detectable under these extremely acidic conditions. This led to the assumption that the iron oxidation chain of this species is inhibited below pH 1.3, preventing re-oxidation of the ferrous iron formed. A pH increase to 1.8 again enabled rapid ferrous iron oxidation (Sand, 1989). These results contradicted the until then prevailing assumption of exclusively anaerobic iron reduction by *At. ferrooxidans* (Brock and Gustafson, 1976). Also, strains belonging to the acidophilic, heterotrophic genus of *Acidiphilium*, were characterized as aerobic iron reducers, using glucose or glycerol as substrate (Johnson and McGinness, 1991). When comparing iron reduction rates, which are often used to obtain tendencies of organisms to reduce iron under certain conditions, it is important to respect what the numbers are referring to. In the case of *Acidiphilium cryptum* JF-5, the non-normalized ferric iron reduction rate of the anaerobic experiment (24.2 mmol L^{-1} day^{-1}) was smaller than at oxygen saturation (46.6 mmol L^{-1} day^{-1}), while normalizing the rate to cell density showed a lower iron reduction rate at oxygen saturation, which was attributed to
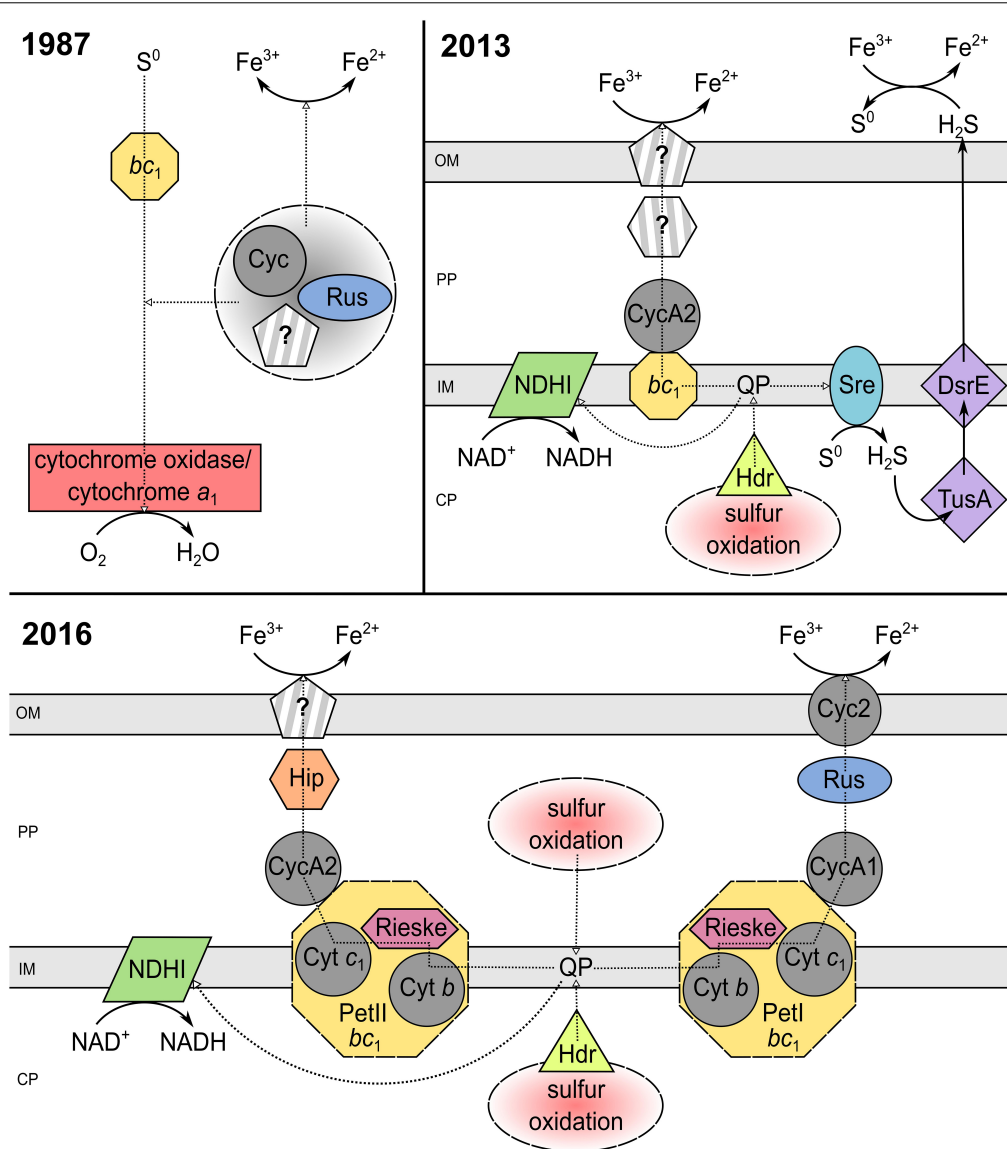
FIGURE 3 | Elucidation of the iron reduction mechanism coupled to sulfur oxidation by *Acidithiobacillus ferrooxidans* from 1987 to 2016. The underlying models of Corbett and Ingledew (1987); Osorio et al. (2013), and Kucera et al. (2016a) were modified to a consistent style. PP, periplasm; CP, cytoplasm; OM, outer membrane; IM, inner membrane; Cyt, cytochrome; Cyc, cytochrome *c*; QP, quinol pool; Rus, rusticyanin; NDHI, NADH dehydrogenase complex; *bc*$_1$, cytochrome *bc*$_1$ complex (encoded by *petI/II*); Sre, sulfur reductase; DsrE/TusA, conserved hypothetical protein (likely sulfur transferase); Hdr, heterodisulfide reductase; Hip, high potential iron-sulfur protein; Rieske, Rieske iron-sulfur protein; ?, unidentified pathway component.

better growth under these conditions (Küsel et al., 2002). Despite that, aerobic iron reduction was growth-related in further species, like *Acidithiobacillus thiooxidans*, *Acidiphilium cryptum*, and *Acidiphilium* sp. SJH (Brock and Gustafson, 1976; Johnson and McGinness, 1991). However, in *Sulfolobus acidocaldarius* and *At. ferrooxidans*, accumulation of ferrous iron was not accompanied by cell growth (Brock and Gustafson, 1976; Sand, 1989). Therefore, it remained unclear whether this was a microbial respiratory process or a non-enzymatic, chemical reduction mediated by metabolic intermediates accumulated in the culture liquor. Additionally, abiotic ferric iron reduction was shown for uninoculated flasks at thermophilic conditions (70°C). Still, the amount of iron reduced in inoculated batches was considerably higher (Brock and Gustafson, 1976).

Insights Into the Aerobic Ferric Iron Reduction Mechanisms in Acidithiobacilli

In recent years, aerobic ferric iron reduction ability was reported for various *Acidithiobacillus* species (**Table 3**) reducing soluble ferric iron and/or mediating the reductive dissolution of ferric iron ores (Marrero et al., 2015, 2017; Johnson et al., 2017; Smith and Johnson, 2018). The results obtained from these studies indicate species-dependent mechanisms for the "latent iron reduction" by acidithiobacilli (Johnson et al., 2021). Thereby, a

FIGURE 4 | Proposed mechanism of iron reduction coupled to hydrogen oxidation by *Acidithiobacillus ferrooxidans*. The underlying scheme is given in Kucera et al. (2020). PP, periplasm; CP, cytoplasm; OM, outer membrane; IM, inner membrane; Cyc, cytochrome *c*; QP, quinol pool; Rus, rusticyanin; NDHI, NADH dehydrogenase complex; *bc₁*, cytochrome bc_1 complex (encoded by *petI/II*); Hip, high potential iron-sulfur protein; ?, unidentified pathway component.

non-growth-related mechanism was assumed for *At. thiooxidans* (Marrero et al., 2015), which contradicted the formerly proposed correlation between ferrous iron accumulation and cell growth in cultures of *At. thiooxidans* (Brock and Gustafson, 1976). Furthermore, cell-free supernatants (pH 1.0) of sulfur-respiring *At. caldus* and *At. ferridurans*, also showed aerobic iron reduction, indicating a cell-independent chemical reduction process (Johnson et al., 2017). However, in the case of *At. thiooxidans*, washed cells were also capable of reducing iron within a short time under aerobic conditions (150–200 mg/L in 6.5 h) regardless of the pH (0.9, 2.5, 3.5) (Marrero et al., 2015). Initially, the aerobic iron reduction observed in *Acidithiobacillus* (especially *At. thiooxidans*) cultures was exclusively assigned to the reduction of ferric iron by sulfur intermediates formed during sulfur respiration under aerobic conditions (Marrero et al., 2015, 2017). However, Johnson et al. (2017, 2021) provided evidence, that aerobic ferric iron reduction mediated by acidithiobacilli is not only attributed to sulfur oxidation, as hydrogen-grown *At. caldus* and tetrathionate- or hydrogen-metabolizing *At. thiooxidans* reduced ferric iron under these conditions as well. Recently, indications were given that an initial hydrophilization of sulfur by acidithiobacilli may enable chemical, ferric iron reduction as tests with bio-activated, heat-sterilized sulfur allowed aerobic iron reduction at 50°C and pH 0.9–1.0 as

well (Johnson et al., 2021). These finding may also explain the pH-independent iron reduction of washed *At. thiooxidans* cells by cell-surface-associated activated sulfur particles. Accordingly, further studies are required to gain more insight into the processes that occur during aerobic iron reduction at low pH. Moreover, the aforementioned aerobic reduction of ferric iron by heterotrophic *Acidiphilum* species (Johnson and McGinness, 1991) implies that there are several mechanisms enabling ferrous iron formation in the presence of oxygen. Whether these mechanisms are enzymatic, non-enzymatic, or a combination of both should be further investigated. In this context, the detection and identification of intermediates of the sulfur metabolism is a major drawback due to their low stability and rapid degradation (Rohwerder, 2002).

APPLICATION AND ENVIRONMENTAL ASPECTS

Reductive dissolution by acidophiles was shown for several ferric iron-containing minerals, as ferric hydroxide [$Fe(OH)_3$], goethite (α-FeOOH), akageneite (β-FeOOH), magnetite (Fe_3O_4), potassium jarosite [$KFe_2(SO_4)_2(OH)_6$], and natrojarosite [$NaFe_2(SO_4)_2(OH)_6$] (Bridge and Johnson, 1998, 2000). These findings opened new possibilities for biohydrometallurgical processes where oxidative bioleaching is not effective or more sustainable methods should be established. Nevertheless, as other metal recovery methods, anaerobic (AnRD) and aerobic reductive dissolution (AeRD) require special operating conditions which have to be considered especially regarding operational costs. The maintenance of an anaerobic process requires a closed system and exclusion of oxygen, e.g., by supply of nitrogen. Further, AnRD is an acid-consuming reaction involving the addition of acid (Marrero et al., 2015), while AeRD can operate as net acid-producing process (Johnson et al., 2021). On the downside, the extremely low pH during AeRD is only tolerated by few microorganisms, which minimizes the number of suitable acidophiles. Moreover, the exclusion of aerobic iron oxidizers from AeRD applications may require an anaerobic phase integrated into the process to avoid interference with iron-oxidizing leptospirilli (Marrero et al., 2015). Still, reductive bioleaching is a promising method to exploit even low-grade ferric iron minerals and to recycle process wastes of traditional hydrometallurgical methods, e.g., laterite tailings (Marrero et al., 2015).

Biohydrometallurgical Application of Microbial Iron Reduction

The success of biotechnological applications in the metallurgical sector is evident from copper production *via* bioleaching in e.g., Chile and elsewhere in the world (Gentina and Acevedo, 2016). Despite the fact that biohydrometallurgical processes are nowadays well established, commercial applications are still limited to sulfidic minerals (Johnson, 2014). In the case of oxidative metal dissolution by acidophiles, the process is enabled by the sulfuric acid produced during sulfur oxidation by the microorganisms and/or the ferric iron generated by

microbial iron oxidation of the mineral (Johnson et al., 2013). The mechanism underlying reductive dissolution of ferric iron minerals is also caused, to some extent, by acidolysis (Marrero et al., 2015). Further, an equilibrium shift between iron in the solid phase and in solution created by microbial ferric iron reduction accelerates mineral dissolution (Bridge and Johnson, 1998). Although further research is required to understand the exact mechanisms involved, reductive bioleaching is already enabling new, economic recovery methods for different ferric iron-associated metals in laboratory scale.

Processing Ferric Iron Ores for Base Metal Recovery

Over the last decade, a lot of research has been conducted on the application of reductive microbial dissolution processes focusing on the extraction of metals from limonitic laterites (**Table 6**). Approximately 70% of world's land-based nickel resources are present in the form of laterite (Dalvi et al., 2004). Laterites are commercially processed by reagent- and energy-demanding pyrometallurgical (smelting) or hydrometallurgical (CARON process or high pressure acid leaching) processes (Dalvi et al., 2004). Nickel laterites consist of two zones, the upper limonite and the lower saprolite zone, of which the limonite one contains base metals associated with ferric iron in the form of modified goethite (Johnson et al., 2013). Bioleaching of limonite by acidithiobacilli is achieved through microbial acid generation, lowering the redox potential under anaerobic conditions, iron reduction under aerobic conditions and activating/"wetting" of naturally hydrophobic sulfur (Johnson et al., 2021). Advantages compared to conventional pyro- and hydrometallurgical methods are cost reduction by using inorganic energy sources (e.g., sulfur) and an ambient operating temperature (Johnson et al., 2013). The extremely low pH during AeRD and AnRD enhances the acidic attack during mineral dissolution and reduces metal precipitation, thereby simplifying downstream processing which is a major advantage compared to neutrophilic iron reduction approaches (Johnson et al., 2013).

According to the diverse composition of laterites, different base metals have already been the target of reductive bioleaching, such as nickel (Hallberg et al., 2011a), copper (Nancucheo et al., 2014), and cobalt (Smith et al., 2017), reaching up to 100% yield. It also became apparent, that the geographical origin of the ore and its mineralogical composition had a major impact on their biohydrometallurgical application (Santos et al., 2020). Initially, the applicability of reductive bioleaching was examined by using pure cultures of *Acidithiobacillus ferrooxidans* operating under anoxic conditions (Hallberg et al., 2011a; Johnson et al., 2013; Nancucheo et al., 2014). Afterward, other acidophiles and consortia as well as different aeration setups (AeRD, Ae-AnRD and AnRD) were investigated for this reductive extraction approach (Marrero et al., 2015; Santos et al., 2020; Johnson et al., 2021). In some processes, also thermotolerant strains were used (Smith et al., 2017; Johnson et al., 2021) which enable higher operation temperatures and thereby enhanced reaction kinetics. However, the use of consortia also showed that the diversity of the microbial community decreases during the reductive bioleaching process, which could sometimes be due to a lack of adaptation of the acidophiles to the mineral, resulting in the predominance of a

few strains (Smith et al., 2017; Santos et al., 2020). Monitoring of cell numbers by qPCR also showed a constant number of active bacteria during the AeRD process and a decrease after 3 days in the case of AnRD and Ae-AnRD (Marrero et al., 2015). Moreover, comparison of reductive limonite dissolution by a mesophilic consortium of *At. ferrooxidans*[T], *At. ferriphilus*[T], *At. ferridurans*[T], *Sulfobacillus acidophilus*[T], and *Sulfobacillus thermosulfidooxidans*[T] showed higher base metal yields in the anaerobic bioreactor, while the aerobic bioreactor required no further acid supply after 9 days (Johnson et al., 2021). Smith et al. (2017) observed the same tendency for cobalt extraction from limonite, which was more efficient using the anaerobic reductive leaching approach and seemed to correlate with manganese dissolution. In contrast, AeRD experiments with pure cultures of non-iron-oxidizing acidithiobacilli (*At. thiooxidans* or *At. caldus*) resulted in higher metal yields compared to AnRD approaches with the same limonite (Marrero et al., 2017; Johnson et al., 2021). Marrero et al. (2017) provided a direct comparison of AnRD and AeRD of nickel laterite overburden by *At. ferrooxidans* or *At. thiooxidans*, respectively, showing approximately the same yields. Accordingly, in this case AeRD was selected as the favorable method since it does not require anoxic conditions, it enables better extraction kinetics, and the extremely low pH reduces iron precipitation (Marrero et al., 2017). The success of bioprocessing of limonitic ores depends on the mineralogy. If large amounts of acid-unstable ferrous iron minerals are present, manganese-associated cobalt is readily solubilized by microbial acid generation under aerobic conditions, whereas ferric iron reduction by acidophiles promotes solubilization of ferric iron minerals and nickel (Johnson et al., 2021). In addition, microbial ferrous iron generation accelerates the reductive dissolution on Mn (IV) minerals (Santos et al., 2020; Johnson and Pakostova, 2021).

All this indicates that, as in any biotechnological process, a large number of parameters need to be optimized to obtain the best possible results. These are highly influenced by the mineral composition, the bioleaching method and the microorganisms involved. Accordingly, pre-investigations concerning the characteristics of the applied acidophiles and ores are crucial. Furthermore, abiotic effects during reductive bioleaching must be considered. In the context of abiotic effects caused by chemical leaching *via* acid dissolution at extremely low pH values used for AnRD and AeRD, it was shown that bacterial catalyzed approaches with iron reducers allow more effective metal dissolution than acidolysis alone (Nancucheo et al., 2014). Reductive bioleaching was also up to 6 times more effective than acid leaching in processing cobalt-containing limonite from Kazakhstan or the Philippines under aerobic conditions (Smith et al., 2017). Further benefits of reductive bioleaching approaches were shown by an initial anaerobic phase during leaching of polymetallic sulfides which had to some extend positive effects on the base metal recovery and enhanced the iron removal noticeable at both 45 and 70°C (Norris et al., 2015).

Additional to the bioleaching of base metals from laterites, iron reduction by acidophiles offers far more application potential. The bio-reductive processes would be especially suitable to recycle scrap electronics or bioremediate

TABLE 6 | Overview of reductive mineral dissolution using extreme acidophiles.

Mineral	Origin	Microorganisms	Method	Yield	References
Nickel laterite	Western Australia	At. ferrooxidans[T]	AnRD	> 80% Ni	Hallberg et al., 2011a
Nickel laterite	Western Australia	At. ferrooxidans[T]	AeRD	~10% Ni	Hallberg et al., 2011a
Limonitic nickel laterite	Australia	At. ferrooxidans[T]	AnRD	46–82% Ni 64–90% Co 75–116% Mn	Johnson et al., 2013
Copper laterite	Carajás Belt, Pará State, Brazil	At. ferrooxidans[T]	AnRD	≤ 78% Cu	Nancucheo et al., 2014
Laterite tailings	Moa, Cuba; from CARON process	At. thiooxidans[T]	AeRD	53% Ni 46% Co	Marrero et al., 2015
Laterite tailings	Moa, Cuba; from CARON process	At. thiooxidans[T] and At. ferrooxidans[T]	AeRD	56% Ni 60% Co	Marrero et al., 2015
Laterite tailings	Moa, Cuba; from CARON process	At. thiooxidans[T] and At. ferrooxidans[T]	Ae-AnRD	53% Ni 58% Co	Marrero et al., 2015
Laterite tailings	Moa, Cuba; from CARON process	At. ferrooxidans[T]	AnRD	58% Ni 56% Co	Marrero et al., 2015
Nickel laterite overburden	Punta Gorda, Moa, Cuba	At. thiooxidans[T]	AeRD	16% Ni 85% Co 74% Mn	Marrero et al., 2017
Nickel laterite overburden	Punta Gorda, Moa, Cuba	At. ferrooxidans[T]	AnRD	16% Ni 78% Co 86% Mn	Marrero et al., 2017
Cobalt-bearing limonitic laterite ore	Shevchenko, Kazakhstan	Consortium mesophilic and thermotolerant bacteria[a]	AnRD	50–70% Ni 89–99% Co	Smith et al., 2017
Cobalt-bearing limonitic laterite ore	Acoje mine, Philippines	Consortium mesophilic and thermotolerant bacteria[a]	AnRD	40% Ni 90% Co	Smith et al., 2017
Cobalt-bearing limonitic materials	Kastoria, Agios Ioannis and Evia mine, Greece	Consortium mesophilic bacteria[b]	AnRD	37–73% Ni 40–50% Co 15–52% Mn	Santos et al., 2020
Limonite ore	Shevchenko, Kazakhstan	Consortium mesophilic bacteria[c]	AnRD	~55% Ni ~90% Co ~95% Mn	Johnson et al., 2021
Limonite ore	Shevchenko, Kazakhstan	Consortium mesophilic bacteria[c]	AeRD	~25% Ni ~20% Co ~40% Mn	Johnson et al., 2021
Limonite ore	Shevchenko, Kazakhstan	At. caldus[T]	AeRD	~100% Ni ~95% Co ~90% Mn	Johnson et al., 2021

[a]At. ferrooxidans[T], At. ferriphilus[T], At. ferridurans[T], "Acidibacillus sulfuroxidans"[T], Sulfobacillus thermosulfidooxidans[T], and Sulfobacillus acidophilus BOR1.
[b]At. ferrooxidans[T], At. ferrooxidans CF3, At. ferriphilus[T], At. ferridurans[T], and Sulfobacillus thermosulfidooxidans[T].
[c]At. ferrooxidans[T], At. ferriphilus[T], At. ferridurans[T], Sulfobacillus acidophilus[T], and Sulfobacillus thermosulfidooxidans[T].
AnRD, anaerobic reductive dissolution; AeRD, aerobic reductive dissolution; Ae-AnRD, aerobic reductive dissolution with an integrated anaerobic phase.

metal-contaminated soil (Johnson et al., 2013). Another approach aimed at the pretreatment of laterite-associated monazite, a rare earth elements-containing phosphate mineral, by ferric iron reduction to improve the monazite exposure for further acid dissolution (Nancucheo et al., 2019). However, due to the innovation of this method, further research is required before it is ready for application.

Introducing a Reductive Bioleaching Circuit: The Ferredox Concept

In 2011, a generic concept for ferric iron mineral processing, Ferredox, was presented for tropical limonitic laterite (du Plessis et al., 2011). In assembly, the model has not yet been applied in a full-scale operation, but it has been proposed as a promising sustainable metal recovery method. The modular process relies on the findings that anaerobic reductive dissolution of goethite can take place under ambient conditions utilizing bacterial catalysis (Hallberg et al., 2011a). This new perspective creates a biohydrometallurgical method, which might allow iron oxide treatment with reduced material and processing costs. Overall, the process consists of four modularized main stages: reductive bioleaching, metal recovery, iron oxidation/precipitation, and reductive acid generation. During the initial reductive bioleaching, iron-oxidizing/reducing

acidithiobacilli couple the oxidation of elemental sulfur to the reduction of ferric iron present in goethite and other minerals, solubilizing target metals. The resulting pregnant leaching solution is subsequently processed to extract the focused base metals. After a two staged iron oxidation process, the formed iron precipitates are utilized to generate the sulfuric acid consumed during the initial reductive dissolution of goethite $[S^0 + 6FeO(OH) + 10H^+ \rightarrow SO_4{}^{2-} + 6Fe^{2+} + 8H_2O]$. In turn, this process will make use of the metabolism of *At. ferrooxidans* once again by anaerobic oxidation of elemental sulfur generating acid $(S^0 + 6Fe^{3+} + 4H_2O \rightarrow SO_4{}^{2-} + 6Fe^{2+} + 8H^+)$. Advantages of the given Ferredox concept are biomass recycling and enhancement due to the use of *At. ferrooxidans* in multiple process steps and increased sustainability due to biologically catalyzed reactions and operation at ambient conditions (du Plessis et al., 2011; Johnson and du Plessis, 2015).

Environmental Impact of Acidophilic Iron Reducers

In an ancient, oxygen-free atmosphere, the usage of ferric iron as electron acceptor may have had a major role and impact on live (Walker, 1987; Slobodkin, 2005). Currently, oxido-reduction of iron has been demonstrated for both pure and mixed cultures of acidophiles (Johnson et al., 1993). It is now known that microbial iron cycling effects inorganic and organic contaminants, carbon cycling, and nutrient mobilization in the environment (Kappler and Straub, 2005). For instance, iron-reducing acidophiles were shown to eliminate the toxicity of some oxy-anions through ferrous iron generation, e.g., the more toxic Cr(VI) (hydrogen chromate, $HCrO_4{}^-$) is reduced to the less harmful Cr(III) by ferrous iron (Johnson et al., 2017).

Again, the dissolution of iron oxides releases metal cations absorbed to the mineral, which may have a negative impact on the biosphere (Kappler and Straub, 2005). Nevertheless, ferrous iron, the product of ferric iron reduction, can in turn act as a strong reductant for organic pollutants, e.g., nitroaromatics (Hofstetter et al., 1999).

The fact that acidophilic iron reducers feature autotrophic as well as heterotrophic metabolisms ensures their function as both, producers and degraders in ecosystems (Kashefi et al., 2004). The resulting complete mineralization of organic matter prevents accumulation of high amounts of organic substances which could have a toxic effect, e.g., small organic acids, on autotrophic acidophilic microorganisms. In acid streamers, dead organic matter and generated iron precipitates settle down to anoxic, organotrophic acidophiles-inhabited zones, which were shown to feature more ferrous than ferric iron due to the occurring oxidation of organic substances coupled to iron reduction (Johnson et al., 1993, 2012). Thus, acidophiles actively contribute to the carbon and iron cycle. In the river Rio Tinto, iron transformation was shown to be mainly enabled by the interaction of the autotrophic iron oxidizers *Leptospirillum ferrooxidans* and *At. ferrooxidans* in aerobic areas and iron reduction by *At. ferrooxidans* (autotrophic) and *Acidiphilium* species (heterotrophic) in anaerobic zones (González-Toril et al., 2003). Investigations of the microbial diversity in sediments of

an acidic lignite mine lake (Lusatia, Germany) revealed further iron reducers and cycling in these environments. Isolates closest related to species of *Acidiphilium*, *Acidocella*, *Ferrimicrobium*, *Acidithiobacillus*, *Alicyclobacillus*, and *Acidobacteriaceae* reduced iron in this habitat, with some of them being found throughout all zones of the sediment (Lu et al., 2010). Likewise, previous investigations of the same coal mine lake emphasized this pH gradient-dependent heterogeneity of ferric iron-reducing acidophiles in the sediment (Blöthe et al., 2008). In addition, two of the isolated *Alicyclobacillus* relatives showed iron cycling capacity in the laboratory by dissolving schwertmannite and re-oxidizing the ferrous iron by forming schwertmannite precipitates again (Lu et al., 2010). Still, the capacity of iron cycling by pure cultures was previously described for other acidophilic isolates (Johnson et al., 1993). Nevertheless, in many acidic ecosystems, a diverse community of phyla is responsible for the conversion of iron and other elements. Depending on the conditions, a closed iron cycle can be independently realized by individual species. According to the manifold influences of acidophilic iron reducers on different ecosystems and material cycles, it is important to study this group of microorganisms and thus to achieve a better awareness of human interventions, especially with regard to anthropogenically induced acidic and metal-rich habitats.

CONCLUSION AND PERSPECTIVES

Although iron reduction processes mediated by acidophiles have been known for nearly 50 years, it is only in the last decade that much interest has developed in their application. This especially included the utilization of well-studied ferric iron reducers of the genus *Acidithiobacillus*. Nevertheless, far more iron-reducing acidophiles have been identified including over 20 different genera. This is justified by the growing impact of this metabolic trait due to the recently known suitability for reductive bioleaching of ferric iron minerals. Accordingly, the characterization of novel acidophilic microorganisms now usually includes testing for iron reduction, which was not necessarily the case in the past. Still, few of the newly discovered iron reducers are investigated in detail. This lack of knowledge offers a great demand for further research particularly with regards to thermophilic conditions which could enhance the leaching kinetics. Moreover, the different mechanisms of ferric iron reduction in acidophiles (anaerobic, micro-aerobic or aerobic) allow diverse operation setups for base metal processing from laterites and a correspondingly higher freedom to customize the metal extraction method according to the mineral composition. Recently, the high potential to use acidophile-promoted reductive dissolution was demonstrated by exceeding base metal extraction from ferric minerals and overburden. Approaches were applied to recycle low-grade wastes from traditional laterite processing by reductive bioleaching. Additionally, first approaches to pretreat rare earth-bearing minerals associated with laterites showed promising results. Besides, ferric iron reduction by acidophiles might further be suitable for bioremediation and scrap electronics recycling.

All the benefits of such an innovative application profit from precise knowledge of the underlying microbiological processes and mineral composition. Both call for suitable devices and expertise in microbiological and mineralogical fields which may in some cases rely on an interdisciplinary cooperation. Further adjustments of this biohydrometallurgical approach require a sufficient understanding of the biochemical iron reduction mechanism, especially concerning acidophiles apart from acidithiobacilli. As *Sulfobacillus thermosulfidooxidans* was shown to be the predominant microorganism (beside *Acidithiobacillus ferrooxidans*) in reductive bioleaching experiments with moderate consortia, this acidophile seems to be an outstanding candidate for detailed investigations of its iron reduction mechanism. The biochemical information obtained for *At. ferrooxidans* offered new knowledge and interesting insights, by showing a reverse usage of iron oxidation chain components. Accordingly, a comprehensive insight into the iron reduction mechanisms of other acidophiles, especially those with iron oxidation pathways other than *At. ferrooxidans*, would be of great interest. The research field of reductive bioleaching would benefit from further studies on additional acidophilic genera and even archaea.

AUTHOR CONTRIBUTIONS

REFERENCES

Ai, C., Liang, Y., Miao, B., Chen, M., Zeng, W., and Qiu, G. (2018). Identification and analysis of a novel gene cluster involves in Fe^{2+} oxidation in *Acidithiobacillus ferrooxidans* ATCC 23270, a typical biomining acidophile. *Curr. Microbiol.* 75, 818–826. doi: 10.1007/s00284-018-1453-9

Amouric, A., Brochier-Armanet, C., Johnson, D. B., Bonnefoy, V., and Hallberg, K. B. (2011). Phylogenetic and genetic variation among Fe(II)-oxidizing acidithiobacilli supports the view that these comprise multiple species with different ferrous iron oxidation pathways. *Microbiology* 157, 111–122. doi: 10.1099/mic.0.044537-0

Auernik, K. S., and Kelly, R. M. (2010). Impact of molecular hydrogen on chalcopyrite bioleaching by the extremely thermoacidophilic archaeon *Metallosphaera sedula*. *Appl. Environ. Microbiol.* 76, 2668–2672. doi: 10.1128/AEM.02016-09

Blöthe, M., Akob, D. M., Kostka, J. E., Göschel, K., Drake, H. L., and Küsel, K. (2008). pH gradient-induced heterogeneity of Fe(III)-reducing microorganisms in coal mining-associated lake sediments. *Appl. Environ. Microbiol.* 74, 1019–1029. doi: 10.1128/AEM.01194-07

Bonnefoy, V. (2010). "Bioinformatics and genomics of iron- and sulfur-oxidizing acidophiles," in *Geomicrobiology: Molecular and Environmental Perspective*, eds L. L. Barton, M. Mandl, and A. Loy (Dordrecht: Springer), 169–192.

Bridge, T. A. M., and Johnson, D. B. (1998). Reduction of soluble iron and reductive dissolution of ferric iron-containing minerals by moderately thermophilic iron-oxidizing bacteria. *Appl. Environ. Microbiol.* 64, 2181–2186. doi: 10.1128/AEM.64.6.2181-2186.1998

Bridge, T. A. M., and Johnson, D. B. (2000). Reductive dissolution of ferric iron minerals by *Acidiphilium* SJH. *Geomicrobiol. J.* 17, 193–206. doi: 10.1080/01490450050121161

Brock, T. D., and Gustafson, J. (1976). Ferric iron reduction by sulfur- and iron-oxidizing bacteria. *Appl. Environ. Microbiol.* 32, 567–571. doi: 10.1128/AEM.32.4.567-571.1976

Bromfield, S. M. (1954). The reduction of iron oxide by bacteria. *J. Soil Sci.* 5, 129–139. doi: 10.1111/j.1365-2389.1954.tb02181.x

Corbett, C. M., and Ingledew, W. J. (1987). Is $Fe^{3+/2+}$ cycling an intermediate in sulphur oxidation by Fe^{2+}-grown *Thiobacillus ferrooxidans*. *FEMS Microbiol. Lett.* 41, 1–6. doi: 10.1111/j.1574-6968.1987.tb02131.x

Coupland, K., and Johnson, D. B. (2008). Evidence that the potential for dissimilatory ferric iron reduction is widespread among acidophilic heterotrophic bacteria. *FEMS Microbiol. Lett.* 279, 30–35. doi: 10.1111/j.1574-6968.2007.00998.x

Dahl, C. (2020). "A biochemical view on the biological sulfur cycle," in *Environmental Technologies to Treat Sulphur Pollution: Principles and Engineering*, ed. P. N. L. Lens (London: IWA Publishing), 55–96.

Dalvi, A. D., Bacon, W. G., and Osborne, R. C. (eds) (2004). "The past and the future of nickel laterites," in *Proceedings of the PDAC 2004 International Conference Trade Show and Investors Exchange*, Toronto, 7–10.

Das, A., Mishra, A. K., and Roy, P. (1992). Anaerobic growth on elemental sulfur using dissimilar iron reduction by autotrophic *Thiobacillus ferrooxidans*. *FEMS Microbiol. Lett.* 97, 167–172. doi: 10.1111/j.1574-6968.1992.tb05457.x

Dopson, M. (2016). "Physiological and phylogenetic diversity of acidophilic bacteria," in *Acidophiles: Life in Extremely Acidic Environments*, eds R. Quatrini and D. B. Johnson (Poole: Caister Academic Press), 79–92.

Dopson, M., Baker-Austin, C., and Bond, P. (2007). Towards determining details of anaerobic growth coupled to ferric iron reduction by the acidophilic archaeon 'Ferroplasma acidarmanus' Fer1. *Extremophiles* 11, 159–168. doi: 10.1007/s00792-006-0029-y

Dopson, M., Baker-Austin, C., Hind, A., Bowman, J. P., and Bond, P. L. (2004). Characterization of *Ferroplasma* isolates and *Ferroplasma acidarmanus* sp. nov., extreme acidophiles from acid mine drainage and industrial bioleaching environments. *Appl. Environ. Microbiol.* 70, 2079–2088. doi: 10.1128/AEM.70.4.2079-2088.2004

du Plessis, C. A., Slabbert, W., Hallberg, K. B., and Johnson, D. B. (2011). Ferredox: a biohydrometallurgical processing concept for limonitic nickel laterites. *Hydrometallurgy* 109, 221–229. doi: 10.1016/j.hydromet.2011.07.005

Ehrenreich, A., and Widdel, F. (1994). Anaerobic oxidation of ferrous iron by purple bacteria, a new type of phototrophic metabolism. *Appl. Environ. Microbiol.* 60, 4517–4526. doi: 10.1128/aem.60.12.4517-4526.1994

Elbehti, A., Brasseur, G., and Lemesle-Meunier, D. (2000). First evidence for existence of an uphill electron transfer through the bc_1 and NADH-Q oxidoreductase complexes of the acidophilic obligate chemolithotrophic ferrous ion-oxidizing bacterium *Thiobacillus ferrooxidans*. *J. Bacteriol.* 182, 3602–3606. doi: 10.1128/JB.182.12.3602-3606.2000

Elbehti, A., Nitschke, W., Tron, P., Michel, C., and Lemesle-Meunier, D. (1999). Redox components of cytochrome *bc*-type enzymes in acidophilic prokaryotes. I. Characterization of the cytochrome bc_1-type complex of the acidophilic ferrous ion-oxidizing bacterium *Thiobacillus ferrooxidans*. *J. Biol. Chem.* 274, 16760–16765. doi: 10.1074/jbc.274.24.16760

Emerson, D., Fleming, E. J., and McBeth, J. M. (2010). Iron-oxidizing bacteria: an environmental and genomic perspective. *Annu. Rev. Microbiol.* 64, 561–583. doi: 10.1146/annurev.micro.112408.134208

Falagán, C., Foesel, B., and Johnson, B. (2017). *Acidicapsa ferrireducens* sp. nov., *Acidicapsa acidiphila* sp. nov., and *Granulicella acidiphila* sp. nov. novel acidobacteria isolated from metal-rich acidic waters. *Extremophiles* 21, 459–469. doi: 10.1007/s00792-017-0916-4

Falagán, C., and Johnson, D. B. (2014). *Acidibacter ferrireducens* gen. nov., sp. nov. an acidophilic ferric iron-reducing gammaproteobacterium. *Extremophiles* 18, 1067–1073. doi: 10.1007/s00792-014-0684-3

Falagán, C., and Johnson, D. B. (2016). *Acidithiobacillus ferriphilus* sp. nov., a facultatively anaerobic iron- and sulfur-metabolizing extreme acidophile. *Int. J. Syst. Evol. Microbiol.* 66, 206–211. doi: 10.1099/ijsem.0.000698

Ferrer, A., Orellana, O., and Levicán, G. (2016). "Oxidative stress and metal tolerance in extreme acidophiles," in *Acidophiles: Life in Extremely Acidic Environments*, eds R. Quatrini and D. B. Johnson (Norfolk: Caister Academic Press), 63–76.

Gadd, G. M. (2009). "Heavy metal pollutants: environmental and biotechnological aspects," in *Encyclopedia of Microbiology*, ed. M. Schaechter (Oxford: Elsevier), 321–334.

Gentina, J. C., and Acevedo, F. (2016). Copper bioleaching in Chile. *Minerals* 6:23. doi: 10.3390/min6010023

Giaveno, M. A., Urbieta, M. S., Ulloa, J. R., Toril, E. G., and Donati, E. R. (2013). Physiologic versatility and growth flexibility as the main characteristics of a novel thermoacidophilic *Acidianus* strain isolated from Copahue geothermal area in Argentina. *Microb. Ecol.* 65, 336–346. doi: 10.1007/s00248-012-0129-4

Giudici-Orticoni, M. T., Guerlesquin, F., Bruschi, M., and Nitschke, W. (1999). Interaction-induced redox switch in the electron transfer complex rusticyanin-cytochrome c_4. *J. Biol. Chem.* 274, 30365–30369. doi: 10.1074/jbc.274.43.30365

Golyshina, O. V., Ferrer, M., and Golyshin, P. N. (2016). "Diversity and physiologies of acidophilic *Archaea*," in *Acidophiles: Life in Extremely Acidic Environments*, eds R. Quatrini and D. B. Johnson (Norfolk: Caister Academic Press), 93–106.

Golyshina, O. V., Yakimov, M. M., Lünsdorf, H., Ferrer, M., Nimtz, M., Timmis, K. N., et al. (2009). *Acidiplasma aeolicum* gen. nov., sp. nov., a euryarchaeon of the family *Ferroplasmaceae* isolated from a hydrothermal pool, and transfer of *Ferroplasma cupricumulans* to *Acidiplasma cupricumulans* comb. nov. *Int. J. Syst. Evol. Microbiol.* 59, 2815–2823. doi: 10.1099/ijs.0.009639-0

González, D., Huber, K. J., Tindall, B., Hedrich, S., Rojas-Villalobos, C., Quatrini, R., et al. (2020). *Acidiferrimicrobium australe* gen. nov., sp. nov., an acidophilic and obligately heterotrophic, member of the *Actinobacteria* that catalyses dissimilatory oxido-reduction of iron isolated from metal-rich acidic water in Chile. *Int. J. Syst. Evol. Microbiol.* 70, 3348–3354. doi: 10.1099/ijsem.0.004179

González-Toril, E., Llobet-Brossa, E., Casamayor, E. O., Amann, R., and Amils, R. (2003). Microbial ecology of an extreme acidic environment, the Tinto River. *Appl. Environ. Microbiol.* 69, 4853–4865. doi: 10.1128/AEM.69.8.4853-4865.2003

Gyure, R. A., Konopka, A., Brooks, A., and Doemel, W. (1990). Microbial sulfate reduction in acidic (pH 3) strip-mine lakes. *FEMS Microbiol. Lett.* 73, 193–201. doi: 10.1111/j.1574-6968.1990.tb03941.x

Hallberg, K. B., González-Toril, E., and Johnson, D. B. (2010). *Acidithiobacillus ferrivorans*, sp. nov.; facultatively anaerobic, psychrotolerant iron-, and sulfur-oxidizing acidophiles isolated from metal mine-impacted environments. *Extremophiles* 14, 9–19. doi: 10.1007/s00792-009-0282-y

Hallberg, K. B., Grail, B. M., du Plessis, C. A., and Johnson, D. B. (2011a). Reductive dissolution of ferric iron minerals: a new approach for bio-processing nickel laterites. *Miner. Eng.* 24, 620–624. doi: 10.1016/j.mineng.2010.09.005

Hallberg, K. B., Hedrich, S., and Johnson, D. B. (2011b). Acidiferrobacter thiooxydans, gen. nov. sp. nov.; an acidophilic, thermo-tolerant, facultatively anaerobic iron- and sulfur-oxidizer of the family *Ectothiorhodospiraceae*. *Extremophiles* 15, 271–279. doi: 10.1007/s00792-011-0359-2

Hawkes, R. B., Franzmann, P. D., O'Hara, G., and Plumb, J. J. (2006). *Ferroplasma cupricumulans* sp. nov., a novel moderately thermophilic, acidophilic archaeon isolated from an industrial-scale chalcocite bioleach heap. *Extremophiles* 10, 525–530. doi: 10.1007/s00792-006-0527-y

Hedrich, S., and Johnson, D. B. (2013b). Aerobic and anaerobic oxidation of hydrogen by acidophilic bacteria. *FEMS Microbiol. Lett.* 349, 40–45. doi: 10.1111/1574-6968.12290

Hedrich, S., and Johnson, D. B. (2013a). *Acidithiobacillus ferridurans* sp. nov., an acidophilic iron-, sulfur- and hydrogen-metabolizing chemolithotrophic gammaproteobacterium. *Int. J. Syst. Evol. Microbiol.* 63, 4018–4025. doi: 10.1099/ijs.0.049759-0

Hedrich, S., Schlömann, M., and Johnson, D. B. (2011). The iron-oxidizing *proteobacteria*. *Microbiology* 157, 1551–1564. doi: 10.1099/mic.0.045344-0

Hofstetter, T. B., Heijman, C. G., Haderlein, S. B., Holliger, C., and Schwarzenbach, R. P. (1999). Complete reduction of TNT and other (poly)nitroaromatic compounds under iron-reducing subsurface conditions. *Environ. Sci. Technol.* 33, 1479–1487. doi: 10.1021/es9809760

Ilbert, M., and Bonnefoy, V. (2013). Insight into the evolution of the iron oxidation pathways. *Biochim. Biophys. Acta* 1827, 161–175. doi: 10.1016/j.bbabio.2012.10.001

Ingledew, W. J. (1982). *Thiobacillus ferrooxidans* the bioenergetics of an acidophilic chemolithotroph. *Biochim. Biophys. Acta* 683, 89–117. doi: 10.1016/0304-4173(82)90007-6

Itoh, T., Yamanoi, K., Kudo, T., Ohkuma, M., and Takashina, T. (2011). *Aciditerrimonas ferrireducens* gen. nov., sp. nov., an iron-reducing thermoacidophilic actinobacterium isolated from a solfataric field. *Int. J. Syst. Evol. Microbiol.* 61, 1281–1285. doi: 10.1099/ijs.0.023044-0

Jiang, V., Khare, S. D., and Banta, S. (2021). Computational structure prediction provides a plausible mechanism for electron transfer by the outer membrane protein Cyc2 from *Acidithiobacillus ferrooxidans*. *Protein Sci.* 30, 1640–1652. doi: 10.1002/pro.4106

Johnson, D. B. (1998). Biodiversity and ecology of acidophilic microorganisms. *FEMS Microbiol. Ecol.* 27, 307–317. doi: 10.1111/j.1574-6941.1998.tb00547.x

Johnson, D. B. (2009). "Extremophiles: acidic environments," in *Encyclopedia of Microbiology*, ed. M. Schaechter (Oxford: Elsevier), 107–126.

Johnson, D. B. (2010). "The biogeochemistry of biomining," in *Geomicrobiology: Molecular and Environmental Perspective*, eds L. L. Barton, M. Mandl, and A. Loy (Dordrecht: Springer), 401–426.

Johnson, D. B. (2014). Biomining - biotechnologies for extracting and recovering metals from ores and waste materials. *Curr. Opin. Biotechnol.* 30, 24–31. doi: 10.1016/j.copbio.2014.04.008

Johnson, D. B., Bacelar-Nicolau, P., Okibe, N., Thomas, A., and Hallberg, K. B. (2009). *Ferrimicrobium acidiphilum* gen. nov., sp. nov. and *Ferrithrix thermotolerans* gen. nov., sp. nov. heterotrophic, iron-oxidizing, extremely acidophilic actinobacteria. *Int. J. Syst. Evol. Microbiol.* 59, 1082–1089. doi: 10.1099/ijs.0.65409-0

Johnson, D. B., and Bridge, T. A. M. (2002). Reduction of ferric iron by acidophilic heterotrophic bacteria: evidence for constitutive and inducible enzyme systems in *Acidiphilium* spp. *J. Appl. Microbiol.* 92, 315–321. doi: 10.1046/j.1365-2672.2002.01535.x

Johnson, D. B., and du Plessis, C. A. (2015). Biomining in reverse gear: using bacteria to extract metals from oxidised ores. *Miner. Eng.* 75, 2–5. doi: 10.1016/j.mineng.2014.09.024

Johnson, D. B., Ghauri, M. A., and McGinness, S. (1993). Biogeochemical cycling of iron and sulphur in leaching environments. *FEMS Microbiol. Rev.* 11, 63–70. doi: 10.1111/j.1574-6976.1993.tb00268.x

Johnson, D. B., Grail, B. M., and Hallberg, K. B. (2013). A new direction for biomining: extraction of metals by reductive dissolution of oxidized ores. *Minerals* 3, 49–58. doi: 10.3390/min3010049

Johnson, D. B., and Hallberg, K. B. (2008). Carbon, iron and sulfur metabolism in acidophilic micro-organisms. *Adv. Microb. Physiol.* 54, 201–255.

Johnson, D. B., Hedrich, S., and Pakostova, E. (2017). Indirect redox transformations of iron, copper, and chromium catalyzed by extremely acidophilic bacteria. *Front. Microbiol.* 8:211. doi: 10.3389/fmicb.2017.00211

Johnson, D. B., Joulian, C., d'Hugues, P., and Hallberg, K. B. (2008). *Sulfobacillus benefaciens* sp. nov., an acidophilic facultative anaerobic *Firmicute* isolated from mineral bioleaching operations. *Extremophiles* 12, 789–798. doi: 10.1007/s00792-008-0184-4

Johnson, D. B., Kanao, T., and Hedrich, S. (2012). Redox transformations of iron at extremely low pH: fundamental and applied aspects. *Front. Microbiol.* 3:96. doi: 10.3389/fmicb.2012.00096

Johnson, D. B., and McGinness, S. (1991). Ferric iron reduction by acidophilic heterotrophic bacteria. *Appl. Environ. Microbiol.* 57, 207–211. doi: 10.1128/aem.57.1.207-211.1991

Johnson, D. B., and Pakostova, E. (2021). Dissolution of manganese (IV) oxide mediated by acidophilic bacteria, and demonstration that manganese (IV) can act as both a direct and indirect electron acceptor for iron-reducing *Acidithiobacillus* spp. *Geomicrobiol. J.* 38, 570–576. doi: 10.1080/01490451.2021.1903624

Johnson, D. B., and Quatrini, R. (2020). Acidophile microbiology in space and time. *Curr. Issues Mol. Biol.* 39, 63–76. doi: 10.21775/cimb.039.063

Johnson, D. B., Smith, S. L., and Santos, A. L. (2021). Bioleaching of transition metals from limonitic laterite deposits and reassessment of the multiple roles of sulfur-oxidizing acidophiles in the process. *Front. Microbiol.* 12:703177. doi: 10.3389/fmicb.2021.703177

Johnson, D. B., Stallwood, B., Kimura, S., and Hallberg, K. B. (2006). Isolation and characterization of *Acidicaldus organivorus*, gen. nov., sp. nov. a novel sulfur-oxidizing, ferric iron-reducing thermo-acidophilic heterotrophic *Proteobacterium*. *Arch. Microbiol.* 185, 212–221. doi: 10.1007/s00203-006-0087-7

Jones, R. M., Hedrich, S., and Johnson, D. B. (2013). *Acidocella aromatica* sp. nov. an acidophilic heterotrophic *Alphaproteobacterium* with unusual phenotypic traits. *Extremophiles* 17, 841–850. doi: 10.1007/s00792-013-0566-0

Jones, R. M., and Johnson, D. B. (2015). *Acidithrix ferrooxidans* gen. nov., sp. nov.; a filamentous and obligately heterotrophic, acidophilic member of the *Actinobacteria* that catalyzes dissimilatory oxido-reduction of iron. *Res. Microbiol.* 166, 111–120. doi: 10.1016/j.resmic.2015.01.003

Kai, M., Yano, T., Tamegai, H., Fukumori, Y., and Yamanaka, T. (1992). *Thiobacillus ferrooxidans* cytochrome *c* oxidase: purification, and molecular and enzymatic features. *J. Biochem.* 112, 816–821. doi: 10.1093/oxfordjournals. jbchem.a123982

Kappler, A., and Straub, K. L. (2005). Geomicrobiological cycling of iron. *Rev. Mineral. Geochem.* 59, 85–108. doi: 10.2138/rmg.2005.59.5

Karavaiko, G. I., Bogdanova, T. I., Tourova, T. P., Kondrat'eva, T. F., Tsaplina, I. A., Egorova, M. A., et al. (2005). Reclassification of 'Sulfobacillus thermosulfidooxidans subsp. thermotolerans' strain K1 as *Alicyclobacillus tolerans* sp. nov. and *Sulfobacillus disulfidooxidans* Dufresne et al. 1996 as *Alicyclobacillus disulfidooxidans* comb. nov., and emended description of the genus *Alicyclobacillus*. *Int. J. Syst. Evol. Microbiol.* 55, 941–947. doi: 10.1099/ijs. 0.63300-0

Kashefi, K., Holmes, D. E., Lovley, D. R., and Tor, J. M. (2004). "Potential importance of dissimilatory Fe(III)-reducing microorganisms in hot sedimentary environments," in *The Subseafloor Biosphere at Mid-Ocean Ridges*, eds W. S. Wilcock, E. F. DeLong, D. S. Kelley, J. A. Baross, and S. Craig Cary (Washington, DC: American Geophysical Union), 199–211.

Kucera, J., Bouchal, P., Cerna, H., Potesil, D., Janiczek, O., Zdrahal, Z., et al. (2012). Kinetics of anaerobic elemental sulfur oxidation by ferric iron in *Acidithiobacillus ferrooxidans* and protein identification by comparative 2-DE-MS/MS. *Antonie Van Leeuwenhoek* 101, 561–573. doi: 10.1007/s10482-011-9670-2

Kucera, J., Bouchal, P., Lochman, J., Potesil, D., Janiczek, O., Zdrahal, Z., et al. (2013). Ferrous iron oxidation by sulfur-oxidizing *Acidithiobacillus ferrooxidans* and analysis of the process at the levels of transcription and protein synthesis. *Antonie Van Leeuwenhoek* 103, 905–919. doi: 10.1007/s10482-012-9872-2

Kucera, J., Lochman, J., Bouchal, P., Pakostova, E., Mikulasek, K., Hedrich, S., et al. (2020). A model of aerobic and anaerobic metabolism of hydrogen in the extremophile *Acidithiobacillus ferrooxidans*. *Front. Microbiol.* 11:610836. doi: 10.3389/fmicb.2020.610836

Kucera, J., Pakostova, E., Janiczek, O., and Mandl, M. (2015). Changes in *Acidithiobacillus ferrooxidans* ability to reduce ferric iron by elemental sulfur. *Adv. Mat. Res.* 1130, 97–100.

Kucera, J., Pakostova, E., Lochman, J., Janiczek, O., and Mandl, M. (2016a). Are there multiple mechanisms of anaerobic sulfur oxidation with ferric iron in *Acidithiobacillus ferrooxidans*? *Res. Microbiol.* 167, 357–366. doi: 10.1016/j. resmic.2016.02.004

Kucera, J., Sedo, O., Potesil, D., Janiczek, O., Zdrahal, Z., and Mandl, M. (2016b). Comparative proteomic analysis of sulfur-oxidizing *Acidithiobacillus ferrooxidans* CCM 4253 cultures having lost the ability to couple anaerobic elemental sulfur oxidation with ferric iron reduction. *Res. Microbiol.* 167, 587–594. doi: 10.1016/j.resmic.2016.06.009

Küsel, K., Roth, U., and Drake, H. L. (2002). Microbial reduction of Fe(III) in the presence of oxygen under low pH conditions. *Environ. Microbiol.* 4, 414–421. doi: 10.1046/j.1462-2920.2002.00314.x

Lovley, D. R. (1991). Dissimilatory Fe(III) and Mn(IV) reduction. *Microbiol. Rev.* 55, 259–287. doi: 10.1128/mr.55.2.259-287.1991

Lu, S., Gischkat, S., Reiche, M., Akob, D. M., Hallberg, K. B., and Küsel, K. (2010). Ecophysiology of Fe-cycling bacteria in acidic sediments. *Appl. Environ. Microbiol.* 76, 8174–8183. doi: 10.1128/AEM.01931-10

Malarte, G., Leroy, G., Lojou, E., Abergel, C., Bruschi, M., and Giudici-Orticoni, M. T. (2005). Insight into molecular stability and physiological properties of the diheme cytochrome CYC$_{41}$ from the acidophilic bacterium *Acidithiobacillus ferrooxidans*. *Biochemistry* 44, 6471–6481. doi: 10.1021/bi048425b

Matin, A. (1999). "pH homeostasis in acidophiles," in *Novartis Foundation Symposium 221 – Bacterial Responses to pH*, eds D. J. Chadwick and G. Cardew (New York, NY: Wiley). doi: 10.1002/9780470515631.ch10

Marrero, J., Coto, O., Goldmann, S., Graupner, T., and Schippers, A. (2015). Recovery of nickel and cobalt from laterite tailings by reductive dissolution under aerobic conditions using *Acidithiobacillus* species. *Environ. Sci. Technol.* 49, 6674–6682. doi: 10.1021/acs.est.5b00944

Marrero, J., Coto, O., and Schippers, A. (2017). Anaerobic and aerobic reductive dissolutions of iron-rich nickel laterite overburden by *Acidithiobacillus*. *Hydrometallurgy* 168, 49–55. doi: 10.1016/j.hydromet.2016.08.012

Masaki, Y., Tsutsumi, K., and Okibe, N. (2018). Iron redox transformation by the thermo-acidophilic archaea from the genus *Sulfolobus*. *Geomicrobiol. J.* 35, 757–767. doi: 10.1080/01490451.2018.1465491

Nancucheo, I., Grail, B. M., Hilario, F., du Plessis, C. A., and Johnson, D. B. (2014). Extraction of copper from an oxidized (lateritic) ore using bacterially catalysed reductive dissolution. *Appl. Microbiol. Biotechnol.* 98, 6297–6305. doi: 10.1007/s00253-014-5687-6

Nancucheo, I., Oliveira, A., Lopes, M., and Johnson, D. B. (2019). Bioreductive dissolution as a pretreatment for recalcitrant rare-earth phosphate minerals associated with lateritic ores. *Minerals* 9:136. doi: 10.3390/min9030136

Norris, P. R., Falagán, C., Moya-Beltrán, A., Castro, M., Quatrini, R., and Johnson, D. B. (2020). *Acidithiobacillus ferrianus* sp. nov. an ancestral extremely acidophilic and facultatively anaerobic chemolithoautotroph. *Extremophiles* 24, 329–337. doi: 10.1007/s00792-020-01157-1

Norris, P. R., Gould, O. J. P., and Ogden, T. J. (2015). Iron solubilization during anaerobic growth of acidophilic microorganisms with a polymetallic sulfide ore. *Miner. Eng.* 75, 77–84. doi: 10.1016/j.mineng.2014.12.004

Norris, P. R., Laigle, L., and Slade, S. (2018). Cytochromes in anaerobic growth of *Acidithiobacillus ferrooxidans*. *Microbiology* 164, 383–394. doi: 10.1099/mic.0. 000616

Ohmura, N., Sasaki, K., Matsumoto, N., and Saiki, H. (2002). Anaerobic respiration using Fe^{3+}, S^0, and H$_2$ in the chemolithoautotrophic bacterium *Acidithiobacillus ferrooxidans*. *J. Bacteriol.* 184, 2081–2087. doi: 10.1128/JB.184. 8.2081-2087.2002

Osorio, H., Mangold, S., Denis, Y., Nancucheo, I., Esparza, M., Johnson, D. B., et al. (2013). Anaerobic sulfur metabolism coupled to dissimilatory iron reduction in the extremophile *Acidithiobacillus ferrooxidans*. *Appl. Environ. Microbiol.* 79, 2172–2181. doi: 10.1128/AEM.03057-12

Ottow, J. C. G., and Glathe, H. (1971). Isolation and identification of iron-reducing bacteria from gley soils. *Soil Biol. Biochem.* 3, 43–55. doi: 10.1016/ 0038-0717(71)90030-7

Plumb, J. J., Haddad, C. M., Gibson, J. A. E., and Franzmann, P. D. (2007). *Acidianus sulfidivorans* sp. nov., an extremely acidophilic, thermophilic archaeon isolated from a solfatara on Lihir Island, Papua New Guinea, and emendation of the genus description. *Int. J. Syst. Evol. Microbiol.* 57, 1418–1423. doi: 10.1099/ijs.0.64846-0

Pronk, J. T., Bruyn, J. C., de Bos, P., and Kuenen, J. G. (1992). Anaerobic growth of *Thiobacillus ferrooxidans*. *Appl. Environ. Microbiol.* 58:2227. doi: 10.1128/aem. 58.7.2227-2230.1992

Pronk, J. T., Liem, K., Bos, P., and Kuenen, J. G. (1991). Energy transduction by anaerobic ferric iron respiration in *Thiobacillus ferrooxidans*. *Appl. Environ. Microbiol.* 57:2063. doi: 10.1128/aem.57.7.2063-2068.1991

Quatrini, R., Appia-Ayme, C., Denis, Y., Ratouchniak, J., Veloso, F., Valdes, J., et al. (2006). Insights into the iron and sulfur energetic metabolism of *Acidithiobacillus ferrooxidans* by microarray transcriptome profiling. *Hydrometallurgy* 83, 263–272. doi: 10.1016/j.hydromet.2006.03.030

Quatrini, R., and Johnson, D. B. (2016). *Acidophiles: Life in Extremely Acidic Environments*. Norfolk: Caister Academic Press.

Rohwerder, T. (2002). *Untersuchungen zur Enzymatischen Oxidation von Elementarschwefel bei Acidophilen Laugungsbakterien*. Thesis. Hamburg: Universität Hamburg.

Sakai, H. D., and Kurosawa, N. (2018). *Saccharolobus caldissimus* gen. nov., sp. nov., a facultatively anaerobic iron-reducing hyperthermophilic archaeon isolated from an acidic terrestrial hot spring, and reclassification of *Sulfolobus solfataricus* as *Saccharolobus solfataricus* comb. nov. and *Sulfolobus shibatae* as *Saccharolobus shibatae* comb. nov. *Int. J. Syst. Evol. Microbiol.* 68, 1271–1278. doi: 10.1099/ijsem.0.002665

Sand, W. (1989). Ferric iron reduction by *Thiobacillus ferrooxidans* at extremely low pH-values. *Biogeochemistry* 7, 195–201. doi: 10.1007/BF00004217

Santos, A. L., Dybowska, A., Schofield, P. F., Herrington, R. J., and Johnson, D. B. (2020). Sulfur-enhanced reductive bioprocessing of cobalt-bearing materials for base metals recovery. *Hydrometallurgy* 195:105396. doi: 10.1016/j.hydromet. 2020.105396

Slobodkin, A. I. (2005). Thermophilic microbial metal reduction. *Microbiology* 74, 501–514. doi: 10.1007/s11021-005-0096-6

Smith, S. L., Grail, B. M., and Johnson, D. B. (2017). Reductive bioprocessing of cobalt-bearing limonitic laterites. *Miner. Eng.* 106, 86–90. doi: 10.1016/j.mineng.2016.09.009

Smith, S. L., and Johnson, D. B. (2018). Growth of *Leptospirillum ferriphilum* in sulfur medium in co-culture with *Acidithiobacillus caldus*. *Extremophiles* 22, 327–333. doi: 10.1007/s00792-018-1001-3

Stolz, A. (2017). "Acidophile," in *Extremophile Mikroorganismen*, ed. A. Stolz (Berlin: Springer), 79–106.

Sugio, T., Domatsu, C., Munakata, O., Tano, T., and Imai, K. (1985). Role of a ferric ion-reducing system in sulfur oxidation of *Thiobacillus ferrooxidans*. *Appl. Environ. Microbiol.* 49, 1401–1406. doi: 10.1128/aem.49.6.1401-1406.1985

Sugio, T., Mizunashi, W., Inagaki, K., and Tano, T. (1987). Purification and some properties of sulfur:ferric ion oxidoreductase from *Thiobacillus ferrooxidans*. *J. Bacteriol.* 169, 4916–4922. doi: 10.1128/jb.169.11.4916-4922.1987

Sugio, T., Taha, T. M., and Takeuchi, F. (2009). Ferrous iron production mediated by tetrathionate hydrolase in tetrathionate-, sulfur-, and iron-grown *Acidithiobacillus ferrooxidans* ATCC 23270 cells. *Biosci. Biotechnol. Biochem.* 73, 1381–1386. doi: 10.1271/bbb.90036

Sugio, T., Wada, K., Mori, M., Inagaki, K., and Tano, T. (1988). Synthesis of an iron-oxidizing system during growth of *Thiobacillus ferrooxidans* on sulfur-basal salts medium. *Appl. Environ. Microbiol.* 54, 150–152. doi: 10.1128/aem.54.1.150-152.1988

Tsuboi, K., Sakai, H. D., Nur, N., Stedman, K. M., Kurosawa, N., and Suwanto, A. (2018). *Sulfurisphaera javensis* sp. nov., a hyperthermophilic and acidophilic archaeon isolated from Indonesian hot spring, and reclassification of *Sulfolobus tokodaii* Suzuki et al. 2002 as *Sulfurisphaera tokodaii* comb. nov. *Int. J. Syst. Evol. Microbiol.* 68, 1907–1913. doi: 10.1099/ijsem.0.002765

Walker, J. C. G. (1987). Was the Archaean biosphere upside down? *Nature* 329, 710–712. doi: 10.1038/329710a0

Wang, R., Lin, J.-Q., Liu, X.-M., Pang, X., Zhang, C.-J., Yang, C.-L., et al. (2019). Sulfur oxidation in the acidophilic autotrophic *Acidithiobacillus* spp. *Front. Microbiol.* 9:3290. doi: 10.3389/fmicb.2018.03290

Yarzábal, A., Brasseur, G., Ratouchniak, J., Lund, K., Lemesle-Meunier, D., DeMoss, J. A., et al. (2002). The high-molecular-weight cytochrome *c* Cyc2 of *Acidithiobacillus ferrooxidans* is an outer membrane protein. *J. Bacteriol.* 184, 313–317. doi: 10.1128/JB.184.1.313-317.2002

Yoshida, N., Nakasato, M., Ohmura, N., Ando, A., Saiki, H., Ishii, M., et al. (2006). *Acidianus manzaensis* sp. nov., a novel thermoacidophilic archaeon growing autotrophically by the oxidation of H_2 with the reduction of Fe^{3+}. *Curr. Microbiol.* 53, 406–411. doi: 10.1007/s00284-006-0151-1

Zhang, R., Hedrich, S., Jin, D., Breuker, A., and Schippers, A. (2021). *Sulfobacillus harzensis* sp. nov., an acidophilic bacterium inhabiting mine tailings from a polymetallic mine. *Int. J. Syst. Evol. Microbiol.* 71:4871. doi: 10.1099/ijsem.0.004871

Zhang, R., Hedrich, S., Ostertag-Henning, C., and Schippers, A. (2018). Effect of elevated pressure on ferric iron reduction coupled to sulfur oxidation by biomining microorganisms. *Hydrometallurgy* 178, 215–223. doi: 10.1016/j.hydromet.2018.05.003

Zhou, H., Zhang, R., Hu, P., Zeng, W., Xie, Y., Wu, C., et al. (2008). Isolation and characterization of *Ferroplasma thermophilum* sp. nov., a novel extremely acidophilic, moderately thermophilic archaeon and its role in bioleaching of chalcopyrite. *J. Appl. Microbiol.* 105, 591–601. doi: 10.1111/j.1365-2672.2008.03807.x

Seroprevalence, Prevalence, and Genomic Surveillance: Monitoring the Initial Phases of the SARS-CoV-2 Pandemic in Betim, Brazil

Ana Valesca Fernandes Gilson Silva[1]*[†], Diego Menezes[2,3][†],
Filipe Romero Rebello Moreira[4][†], Octávio Alcântara Torres[1],
Paula Luize Camargos Fonseca[2,3], Rennan Garcias Moreira[5], Hugo José Alves[2,3],
Vivian Ribeiro Alves[1], Tânia Maria de Resende Amaral[1], Adriano Neves Coelho[1],
Júlia Maria Saraiva Duarte[3], Augusto Viana da Rocha[1],
Luiz Gonzaga Paula de Almeida[6], João Locke Ferreira de Araújo[2,3],
Hilton Soares de Oliveira[1], Nova Jersey Cláudio de Oliveira[1], Camila Zolini[4],
Jôsy Hubner de Sousa[7], Elizângela Gonçalves de Souza[1], Rafael Marques de Souza[2,3],
Luciana de Lima Ferreira[2,3], Alexandra Lehmkuhl Gerber[6],
Ana Paula de Campos Guimarães[6], Paulo Henrique Silva Maia[1],
Fernanda Martins Marim[2,3], Lucyene Miguita[8], Cristiane Campos Monteiro[1],
Tuffi Saliba Neto[1], Fabrícia Soares Freire Pugêdo[1], Daniel Costa Queiroz[2,3],
Damares Nigia Alborguetti Cuzzuol Queiroz[1], Luciana Cunha Resende-Moreira[9],
Franciele Martins Santos[7], Erika Fernanda Carlos Souza[1], Carolina Moreira Voloch[4],
Ana Tereza Vasconcelos[6], Renato Santana de Aguiar[2,3,10]* and Renan Pedra de Souza[2,3]*

[1] Escola de Saúde Pública de Betim, Betim, Brazil, [2] Programa de Pós Graduação em Genética, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [3] Laboratório de Biologia Integrativa, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [4] Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, [5] Centro de Laboratórios Multiusuários, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [6] Laboratório Nacional de Computação Científica, Petropólis, Brazil, [7] Programa de Pós-graduação em Biologia Celular, Departamento de Morfologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [8] Departamento de Patologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [9] Departamento de Botânica, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, [10] Instituto D'Or de Pesquisa e Ensino (IDOR), Rio de Janeiro, Brazil

The COVID-19 pandemic has created an unprecedented need for epidemiological monitoring using diverse strategies. We conducted a project combining prevalence, seroprevalence, and genomic surveillance approaches to describe the initial pandemic stages in Betim City, Brazil. We collected 3239 subjects in a population-based age-, sex- and neighborhood-stratified, household, prospective; cross-sectional study divided into three surveys 21 days apart sampling the same geographical area. In the first survey, overall prevalence (participants positive in serological or molecular tests) reached 0.46% (90% CI 0.12–0.80%), followed by 2.69% (90% CI 1.88–3.49%) in the second survey and 6.67% (90% CI 5.42–7.92%) in the third. The underreporting reached 11, 19.6, and 20.4 times in each survey. We observed increased odds to test positive in females compared to males (OR 1.88 95% CI 1.25–2.82), while the single best predictor for positivity was ageusia/anosmia (OR 8.12, 95% CI 4.72–13.98). Thirty-five SARS-CoV-2

genomes were sequenced, of which 18 were classified as lineage B.1.1.28, while 17 were B.1.1.33. Multiple independent viral introductions were observed. Integration of multiple epidemiological strategies was able to adequately describe COVID-19 dispersion in the city. Presented results have helped local government authorities to guide pandemic management.

INTRODUCTION

Since its emergence in December 2019, the new human coronavirus has had a tremendous impact on humanity due to the pandemic nature of its infection, called COVID-19 (Zhou et al., 2020). The SARS-CoV-2 pathogen was described on January 24, 2020. In Brazil, the first case of COVID-19 was reported on February 26, 2020, in the city of São Paulo (Araujo et al., 2020). The virus spread rapidly, and the country had the highest number of cases and deaths in Latin America, experiencing its first peak wave in late July 2020. Although most patients were identified in the most prominent Brazilian cities, São Paulo and Rio de Janeiro, dispersion to other municipalities was quickly reported. Betim, a town located in the Minas Gerais State in Brazil with an estimated population of 439,340 in 2019, had its first reported SARS-CoV-2 case on March 23, 2020, in two patients returning from Europe. Two months later, on May 23, 2020, only 73 confirmed cases had been reported, although 4380 suspected cases were identified in public databases indicating limited testing availability.

Brazilian public healthcare system has prioritized testing subjects with symptoms due to scarce diagnostic tests, particularly in the early days of the pandemic. Since data suggest that symptomatic cases represent a fraction of persons infected with SARS-CoV-2, official statistics were expected to be underestimated (Wu et al., 2020). Several aspects may influence COVID-19 symptom presentation (Araújo et al., 2021; Rossi et al., 2021). Epidemiological surveillance using prevalence studies is needed to evaluate the true extent of SARS-CoV-2 dispersion, significantly extending testing to asymptomatic subjects. Combining serological and molecular tests may be a more robust strategy to uncover viral diffusion in a territory, avoiding each test's kinetic detection limitations. Valid prevalence and seroprevalence estimates for a population rely on two major factors: (i) a representative population sample and (ii) accurate diagnostic testing (Byambasuren et al., 2021).

While the epidemiological investigation is essential for controlling COVID-19, genomic surveillance is crucial. Robust SARS-CoV-2 variant monitoring can track viral evolution, detect new variants, describe patterns and clusters of transmission, outbreak tracking, among others. Therefore, it can provide actionable information on implementing a more targeted public health strategy that addresses local priorities through stakeholder engagement and mitigation efforts (Robishaw et al., 2021). We conducted a study combining seroprevalence, prevalence, and genomic surveillance approaches to understand the SARS-CoV-2 epidemic spread in Betim city.

MATERIALS AND METHODS

Seroprevalence and Prevalence

The Research Ethics Committee approved the present experiment under protocol CAAE 31459220.2.0000.5651. We conducted a population-based age-, sex- and neighborhood-stratified, household, prospective; cross-sectional study repeated every 21 days in the same geographic area to determine the extent of SARS-CoV-2 transmission in Betim, Minas Gerais, Brazil (**Figures 1A,B**). All populated areas in the city were sampled. Three surveys were held: June 3–5, June 23–25, and July 13–15, 2020. The sample size ($n = 1,080$ each survey) was estimated considering dichotomous outcome (positive or negative), the population of 439,340 inhabitants, the confidence level of 90%, the maximum margin of error of 2.5%, and lack of *a priori* information on the prevalence of SARS-COV-2 in the municipality's population (the latter represented by $p = q = 0.5$) and using the equation below:

$$n = \frac{Z_{\frac{\alpha}{2}}^2 \cdot \hat{p} \cdot \hat{q} \cdot N}{E^2 \cdot (N - 1) + Z_{\frac{\alpha}{2}}^2 \cdot \hat{p} \cdot \hat{q}}$$

Random sampling was employed to ensure representativeness of the population, stratified by sex, age (0–5; 6–19; 20–39; 40–59, and 60 years or older) and city neighborhoods (Centro, Alterosas, Imbiruçu, Norte, Teresópolis, PTB, Citrolândia, Vianópolis, Icaivera, and Petrovale). Every census tract (population stratum created by Governmental agencies) was sampled with at least one address. In case of refusal or closed households, the closest home was selected. Thirty-six teams (one driver, one nurse, and one community health worker) worked on active sampling subjects in 1080 addresses during 3 days. Clinical and epidemiological data were obtained using a questionnaire during interviews with participants or their legal guardians who signed the Informed Consent. Biological samples were collected using a nasal swab to conduct RT-PCR and capillary blood obtained by fingerstick for the serological test.

RT-PCR to detect SARS-CoV-2 RNA was initially conducted in pools of 10 samples (Lohse et al., 2020). Whenever pools were positive, individual samples were examined. Molecular diagnosis was established according to the CDC 2019-Novel Coronavirus Real-Time RT-PCR Diagnostic Panel (N1, N2, and RNP primers). Serological tests were conducted using the SARS-CoV-2 Antibody Test (Guangzhou Wondfo Biotech Co., Ltd. Guangzhou, Guangdong, China) that detects IgM/IgG

FIGURE 1 | Sampling strategy throughout Betim city. **(A)** Betim's geographical location (white area) in the Minas Gerais State (blue area). **(B)** Sampling locations for each survey (*n* = 1080). We sampled all populated areas in the town. Areas without points indicate non-populated areas.
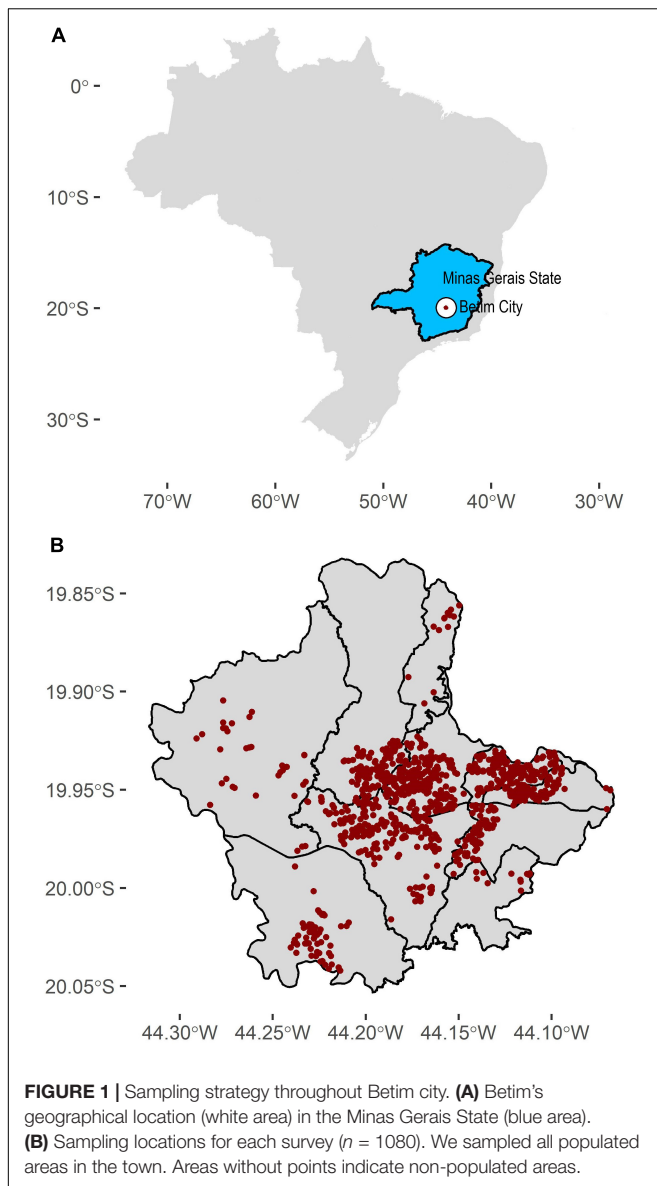
antibodies. The same test was used in a previous study in Brazil (Hallal et al., 2020). Reported sensitivity is 86.43% (95% CI: 82.41∼89.58%) and specificity 99.57% (95% CI: 97.63∼99.92%). We have validated antibody tests using serum samples from subjects who were SARS-CoV-2 positive confirmed with RT-PCR.

Associations of each variable of interest with surveys and positive status were assessed using chi-square tests. Odds ratios were estimated using logistic regression with the *glm* function. Spatial geostatistical modeling and prediction were carried out using the *gstat* and *predict* functions from the *gstat* package. All analyses were carried out in R software (version 4.1.1).

Genomic Surveillance

Whole viral genome amplification and DNA library preparation was carried out as described elsewhere (Moreira et al., 2021). Briefly, QIAseq SARS-CoV-2 Primer Panel—QIAGEN kit was

used to amplify positive samples, following manufacturer instructions. In total, 39 of the 84 detectable samples were eligible for library preparation based on their CTs ≤ 30. Library concentration was measured using the QIAseq Library Quant Assay—QIAGEN kit, and the fragment integrity and size were evaluated using Bioanalyzer (Agilent Technologies, Waldbronn, DE). Sequencing was carried out on a MiSeq (Illumina, San Diego, CA, United States).

The raw data generated were filtered by Trimmomatic v0.39 (Bolger et al., 2014), which trimmed low-quality bases (Phred score < 30) and removed short reads (< 50 nucleotides) as well as adapters and primer sequences. Reads were then mapped against the SARS-CoV-2 reference genome (accession number: NC_045512.2) with Bowtie2 (Langmead et al., 2009). The resulting BAM files were manipulated with SAMtools, BCFtools (Li et al., 2009), and BEDtools (Quinlan and Hall, 2010) to generate consensus genome sequences. Bases with less than 10x sequencing depth were masked. In total, 35 of the 39 genome sequences presented coverage greater than 79% and average sequencing depth greater than 200x (**Supplementary Table 1**). The 35 consensus genome sequences were submitted to the PANGOLIN 2.0 lineage classification tool (database version February 2, 2021) (Rambaut et al., 2020).

To confirm the PANGOLIN identification and further contextualize the diversity of lineages circulating in Betim, we performed a set of phylogenetic analyses. First, a global dataset was assembled from a subset of high-quality data available on GISAID and the newly generated genomes (*n* = 3,814). This dataset contained all Brazilian sequences and one per week for each country, as available on GISAID until January 12, 2021. These sequences were aligned with MAFFT v7.475 (Katoh and Standley, 2013), and a maximum likelihood tree was inferred on IQ-Tree 2 (Minh et al., 2020), under the GTR+F+I+G4 model (Tavaré, 1986; Yang, 1994). Shimoidara-Hasegawa approximate likelihood ratio test (SH-aLRT) was used to assess branches' statistical support (Guindon et al., 2010).

Two subsets of the previous dataset were assembled to explore the temporal dynamics of introduction and circulation of SARS-CoV-2 in Betim, comprehending sequences belonging to lineages B.1.1.28 (*n* = 258) and B.1.1.33 (*n* = 284). The parameterization of the phylogeographic model was set to be primarily informative concerning introductions of SARS-CoV-2 in Betim. Therefore, we set the model with six discrete categories: Betim City, Minas Gerais State, Rio de Janeiro State, São Paulo State, other Brazilian States, and foreign sequences. These locations were represented by 18, 2, 22, 71, 79, and 66 sequences in dataset B.1.1.28 while B.1.1.33 dataset composition was 17, 20, 53, 52, 73, and 69 sequences from each region, respectively.

Maximum likelihood trees were inferred from these datasets, and their temporal signal was evaluated with tempest v1.5.3 (Rambaut et al., 2016). Time scaled phylogenies were then inferred from these datasets with BEAST v1.10.4 (Suchard et al., 2018), using: (*i*) the HKY+I+G4 nucleotide substitution model (Yang, 1994), (*ii*) the strict molecular clock model, (*iii*) the non-parametric coalescent skygrid tree prior (Gill et al., 2013) and (*iv*) a symmetric discrete phylogeographic model (Lemey et al., 2009). A normal prior distribution (mean = 1.13×10^{-3}; std = $5.1 \times$

10^{-4}) on clock rate was assumed, based on a previous estimate (Candido et al., 2020). The cutoff values of the skygrid tree prior were set based on the previously estimated dates for the emergence of each lineage (Candido et al., 2020). The number of grids of the tree priors was set to match the approximate number of weeks comprehended between the estimated dates for lineages' emergence and the dates of the most recently sampled sequences (41 weeks, both datasets). Two and three independent chains of 200 million generations sampling every 10,000 states were performed for datasets B.1.1.33 and B.1.1.28, respectively. Tracer v1.7.1 (Rambaut et al., 2018) was used to verify mixing and convergence of chains (effective sample size > 200 for all parameters), which were then combined with logcombiner v1.10.4 after 10% burning removal. Maximum clade credibility trees were generated with treeannotator v1.10.4. All logs and trees are available in https://github.com/LBI-lab/SARS-CoV-2_phylogenies.git.

RESULTS

Seroprevalence and Prevalence

Evaluation of clinical and epidemiological data showed no significant difference for the presence of any prior health condition across surveys (pneumopathy, chronic neurological disease, pregnant, postpartum, chronic cardiovascular disease, chronic kidney disease, obesity, asthma, immunodepression, chronic liver disease, diabetes, hypertension, transplanted, cancer or any comorbidity) indicating proper sampling was conducted since there was no reason to find significant differences in the period (**Table 1**). Four symptoms (cough, sore throat, myalgia, and rhinorrhea) and contact with a symptomatic person increased while international travel decreased. Prevalence and seroprevalence increased across surveys.

Sampling was conducted in the early stages of the pandemic, as seen in the number of absolute reported cases (**Figure 2A**). Cumulative confirmed cases were underestimated (**Figure 2B**). In the first survey, overall prevalence (participants positive in serological or molecular tests) reached 0.46% (90% CI 0.12–0.80%), followed by 2.69% (90% CI 1.88–3.49%) in the second survey and 6.67% (90% CI 5.42–7.92%) in the third. The underreporting was obtained by the difference between survey prevalence and official data, and its magnitude reached 11, 19.6, and 20.4 times (**Figure 2B**). Overall prevalence increase was observed across most administrative regions (**Figures 2C,D**). Active transmission areas (RT-PCR positive participants) were observed increasing across time (**Figures 3A–C**). By the third survey, almost all populated city areas were likely to have viral circulation (**Figure 3C**).

We have also evaluated whether clinical and epidemiological variables were associated with molecular or serological test positivity (**Table 2**). Several significant results were observed, mostly with reported symptoms (fever, cough, sore throat, dyspnoea, myalgia, rhinorrhea, respiratory discomfort, nausea/vomit, headache, prostration, ageusia/ anosmia). We also observed increased odds to test positive in females compared to males (OR 1.88 95% CI 1.25–2.82) and clear enrichment

of positive cases in certain city regions (e.g., Imbiruçu and Terezópolis). Surprisingly, people with obesity were more likely to be positive (OR 3.33, 95% CI 1.68–6.59). The single best predictor for positivity was ageusia/anosmia (OR 8.12, 95% CI 4.72–13.98). Non-significant associations were also found (**Supplementary Table 2**).

Genomic Viral Surveillance

In total, 35 novel SARS-CoV-2 genome sequences were obtained (GISAID EPI_ISL_5416087-5416121). The sequences were classified by PANGOLIN 2.0 to assess the genetic diversity of SARS-CoV-2 circulating in Betim. 18 of the 35 genomes were classified as lineage B.1.1.28, while 17 were B.1.1.33 (*Probability* = 1.0). Further, a maximum likelihood tree was inferred from the global dataset GISAID (Shu and McCauley, 2017). No difference in the dispersion pattern was observed across lineages.

The analysis supported these results, revealing sequences from the Betim cluster within several clades of these lineages confirming the circulation of (B.1.1.28 and B.1.1.33 during the first wave of COVID-19 pandemics in the city (**Figure 4**). The spread of Betim sequences across the tree suggests multiple independent introductions occurred in the town. Further, eight clades majorly composed by Betim sequences were inferred with variable degrees of statistical support (median SH-aLRT = 82.75, range: 0–100), suggesting the occurrence of local transmission in the city after initial introduction events. In addition to these clusters, nine introductions supported by single sequences have also been detected. Most Betim sequences or clusters are closely related to sequences from Rio de Janeiro and São Paulo, two neighboring States connected by highways to Minas Gerais. To formally assess the dynamics of introduction and spread of SARS-CoV-2 in Betim, separated datasets for lineages B.1.1.28 and B.1.1.33 were evaluated. Regression between sampling times and genetic distances revealed both datasets had moderate temporal signal (B.1.1.28: R^2 = 0.49; B.1.1.33: R^2 = 0.58), justifying molecular clock analysis.

The time-scaled phylogeographic analysis performed with dataset B.1.1.28 suggests this lineage emerged on February 22, 2020, in São Paulo (95% highest posterior density, HPD: February 11, 2020–March 05, 2020; geographic model posterior probability, PP = 0.91), later spreading to other Brazilian states (**Figure 5A**). The phylogeny reveals that two introduction events, dated between April 19, 2020 (95% HPD: April 17, 2020–May 11, 2020) and April 22, 2020 (95% HPD: April 20, 2020–May 27, 2020), led to the emergence of Betim clusters (harboring between two and six sequences). Additionally, four introductions related to single sequences have been detected. The phylogeographic model suggests that three introductions occurred from another Brazilian region to Betim, in addition to other single events from RJ, another one from SP, and another from foreign sequences. All events presented high statistical support (PP > 92% for all events).

The phylogeographic reconstruction performed for dataset B.1.1.33 infers the origin of this lineage on February 06, 2020, in Rio de Janeiro (95% HPD: January 14, 2020–February 25, 2020, PP = 0.78). The model supports the occurrence of many

TABLE 1 | Clinical and epidemiological data obtained from participants.

Variable	Level	Overall n (%)	First survey n (%)	Second survey n (%)	Third survey n (%)	*p*-value
Administrative Regions	Alterosas	634 (19.6%)	198 (18.4%)	218 (20.2%)	218 (20.2%)	0.9584
	Citrolândia	219 (6.8%)	83 (7.7%)	68 (6.3%)	68 (6.3%)	
	Icaivera	62 (1.9%)	20 (1.9%)	21 (1.9%)	21 (1.9%)	
	Imbiruçu	565 (17.4%)	183 (17.0%)	191 (17.7%)	191 (17.7%)	
	Norte	333 (10.3%)	111 (10.3%)	111 (10.3%)	111 (10.3%)	
	Petrovale	41 (1.3%)	13 (1.2%)	14 (1.3%)	14 (1.3%)	
	PTB	290 (9.0%)	108 (10.0%)	91 (8.4%)	91 (8.4%)	
	Sede	583 (18.0%)	201 (18.6%)	191 (17.7%)	191 (17.7%)	
	Terezópolis	319 (9.8%)	109 (10.1%)	105 (9.7%)	105 (9.7%)	
	Vianópolis	193 (6.0%)	53 (4.9%)	70 (6.5%)	70 (6.5%)	
Sex	Female	1628 (50.3%)	548 (50.8%)	536 (49.6%)	544 (50.4%)	0.8619
Age range	0–5	217 (6.7%)	71 (6.6%)	73 (6.8%)	73 (6.8%)	1.0000
	6–19	650 (20.1%)	218 (20.2%)	217 (20.1%)	215 (19.9%)	
	20–39	1067 (32.9%)	354 (32.8%)	355 (32.9%)	358 (33.1%)	
	40–59	871 (26.9%)	291 (27.0%)	289 (26.8%)	291 (26.9%)	
	Above 60	434 (13.4%)	145 (13.4%)	146 (13.5%)	143 (13.2%)	
Pneumopathy	Yes	30 (0.9%)	7 (0.6%)	13 (1.2%)	10 (0.9%)	0.4042
Chronic neurological disease	Yes	39 (1.2%)	16 (1.5%)	10 (0.9%)	13 (1.2%)	0.4948
Pregnant	Yes	28 (0.9%)	10 (0.9%)	11 (1.0%)	7 (0.6%)	0.6257
Postpartum	Yes	9 (0.3%)	2 (0.2%)	3 (0.3%)	4 (0.4%)	0.7165
Chronic cardiovascular disease	Yes	96 (3.0%)	34 (3.2%)	39 (3.6%)	23 (2.1%)	0.1154
Chronic kidney disease	Yes	50 (1.5%)	24 (2.2%)	12 (1.1%)	14 (1.3%)	0.0799
Obesity	Yes	105 (3.2%)	33 (3.1%)	37 (3.4%)	35 (3.2%)	0.8903
Asthma	Yes	173 (5.3%)	65 (6.0%)	58 (5.4%)	50 (4.6%)	0.3537
Immunodepression	Yes	22 (0.7%)	9 (0.8%)	5 (0.5%)	8 (0.7%)	0.5507
Chronic liver disease	Yes	15 (0.5%)	4 (0.4%)	7 (0.6%)	4 (0.4%)	0.5478
Diabetes	Yes	228 (7.0%)	78 (7.2%)	74 (6.9%)	76 (7.0%)	0.9430
Hypertension	Yes	563 (17.4%)	190 (17.6%)	186 (17.2%)	187 (17.3%)	0.9698
Transplanted	Yes	4 (0.1%)	2 (0.2%)	1 (0.1%)	1 (0.1%)	0.7780
Cancer	Yes	23 (0.7%)	10 (0.9%)	8 (0.7%)	5 (0.5%)	0.4342
Any comorbidity	Yes	955 (29.5%)	327 (30.3%)	320 (29.6%)	308 (28.5%)	0.6552
Fever	Yes	224 (6.9%)	66 (6.1%)	70 (6.5%)	88 (8.1%)	0.1398
Cough	Yes	648 (20.0%)	185 (17.1%)	213 (19.7%)	250 (23.1%)	**0.0022**
Sore throat	Yes	397 (12.3%)	112 (10.4%)	125 (11.6%)	160 (14.8%)	**0.0051**
Dyspnoea	Yes	141 (4.4%)	49 (4.5%)	46 (4.3%)	46 (4.3%)	0.9336
Myalgia	Yes	284 (8.8%)	74 (6.9%)	99 (9.2%)	111 (10.3%)	**0.0165**
Rhinorrhea	Yes	717 (22.1%)	205 (19.0%)	240 (22.2%)	272 (25.2%)	**0.0025**
Respiratory discomfort	Yes	188 (5.8%)	63 (5.8%)	58 (5.4%)	67 (6.2%)	0.7084
Nausea/vomit	Yes	120 (3.7%)	37 (3.4%)	39 (3.6%)	44 (4.1%)	0.7156
Headache	Yes	790 (24.4%)	244 (22.6%)	259 (24.0%)	287 (26.6%)	0.0936
Prostration	Yes	188 (5.8%)	60 (5.6%)	51 (4.7%)	77 (7.1%)	0.0523
Diarrhea	Yes	211 (6.5%)	59 (5.5%)	76 (7.0%)	76 (7.0%)	0.2336
Conjunctivitis	Yes	32 (1.0%)	13 (1.2%)	11 (1.0%)	8 (0.7%)	0.5478
Ageusia/anosmia	Yes	101 (3.1%)	30 (2.8%)	30 (2.8%)	41 (3.8%)	0.2914
Loss of voice	Yes	56 (1.7%)	18 (1.7%)	13 (1.2%)	25 (2.3%)	0.1381
Sought health assistance	Hospital	138 (4.3%)	41 (3.8%)	41 (3.8%)	56 (5.2%)	0.1492
	Basic Health Unit	129 (4.0%)	42 (3.9%)	41 (3.8%)	46 (4.3%)	
	Emergency Care Unit	127 (3.9%)	38 (3.5%)	35 (3.2%)	54 (5.0%)	
	None	2845 (87.8%)	958 (88.8%)	963 (89.2%)	924 (85.6%)	
Admitted to a health institution	Yes	38 (1.2%)	11 (1.0%)	12 (1.1%)	15 (1.4%)	0.7085
International travel	Yes	14 (0.4%)	10 (0.9%)	4 (0.4%)	0 (0.0%)	**0.0043**
Household contact with symptomatic person	Yes	640 (19.8%)	157 (14.6%)	193 (17.9%)	290 (26.9%)	**<0.0001**
Sorological test	Reactive	39 (1.2%)	3 (0.3%)	8 (0.7%)	28 (2.6%)	**<0.0001**
	Non-reactive	3200 (98.8%)	1076 (99.7%)	1072 (99.3%)	1052 (97.4%)	
PCR test	Detected	84 (2.6%)	2 (0.2%)	22 (2.0%)	60 (5.6%)	**<0.0001**
	Undetected	3112 (96.1%)	1035 (95.9%)	1057 (98.0%)	1020 (94.4%)	
	Indeterminate	42 (1.3%)	42 (3.9%)	0 (0.0%)	0 (0.0%)	
Prevalence	Sorological reactive and/or PCR detected	106 (3.3%)	5 (0.5%)	29 (2.7%)	72 (6.7%)	**<0.0001**

Bolded p-values indicate p < 0.05.
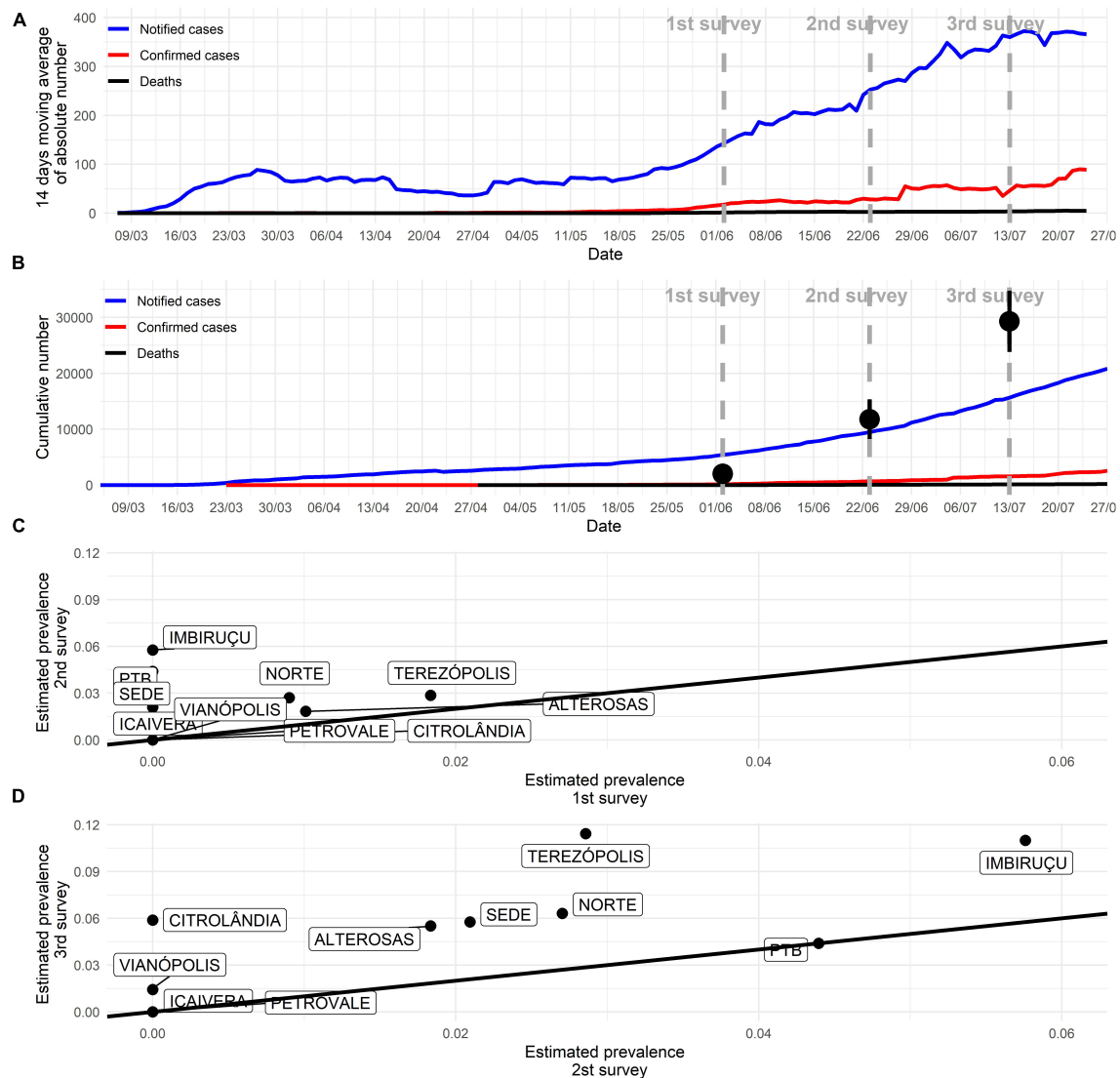
FIGURE 2 | COVID-19 pandemic progression in Betim. **(A)** The absolute number of new cases according to official city statistics. **(B)** The cumulative number of cases according to official city statistics. Black dots indicate estimated overall prevalence (immunological and molecular tests) in the current study with its 95% confidence interval. Distance from black dots and red curve represent underreporting. **(C,D)** Overall prevalence (immunological and molecular tests) comparison in each of the 10 administrative regions of the city across successive surveys. An increase was observed in most areas from the first to the second survey and, more substantially, from the second to the third survey.

Betim clusters. One cluster comprises four sequences, dating to May 27, 2020 (95% HPD: May 01, 2020–June 03, 2020) grouped with other sequences from other Brazilian regions and foreign. The model has also estimated eight introductions supported by single sequences. According to our phylogeny, the B.1.1.33 introductions came from different locations, such as the states of Rio de Janeiro, São Paulo, Minas Gerais, other Brazilian regions, and foreign sequences (PP > 0.81 for all events) (**Figure 5B**). The patterns reconstructed by both phylogeographic inferences are consistent, indicating the emergence of lineages B.1.1.28 and B.1.1.33 was followed by multiple importation events to diverse regions within the country, likely driven by human mobility. Additionally, evolutionary rate estimates also differed between

datasets (B.1.1.28: 8.6372×10^{-4}, 95% HPD: 7.8379×10^{-4}–9.4559×10^{-4}; B.1.1.33: 6.8743×10^{-4}, 95% HPD 6.1784×10^{-4}–7.5446×10^{-4}).

DISCUSSION

Betim is a medium-sized Brazilian city (439,340 inhabitants, 343 thousand square kilometers) crossed by national roads connecting major Brazilian cities and serving as a local hub for the Brazilian Public Health System. Understanding its pandemic dynamic may provide relevant information for municipalities with similar features. Here, we estimated the overall prevalence

FIGURE 3 | Spatial distribution of active infections across three surveys in Betim. **(A–C)** Dispersion of positive molecular tests across each survey. In the third survey **(C)**, most populated areas of the city already had a non-null probability of presenting residents with a positive molecular test.
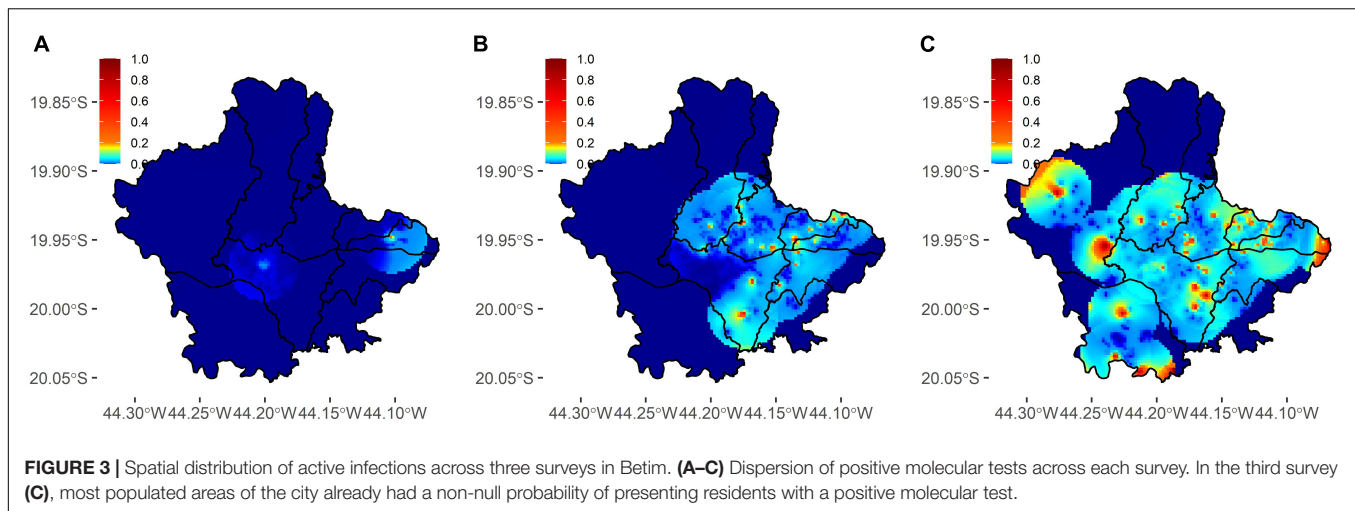
of active infections, seroprevalence and conducted genomic surveillance before the first pandemic wave in Betim. Brazilian molecular diagnostic capacity was insufficient in the first months of the pandemic (Grotto et al., 2020). Therefore, COVID-19 cases may have been included in the official statistics as severe acute respiratory infection cases with unknown etiology. Data until May 2020 indicated a positive association between higher per-capita income and molecular COVID-19 diagnosis, while the severe acute respiratory infection cases with unknown etiology were associated with lower per-capita income, suggesting a possible diagnosis bias related to economic status (de Souza et al., 2020). Inadequate diagnosis availability may lead to underreporting (Kupek, 2021). Our data estimated underreporting rates up to 20 times.

No studies have been conducted in Brazil evaluating active infection prevalence using adequate sampling. Our study design was inspired by previous research conducted in Santa Clara, United States, using pooled samples (Hogan et al., 2020). Pooled PCR tests were initially suggested to be used in asymptomatic people (Lohse et al., 2020) and later were recommended for surveillance studies in populations with low infection prevalence (Mutesa et al., 2021). Seroprevalence studies were conducted during the first wave in Brazil that peaked in July 2020. Two of the highest city seroprevalences reported during the period were Boa Vista (25.4% in June 2020) (Hallal et al., 2020) and São Luiz (40.4% between the end of July and August 2020) (da Silva et al., 2020), both in the northern area of the country. A nationwide survey carried out in May and June 2020 presented seroprevalence lower than 2% during both surveys in all sampled cities neighboring Betim (less than 200 km), corroborating our findings (Hallal et al., 2020). Furthermore, seroprevalences higher than 10% were solely found in towns in the North Region (Hallal et al., 2020). In December 2020, Manaus, the largest city in the North Region, experienced a resurgence of COVID-19 (Sabino et al., 2021) despite high seroprevalence (Buss et al., 2021), likely due to the gamma variant (Faria et al., 2021).

Previous seroprevalence studies have indicated ethnic and socioeconomic bias for SARS-CoV-2 infection in Brazil since

the pandemic's beginning (Amorim Filho et al., 2020; Horta et al., 2020). Results from Rio de Janeiro in April 2020 indicated that younger blood donors with lower education levels were more likely to test positive for SARS-CoV-2 antibodies (Amorim Filho et al., 2020). A nationwide study revealed that the poorest quintile was 2.16 times more likely to test positive with the lowest risks among white, educated, and wealthy individuals (Horta et al., 2020). Likewise, we found one of the highest prevalences in the poorest neighborhood, Terezópolis, that include the largest slum of the city where more than 23,000 people live.

Further modeling results showed higher infection rates among young adults, lower socioeconomic status, and people without healthcare access in the less developed North and Northeast areas until August 2020 (de Lima et al., 2021). Betim also presents most of its inhabitants with less than 59 years (90.7%), but no age effect was observed in the infectivity rates. Increased female infection odds were observed, although previous reports indicated a gender predisposition toward death in some Brazilian regions with higher male risk (Baqui et al., 2020). One possible explanation could be that 70% of the global health workforce are women (Lotta et al., 2021) and a gender bias of pandemic perception and attitude (Galasso et al., 2020).

COVID-19 diffusion presents strong socio-spatial determinants. Relocation diffusion from more- to less-developed regions and hierarchical diffusion from countries with higher population and density were relevant since early 2020 (Sigler et al., 2021). Data indicated a similar pattern in the São Paulo State with contiguous diffusion from the capital metropolitan area and hierarchical with long-distance spread through major highways that connects São Paulo city with cities of regional relevance (Fortaleza et al., 2021). Modeling results revealed that São Paulo city may have accounted for more than 85% of the initial case spread in the entire country (Nicolelis et al., 2021). Betim is directly connected to São Paulo city by a main national highway which may have contributed to COVID-19 diffusion.

Genomic surveillance is a powerful tool to elucidate viral dispersion patterns. The first sequencing work conducted in

TABLE 2 | Significant clinical and epidemiological data associations with a positive test (serological or molecular).

Variable	Level	Positive	Negative	p-value
Survey	First	5 (4.7%)	1074 (34.3%)	<0.0001
	Second	29 (27.4%)	1051 (33.5%)	
	Third	72 (67.9%)	1008 (32.2%)	
Administrative regions	Alterosas	18 (17.0%)	616 (19.7%)	0.0024
	Citrolândia	4 (3.8%)	215 (6.9%)	
	Icaivera	0 (0.0%)	62 (2.0%)	
	Imbiruçu	32 (30.2%)	533 (17.0%)	
	Norte	11 (10.4%)	322 (10.3%)	
	Petrovale	0 (0.0%)	41 (1.3%)	
	PTB	8 (7.5%)	282 (9.0%)	
	Sede	15 (14.2%)	568 (18.1%)	
	Terezópolis	17 (16.0%)	302 (9.6%)	
	Vianópolis	1 (0.9%)	192 (6.1%)	
Sex	Female	69 (65.1%)	1559 (49.8%)	0.0026
Fever	No	88 (83.0%)	2927 (93.4%)	<0.0001
	Yes	18 (17.0%)	206 (6.6%)	
Cough	No	73 (68.9%)	2518 (80.4%)	0.0053
	Yes	33 (31.1%)	615 (19.6%)	
Sore throat	No	77 (72.6%)	2765 (88.3%)	<0.0001
	Yes	29 (27.4%)	368 (11.7%)	
Dyspnoea	No	96 (90.6%)	3002 (95.8%)	0.0180
	Yes	10 (9.4%)	131 (4.2%)	
Myalgia	No	72 (67.9%)	2883 (92.0%)	<0.0001
	Yes	34 (32.1%)	250 (8.0%)	
Rhinorrhea	No	70 (66.0%)	2452 (78.3%)	0.0041
	Yes	36 (34.0%)	681 (21.7%)	
Respiratory discomfort	No	90 (84.9%)	2961 (94.5%)	<0.0001
	Yes	16 (15.1%)	172 (5.5%)	
Nausea/vomit	No	94 (88.7%)	3025 (96.6%)	<0.0001
	Yes	12 (11.3%)	108 (3.4%)	
Headache	No	50 (47.2%)	2399 (76.6%)	<0.0001
	Yes	56 (52.8%)	734 (23.4%)	
Prostration	No	83 (78.3%)	2968 (94.7%)	<0.0001
	Yes	23 (21.7%)	165 (5.3%)	
Ageusia/anosmia	No	87 (82.1%)	3051 (97.4%)	<0.0001
	Yes	19 (17.9%)	82 (2.6%)	
Obesity	No	96 (90.6%)	3038 (97.0%)	<0.0001
	Yes	10 (9.4%)	95 (3.0%)	
Sought health assistance	Hospital	8 (7.5%)	130 (4.1%)	0.0032
	None	81 (76.4%)	2764 (88.2%)	
	Basic Health Unit	8 (7.5%)	121 (3.9%)	
	Emergency Care Unit	9 (8.5%)	118 (3.8%)	
Household contact with symptomatic person	No	71 (67.0%)	2528 (80.7%)	0.0007
	Yes	35 (33.0%)	605 (19.3%)	

Non-significant associations are presented in **Supplementary Table 1**. Bolded p-values indicate p < 0.05.

Brazil evaluated the first six positive individuals and reported the same predominant lineages in Italy (Jesus et al., 2020). Later, a study with samples collected until late April 2020 from different country areas showed the dominance of clade B-derived lineages. At the national level, the respective frequency of these clades was seen in a 98.98%/1.02% ratio (Candido et al., 2020). In

Minas Gerais State, A lineages represented 2.5% of the infections, B.1 appeared in 92.5% of the samples, and B was responsible for 5% of the cases (Xavier et al., 2020). The exclusivity of lineages B.1.1.28 and B.1.1.33 circulating in Betim-MG from June to July 2020, given that multiple introductions from different country regions were demonstrated, is representative of the
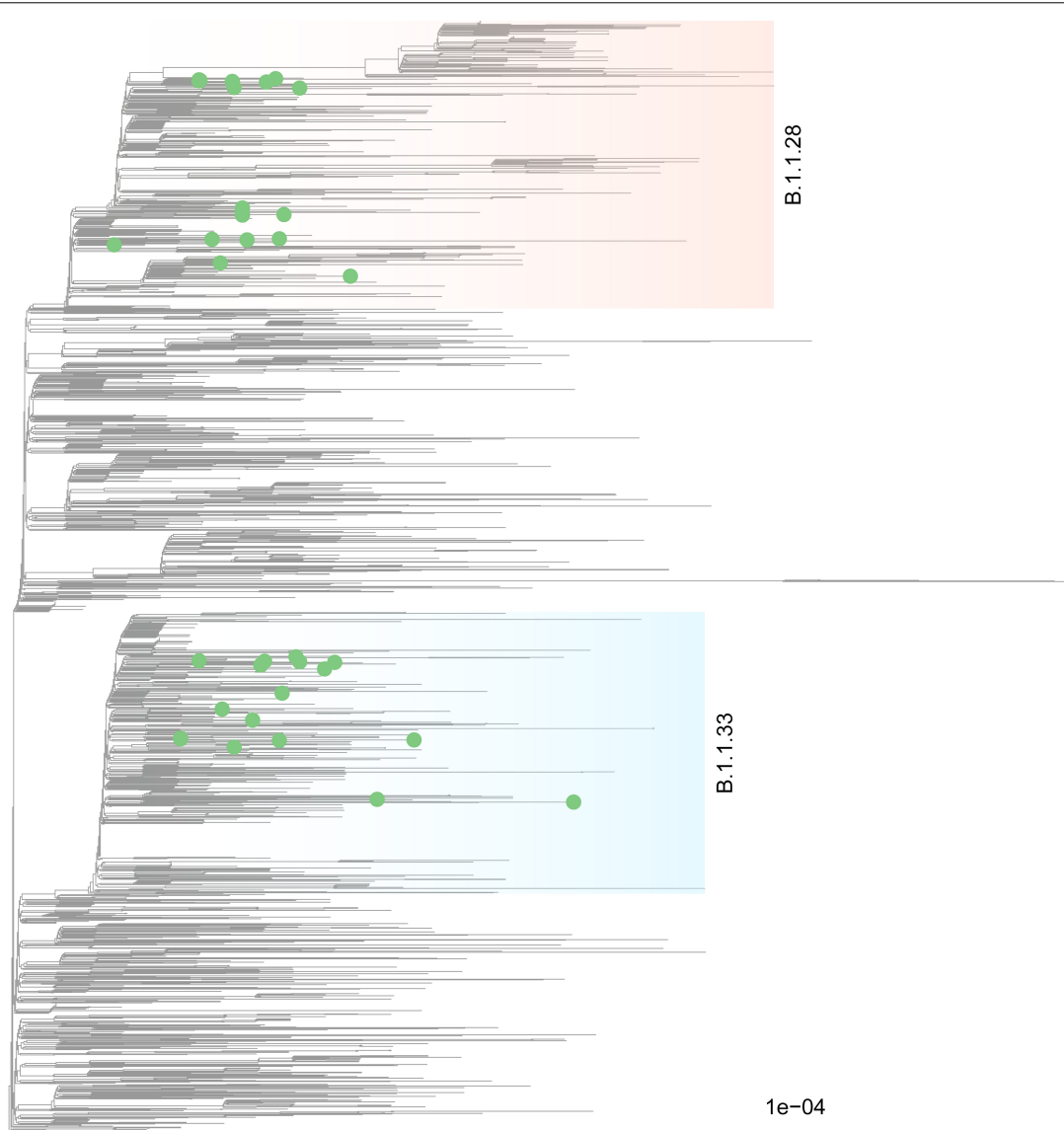
FIGURE 4 | Phylogenetic characterization of SARS-CoV-2 genomes characterized in Betim. A maximum-likelihood tree was inferred on IQ-Tree under the GTR+F+I+G4 model with a comprehensive reference dataset, encompassing all Brazilian sequences plus one international sequence per country per week, from late 2019 to January 12 2021 (*n* = 3,814). The phylogeny depicted exhibits a subtree of 2,023 tips that harbors all relevant diversity considered for this study, mainly lineages B.1.1.28 (light salmon) and B.1.1.33 (light blue) where the novel genome sequences sparsely clustered. Tip shapes mark sequences characterized in this study. The scale bar indicates average nucleotide substitutions per site.

extent of these lineages' dominance in the Brazilian scenario at the moment. Independent introductions also emphasize the importance of inter-state mobility barriers as a measure to control the epidemic.

Our study presents some limitations. First, the household survey is less likely to sample severe cases, thus underestimating symptomatic COVID-19. Second, all clinical data were self-reported, which may lead to reporting bias (Baker et al., 2004). Third, we could not sequence all PCR positive samples due to the low viral load and sequencing technology employed, but we do not expect biased frequency estimation since no differences in mean viral load was reported across B.1.1.28

and B.1.1.33 lineages. A different scenario, later in 2020, was observed after variants of concern detection that led to higher mean viral loads (e.g., P.1 or gamma variant) (Faria et al., 2021). Fourth, our study provides a limited picture of the local epidemic because of the short period across surveys, although it was the moment when the city had less reliable data regarding pandemic progression. Nevertheless, our study shows the potential to integrate different epidemiological inquiries (prevalence, seroprevalence, and genomic surveillance) to describe pandemic dispersion adequately. Moreover, our findings present original and relevant evidence that has helped local government authorities to guide pandemic management.
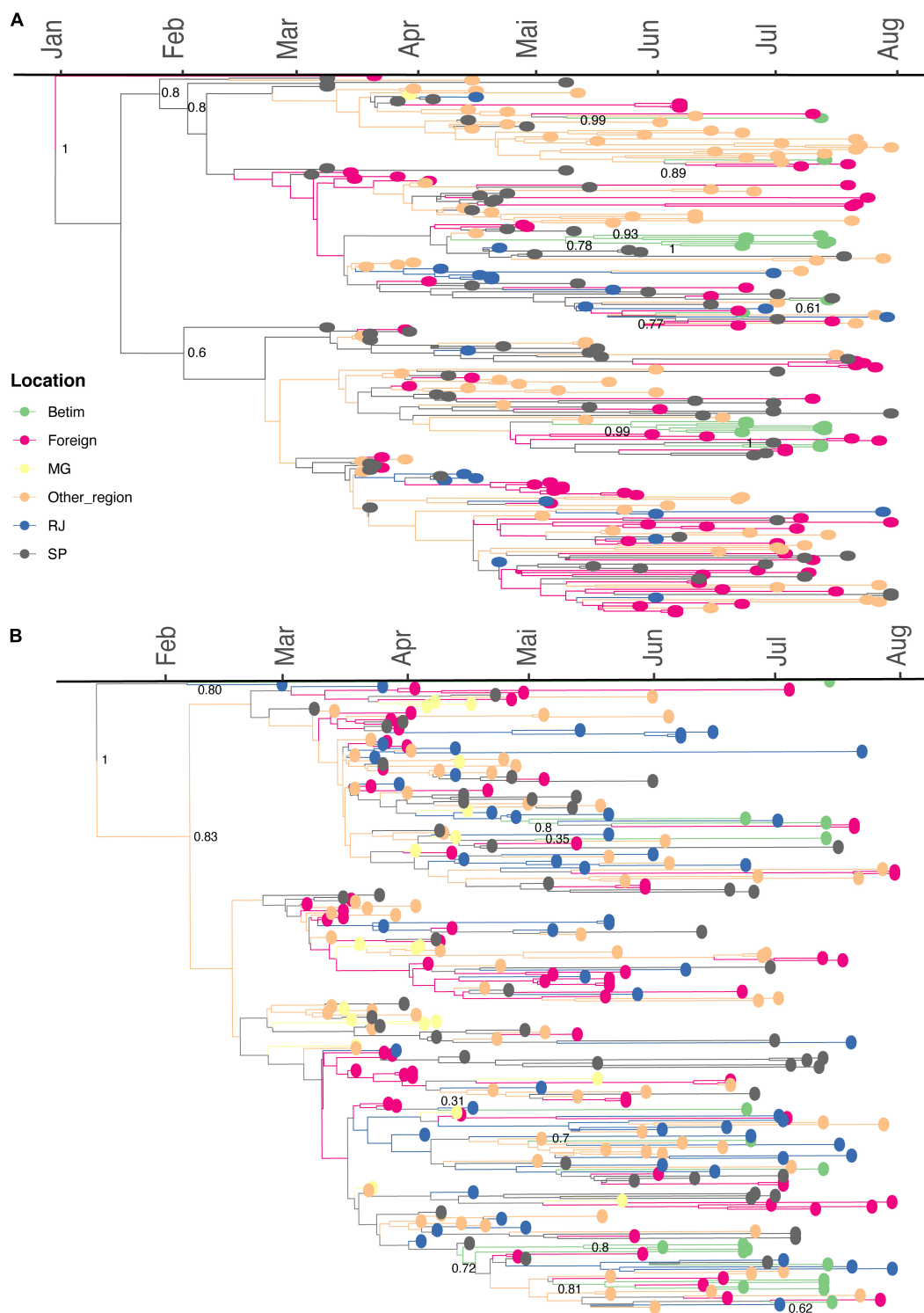
FIGURE 5 | Spread of B.1.1.28 and B.1.1.33 lineages in Betim city. **(A)** Time-resolved maximum clade credibility phylogeny from a dataset comprehending 240 publicly available B.1.1.28 sequences and the 18 genomes generated in this study. **(B)** Time-resolved maximum clade credibility phylogeny from a dataset including 267 publicly available B.1.1.33 sequences and the 17 genomes generated in this study. For both analyses, the HKY+I+G4 nucleotide substitution model was used. The diamond indicates sequences from Betim city obtained in this study. The trees inferred are available on https://github.com/LBI-lab/SARS-CoV-2_phylogenies.git.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Betim Ethics Research Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AVFGS, OT, VA, TA, AC, AR, HO, NO, EGS, PM, CM, TN, FP, EFS, RA, and RPS: study design. AVFGS, DM, FRM, OT, PF, RM, HA, VA, TA, AC, JMS, AR, LA, JA, HO, NO, CZ, JHS, EGS, RS, LF, AL, APG, PM, FMM, LM, CM, TN, FP, DCQ, DNQ, LR-M, FS, EFS, CV, AV, RA, and RPS: experimental procedures. AVFGS, DM, FRM, OT, PF, and RPS: manuscript draft. All authors revised and approved the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.799713/full#supplementary-material

REFERENCES

Amorim Filho, L., Szwarcwald, C. L., Mateos Sog, S., Leon, A. C. M. P., Medronho, R. A., Veloso, V. G., et al. (2020). Seroprevalence of anti-SARS-CoV-2 among blood donors in Rio de Janeiro, Brazil. *Rev. Saude Publica* 54:69. doi: 10.11606/s1518-8787.2020054002643

Araujo, D. B., Machado, R. R. G., Amgarten, D. E., de, F., Malta, M., de Araujo, G. G., et al. (2020). SARS-CoV-2 isolation from the first reported patients in Brazil and establishment of a coordinated task network. *Mem. Inst. Oswaldo Cruz* 115:e200342. doi: 10.1590/0074-02760200342

Araújo, J. L., Menezes, D., Saraiva-Duarte, J. M., Ferreira, L., Aguiar, R., and Souza, R. (2021). Systematic review of host genetic association with Covid-19 prognosis and susceptibility: What have we learned in 2020? *Rev. Med. Virol.* e2283. doi: 10.1002/rmv.2283

Baker, M., Stabile, M., and Deri, C. (2004). What do self-reported, objective, measures of health measure? *J. Hum. Resour.* 39:1067. doi: 10.2307/3559039

Baqui, P., Bica, I., Marra, V., Ercole, A., and van der Schaar, M. (2020). Ethnic and regional variations in hospital mortality from COVID-19 in Brazil: a cross-sectional observational study. *Lancet Glob. Health* 8, e1018–e1026. doi: 10.1016/S2214-109X(20)30285-0

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Buss, L. F., Prete, C. A., Abrahim, C. M. M., Mendrone, A., Salomon, T., de Almeida-Neto, C., et al. (2021). Three-quarters attack rate of SARS-CoV-2 in the Brazilian amazon during a largely unmitigated epidemic. *Science* 371, 288–292. doi: 10.1126/science.abe9728

Byambasuren, O., Dobler, C. C., Bell, K., Rojas, D. P., Clark, J., McLaws, M.-L., et al. (2021). Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: Systematic review. *PLoS One* 16:e0248946. doi: 10.1371/journal.pone.0248946

Candido, D. S. I, Claro, M., de Jesus, J. G., Souza, W. M., Moreira, F. R. R. R., Dellicour, S., et al. (2020). Evolution and epidemic spread of SARS-CoV-2 in Brazil. *Science* 369, 1255–1260. doi: 10.1126/SCIENCE.ABD2161

da Silva, A. A. M., Lima-Neto, L. G., Azevedo, C. M. P. E. S., Costa, L. M. M. D., Bragança, M. L. B. M., Barros Filho, A. K. D., et al. (2020). Population-based seroprevalence of SARS-CoV-2 and the herd immunity threshold in Maranhão. *Rev. Saude Publica* 54:131. doi: 10.11606/s1518-8787.2020054003278

de Lima, E. E. C., Gayawan, E., Baptista, E. A., and Queiroz, B. L. (2021). Spatial pattern of COVID-19 deaths and infections in small areas of Brazil. *PLoS One* 16:e0246808. doi: 10.1371/journal.pone.0246808

de Souza, W. M., Buss, L. F., da, D., Candido, S., Carrera, J.-P., Li, S., et al. (2020). Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. *Nat. Hum. Behav.* 4, 856–865. doi: 10.1038/s41562-020-0928-4

Faria, N. R., Mellan, T. A., Whittaker, C., Claro, I. M., Candido, D. D. S., Mishra, S., et al. (2021). Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* 372, 815–821. doi: 10.1126/science.abh2644

Fortaleza, C. M. C. B., Guimarães, R. B., Catão, R. C., Ferreira, C. P., Berg de Almeida, G., Nogueira Vilches, T., et al. (2021). The use of health geography modeling to understand early dispersion of COVID-19 in São Paulo, Brazil. *PLoS One* 16:e0245051. doi: 10.1371/journal.pone.0245051

Galasso, V., Pons, V., Profeta, P., Becher, M., Brouard, S., and Foucault, M. (2020). Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries. *Proc. Natl. Acad. Sci. U.S.A.* 117, 27285–27291. doi: 10.1073/pnas.2012520117

Gill, M. S., Lemey, P., Faria, N. R., Rambaut, A., Shapiro, B., and Suchard, M. A. (2013). Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* 30, 713–724.

Grotto, R. M. T., Santos Lima, R., de Almeida, G. B., Ferreira, C. P., Guimarães, R. B., Pronunciate, M., et al. (2020). Increasing molecular diagnostic capacity and COVID-19 incidence in Brazil. *Epidemiol. Infect.* 148:e178. doi: 10.1017/S0950268820001818

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Hallal, P. C., Hartwig, F. P., Horta, B. L., Silveira, M. F., Struchiner, C. J., Vidaletti, L. P., et al. (2020). SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. *Lancet Glob. Health* 8, e1390–e1398. doi: 10.1016/S2214-109X(20)30387-9

Hogan, C. A., Sahoo, M. K., and Pinsky, B. A. (2020). Sample Pooling as a strategy to detect community transmission of SARS-CoV-2. *JAMA* 323:1967. doi: 10.1001/jama.2020.5445

Horta, B., Silveira, M., Barros, A., Barros, F., Hartwig, F., Dias, M., et al. (2020). Prevalence of antibodies against SARS-CoV-2 according to socioeconomic and ethnic status in a nationwide Brazilian survey. *Rev. Panam. Salud Pública* 44, 1–7. doi: 10.26633/RPSP.2020.135

Jesus, J. G., Sacchi, C., Candido, D. D. S., Claro, I. M., Sales, F. C. S., Manuli, E. R., et al. (2020). Importation and early local transmission of COVID-19 in Brazil, 2020. *Rev. Inst. Med. Trop. Sao Paulo* 62, e30. doi: 10.1590/s1678-9946202062030

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kupek, E. (2021). How many more? Under-reporting of the COVID-19 deaths in Brazil in 2020. *Trop. Med. Int. Health* 26, 1019–1028. doi: 10.1111/tmi.13628

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25

Lemey, P., Rambaut, A., Drummond, A. J., and Suchard, M. A. (2009). Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5:e1000520. doi: 10.1371/journal.pcbi.1000520

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Lohse, S., Pfuhl, T., Berkó-Göttel, B., Rissland, J., Geißler, T., Gärtner, B., et al. (2020). Pooling of samples for testing for SARS-CoV-2 in asymptomatic people. *Lancet Infect. Dis.* 20, 1231–1232. doi: 10.1016/S1473-3099(20)30362-5

Lotta, G., Fernandez, M., Pimenta, D., and Wenham, C. (2021). Gender, race, and health workers in the COVID-19 pandemic. *Lancet* 397:1264. doi: 10.1016/S0140-6736(21)00530-4

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Moreira, F. R. R., Bonfim, D. M., Zauli, D. A. G., Silva, J. P., Lima, A. B., Malta, F. S. V., et al. (2021). Epidemic spread of SARS-CoV-2 lineage B.1.1.7 in Brazil. *Viruses* 13: 984. doi: 10.3390/v13060984

Mutesa, L., Ndishimye, P., Butera, Y., Souopgui, J., Uwineza, A., Rutayisire, R., et al. (2021). A pooled testing strategy for identifying SARS-CoV-2 at low prevalence. *Nature* 589, 276–280. doi: 10.1038/s41586-020-2885-5

Nicolelis, M. A. L., Raimundo, R. L. G., Peixoto, P. S., and Andreazzi, C. S. (2021). The impact of super-spreader cities, highways, and intensive care availability in the early stages of the COVID-19 epidemic in Brazil. *Sci. Rep.* 11:13001. doi: 10.1038/s41598-021-92263-3

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in bayesian phylogenetics using tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Rambaut, A., Holmes, E. C., O'Toole, Á, Hill, V., McCrone, J. T., Ruis, C., et al. (2020). A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* 5, 1403–1407. doi: 10.1038/s41564-020-0770-5

Rambaut, A., Lam, T. T., Max Carvalho, L., and Pybus, O. G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* 2:vew007. doi: 10.1093/ve/vew007

Robishaw, J. D., Alter, S. M., Solano, J. J., Shih, R. D., DeMets, D. L., Maki, D. G., et al. (2021). Genomic surveillance to combat COVID-19: challenges

and opportunities. *Lancet Microb.* 2, e481–e484. doi: 10.1016/S2666-5247(21)00121-X

Rossi, ÁD., de Araújo, J. L. F., de Almeida, T. B., Ribeiro-Alves, M., de Almeida Velozo, C., de Almeida, J. M., et al. (2021). Association between ACE2 and TMPRSS2 nasopharyngeal expression and COVID-19 respiratory distress. *Sci. Rep.* 11:9658. doi: 10.1038/s41598-021-88944-8

Sabino, E. C., Buss, L. F., Carvalho, M. P. S., Prete, C. A., Crispim, M. A. E., Fraiji, N. A., et al. (2021). Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* 397, 452–455. doi: 10.1016/S0140-6736(21)00183-5

Shu, Y., and McCauley, J. (2017). GISAID: global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* 22:30494. doi: 10.2807/1560-7917.ES.2017.22.13.30494

Sigler, T., Mahmuda, S., Kimpton, A., Loginova, J., Wohland, P., Charles-Edwards, E., et al. (2021). The socio-spatial determinants of COVID-19 diffusion: the impact of globalisation, settlement characteristics and population. *Global. Health* 17:56. doi: 10.1186/s12992-021-00707-2

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 4:vey016. doi: 10.1093/ve/vey016

Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Am. Math. Soc. Lect. Math. Life Sci.* 17, 57–86.

Wu, S. L., Mertens, A. N., Crider, Y. S., Nguyen, A., Pokpongkiat, N. N., Djajadi, S., et al. (2020). Substantial underestimation of SARS-CoV-2 infection in the United States. *Nat. Commun.* 11:4507. doi: 10.1038/s41467-020-18272-4

Xavier, J., Giovanetti, M., Adelino, T., Fonseca, V., Barbosa da Costa, A. V., Ribeiro, A. A., et al. (2020). The ongoing COVID-19 epidemic in Minas Gerais, Brazil: insights from epidemiological data and SARS-CoV-2 whole genome sequencing. *Emerg. Microbes Infect.* 9, 1824–1834. doi: 10.1080/22221751.2020.1803146

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* 39, 306–314. doi: 10.1007/BF00160154

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273. doi: 10.1038/s41586-020-2012-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Check for
updates

Improving Taxonomic Delimitation of Fungal Species in the Age of Genomics and Phenomics

Ashley Stengel[1,2,3]*, Kimberly M. Stanke[1,4], Amanda C. Quattrone[1,5,6] and Joshua R. Herr[3,5,6]*

[1] Complex Biosystems Interdisciplinary Life Sciences, University of Nebraska-Lincoln, Lincoln, NE, United States,
[2] Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, United States, [3] Department of Plant Pathology, University of Nebraska-Lincoln, Lincoln, NE, United States, [4] Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, United States, [5] School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE, United States, [6] Center for Plant Science Innovation, University of Nebraska-Lincoln, Lincoln, NE, United States

Species concepts have long provided a source of debate among biologists. These lively debates have been important for reaching consensus on how to communicate across scientific disciplines and for advancing innovative strategies to study evolution, population biology, ecology, natural history, and disease epidemiology. Species concepts are also important for evaluating variability and diversity among communities, understanding biogeographical distributions, and identifying causal agents of disease across animal and plant hosts. While there have been many attempts to address the concept of species in the fungi, there are several concepts that have made taxonomic delimitation especially challenging. In this review we discuss these major challenges and describe methodological approaches that show promise for resolving ambiguity in fungal taxonomy by improving discrimination of genetic and functional traits. We highlight the relevance of eco-evolutionary theory used in conjunction with integrative taxonomy approaches to improve the understanding of interactions between environment, ecology, and evolution that give rise to distinct species boundaries. Beyond recent advances in genomic and phenomic methods, bioinformatics tools and modeling approaches enable researchers to test hypothesis and expand our knowledge of fungal biodiversity. Looking to the future, the pairing of integrative taxonomy approaches with multi-locus genomic sequencing and phenomic techniques, such as transcriptomics and proteomics, holds great potential to resolve many unknowns in fungal taxonomic classification.

Keywords: fungal diversity, eco-evolutionary theory, integrative taxonomy, species delimitation, omics, bioinformatics

INTRODUCTION: SPECIES CONCEPTS AND THE ROLE OF ECO-EVOLUTIONARY DYNAMICS

The overarching goal of species concepts is to provide a framework for the effective and reliable classification of organisms into logical categories. Decades of discussion have given rise to several predominant species concepts (**Figure 1**; reviewed in Zachos, 2016). Most notably, these include

GRAPHICAL ABSTRACT | Enhanced resolution of spices boundaries.
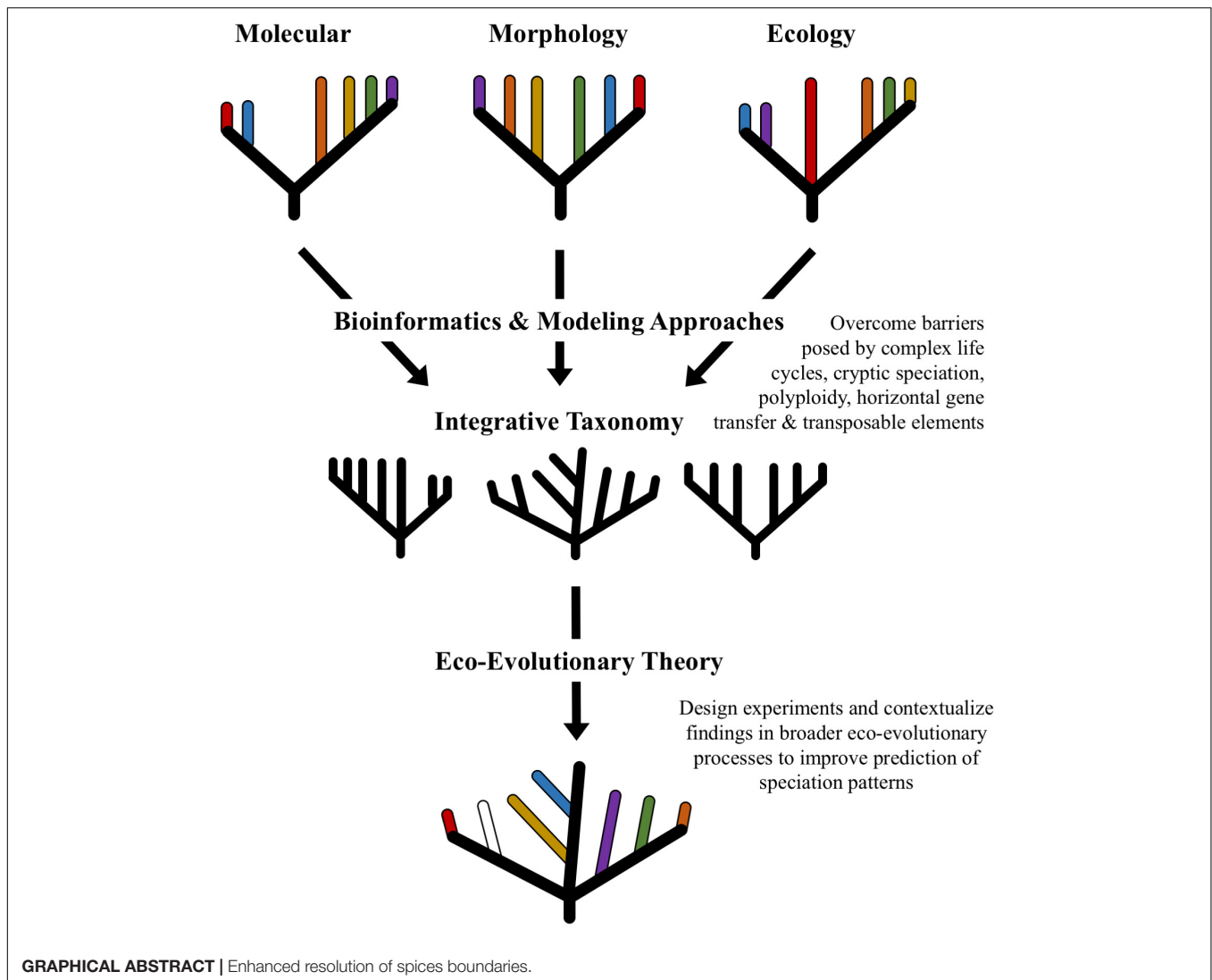
concepts based on biological (Mayr, 1996), phylogenetic (Nixon and Wheeler, 1990; Taylor et al., 2000; de Queiroz, 2005), morphological (Michener, 1970; Bérubé and Dessureault, 1989; Nakasone, 1996), and ecological criteria (Van Valen, 1976; Schmidt et al., 2014). These concepts have been recently reviewed by Xu (2020), addressed for plant pathogenic fungi by Cai et al. (2011), and applied to Arbuscular mycorrhizal fungi (AMF) by Bruns et al. (2018). Overall, there is growing recognition that a single set of criteria do not sufficiently describe the diversity seen among fungal lineages.

In the last 15 years, efforts have been made to shift the discussion of species delimitation away from a preset list of criteria and toward an understanding of the community-level processes driving patterns of diversification (de Queiroz, 2007). In fact, the General Lineage Concept (de Queiroz, 2005) arose to highlight the role of history (i.e., evolution) and circumstance (i.e., environment) in driving differences between the contingent properties of species. Through this framework, evolutionary groups are distinguishable due to their shared history and

processes of adaptation and genetic drift. These eco-evolutionary processes give rise to contingent properties that scientists may use to trace and separate species (de Queiroz, 2007). In recent years, discussion of eco-evolutionary dynamics (Brunner et al., 2019) and its application to microbial communities (Saikkonen et al., 2020) have gained attention. Eco-evolutionary dynamics recognize the role of ecological properties (e.g., spatial heterogeneity, habitat conditions, dispersal, and assembly patterns) and evolutionary trajectories (e.g., life history traits, co-evolution mosaics, intra- and inter-population selection) in reciprocally influencing one another (Brunner et al., 2019). As such, applying an eco-evolutionary framework to microbial interactions has offered new insight into the processes shaping communities and species interactions (McDonald et al., 2020). For example, the framework has given rise to the coupling of molecular and microbiological tools with theoretical approaches toward genetics, phenotypic evolution, species co-evolution, and ecology (Saikkonen et al., 2020). This has improved the identification of keystone taxa, the elucidation of processes
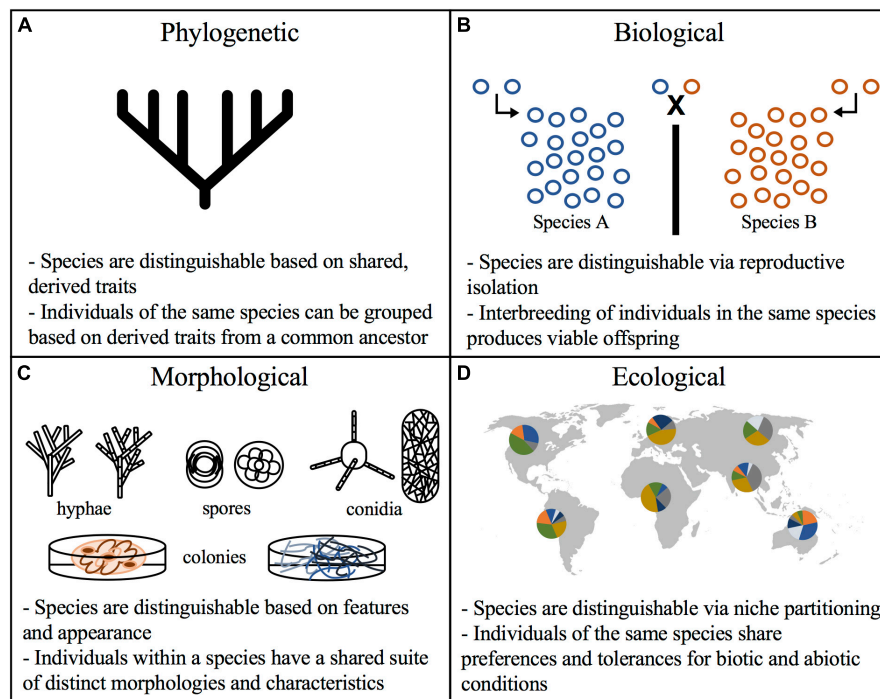
FIGURE 1 | Species concepts criteria for fungal delimitation. Predominant species concepts and the features used to distinguish species. **(A)** Phylogenetic species concept relies on parsimony in reconstructions of taxa relatedness and groups organisms by shared traits; **(B)** biological species concept draws boundaries based on reproductive isolation wherein organisms incapable of producing viable offspring are considered separate species; **(C)** morphological species concept distinguishes species using macro and micro morphology, including colony appearance, pigmentation, conidia, mycelia, and spores; **(D)** ecological species concept evaluates biogeographical ranges and potential niches to delimit taxonomic groups.

shaping host-microbe interactions, and the assignment of putative functions to communities (Cordovez et al., 2019). With regards to fungal species, eco-evolutionary dynamics offer a conceptual framework for researching core ecological and evolutionary processes yielding insight into patterns of biological diversity that shape metapopulations, and ultimately improve the delimitation of species boundaries.

Similar to eco-evolutionary theory, integrative taxonomy is a conceptual framework that recognizes the roles of ecological and evolutionary processes in shaping species boundaries. Therefore, applying recent advances in eco-evolutionary frameworks for microbiome diversity with integrative taxonomy approaches offers great potential for improving research design and interpretation. In practice, integrative taxonomy aims to combine discipline specific observations with broad technical evidence to develop more holistic conclusions regarding evolution and species delimitation (Dayrat, 2005). This approach builds on coalescent theory (Avise and Wollenberg, 1997), combining criteria from various species concepts (e.g., biological, phylogenetic, morphological, etc.) in order to delimit taxa in a way that bridges disciplines and methods of study. Integrative taxonomy has also been referred to as a "multisource approach" (Schlick-Steiner et al., 2010) and has been applied across several eukaryotic lineages, with recent examples including plants (Prata et al., 2018), insects (Lamarre et al., 2016), and fish (Pugedo et al., 2016). Using eco-evolutionary theory to contextualize integrative taxonomy approaches will enable hypothesis testing which can

improve discrimination of genetic, phenotypic, and functional traits that span the fungal tree of life. To better understand the utility of integrative taxonomy and eco-evolutionary framing for research, we highlight how key features of fungal biology have posed challenges for taxonomic delimitation and study. Keeping these challenges in mind, we then address how advances in sequencing, "omics," bioinformatics, and modeling can be applied within the eco-evolutionary and integrative taxonomy frameworks to resolve fungal species boundaries.

CHALLENGES TO DEFINING FUNGAL SPECIES

Complex Life Cycles and Cryptic Species

One of the major obstacles facing the delimitation of fungal species is variation in morphology. The complexity and diversity in life cycles exhibited by fungi has led researchers to misidentify species in two major ways–through dual-naming conventions or cryptic species (Taylor, 2011). The type species described for the sexual stage (teleomorph) and the asexual stage (anamorph), may be very morphologically different. Molecular analyses have shown seemingly disparate taxa to represent the same species. One such example is that of *Curvularia*– an important species complex of phytopathogens that are destructive largely in grasses and cereal plants. This group of filamentous fungi have a *Cochliobolus* teleomorphic stage

that has resulted in several changes in taxonomy as well as confusion with a related fungal genus *Bipolaris* (Manamgoda et al., 2012). Similarly, medically relevant fungal organisms may exhibit different biochemical and morphological traits in their anamorph-teleomorph phases, which has led to difficulty in identifying clinically relevant strains (Latouche et al., 1997). Detecting species across life stages in complex cycles is crucial for capturing biodiversity among fungal lineages and improving our understanding of eco-evolutionary pressures that drive natural variation among populations.

A second obstacle is that fungal species may be easily misidentified when two specimens exhibit identical morphologies and/or very closely related phenotypic traits yet show clearly distinct genetic profiles. Often termed "cryptic species," this moniker highlights that species boundaries in fungi are not always clearly defined (Bickford et al., 2007). *Fusarium subglutinans*, which has more than 20 individuals described as part of the species complex, is a group of well documented plant pathogens demonstrating similar morphology, but distinct biological, ecological, and phylogenetic relationships (Dugan and Everhart, 2016). Similarly, recent evidence shows that *Fusarium* species causing Fusarium Head Blight disease in wheat and other grains are capable of shifting geographic ranges or altering mycotoxin production, which poses a significant challenge to crop yields worldwide (Valverde-Bogantes et al., 2020). Another important example includes the *Aspergillus niger* segregates, which display differing metabolite production, diverse preferential environmental niche adaptation, and broad host ranges despite highly similar and at times indistinguishable morphological characteristics (Howard et al., 2011). Improving the resolution of species boundaries is important for elucidating environmental preferences and evaluating the potential for toxin production which are essential for developing effective disease management practices.

Polyploidy and Transposable Elements

Polyploidy arising from partial and whole genome duplication events (e.g., auto-polyploidy) has been widely studied in fungi. While specific mechanisms and functional relationships are not fully characterized, there is evidence suggesting that fungi may gain some benefits from harboring larger genomes. For example, changes in ploidy can increase organismal fitness during periods of stress and may facilitate acclimatization to changing environments (Todd et al., 2017). *Rhizophagus irregularis* demonstrates higher densities of poly allelic single-nucleotide polymorphisms, contributing to high within-isolate variability (Wyss et al., 2016). This could be attributed to heterokaryotic states among the population that give rise to differing gene copy numbers and divergence in copies among isolates (Tisserant et al., 2013). This variation further complicates that ability to draw distinctions between isolates and clearly demarcate species boundaries.

Polyploidy can also arise through the merging of genomic content between different species, called allopolyploidy. An evaluation of more than 600 genome assemblies from fungal proteomes and genomes revealed that transposable elements cluster together and contribute to larger fungal genomes

(Muszewska et al., 2019). In plant pathogenic fungi, horizontal gene transfer and horizontal chromosome transfer has been observed as a means to expand host ranges (Mehrabi et al., 2011). Furthermore, aneuploidy events can cause the loss of transposable elements which in turn influence fungal lifestyle. The strong selection pressures for and widespread presence of ploidy among fungal lineages pose significant challenges to identifying and characterizing species. Furthermore, the mechanisms driving changes in ploidy are not completely understood. This gap in knowledge draws attention to the importance of developing robust methods that can delimit species in a reliable way.

EMERGING TECHNIQUES AND STRATEGIES

Multi-Locus Sequencing

In addition to their contribution to genomics, the advent of high-throughput sequencing approaches has brought about many advances in the way we study fungal communities, particularly those found in natural and agricultural environments. Recent reviews have discussed the role of DNA sequencing to identify bacteria and fungi in natural communities (Inderbitzin et al., 2020), as well as sampling and laboratory protocols, and analysis techniques to garner more reliable insight into mycobiome diversity (Nilsson et al., 2019). Tekpinar and Kalmer (2019) review various molecular markers and assess their utility in fungal identification and phylogeny construction, concluding that the ITS operon can capture a high degree of interspecies variability. However, sequence variation within the same genome can pose a significant challenge for correctly assigning phylogeny to ITS rRNA sequences, as is the case for many arbuscular mycorrhizal fungi (AMF), including *Glomus intraradices* (Börstler et al., 2008). For AMF phylogenies the ITS operon has underperformed in comparison to 18S rRNA (Hart et al., 2015) due to high sequence length variability (Reich and Labes, 2017) and low sequencing depth (Berruti et al., 2017). As such, the use of the ITS operon in combination with protein-coding genes may be more effective for species-level identification of fungi (Lücking et al., 2020), including AMF (Wyss et al., 2016). The multi-gene or multi-locus sequencing approach has gained attention in distinguishing fungal taxa at a finer scale, thereby allowing hypothesis testing in an evolutionary context that also acknowledges phenotypic characters (Matute and Sepúlveda, 2019). This concordance approach, also called phylogenomics, allows different phylogenetic trees to be constructed based on the combinations of genes used to infer phylogenetic relationships (Herr et al., 2015; Hibbett et al., 2016).

A major advantage to using multi-locus phylogenetic analysis is the ability to resolve cryptic species. This is largely due to the higher degree of homoplasy among closely related cryptic species that enables haplotype-level discrimination not feasible between more distantly related taxa (Sato et al., 2020). Among closely related groups, the overall variation across the ITS regions may either not be enough or be too much for accurate discrimination, therefore, incorporating additional

genetic information may improve partitioning between lineages. Examples of improved delimitation using multi-locus sequencing extend across fungal phyla and fungal lifestyles (Somma et al., 2019; Li et al., 2020; Nie et al., 2020). In fact, parallel sequencing of single copy nuclear genes and mitochondrial (mtLSU, mtSSU) and nuclear (ITS1, ITS2) ribosomal regions was able to distinguish haplotypes of the wood-decomposing Basidiomycete *Hypholoma fasciculare* species complex (Sato et al., 2020). Similarly, concordance analysis based on whole genome sequencing indicated the existence of three distinct host-specialized species of ant-pathogenic fungi from the cryptic species complex *Ophiocordyceps unilateralis* (Kobmoo et al., 2019). Expanding the pool of genes used for phylogenetic reconstruction is an important strategy for improving our understanding of evolutionary relationships between fungal species, including resolving unknown signals within cryptic species complexes (**Figure 2A**).

Omics and Modeling

The emergence of species along evolutionary trajectories is shaped by a suite of stochastic processes (e.g., random mutations, chance events, population size, modes of reproduction, selection pressures from abiotic factors, etc.) that may give rise to distinct genetic or phenotypic signatures (**Figure 2E**). As such, meta-"omics" techniques (i.e., genomics, transcriptomics, proteomics, metabolomics) provide another avenue for discovering the breadth of fungal biodiversity and for drawing clearer species boundaries (Kuske et al., 2015). These strategies are especially useful for overcoming the challenges posed by polyploidy and cryptic species, as elucidation of the genetic and functional traits of organisms provides greater insight into phylogenomic and eco-evolutionary patterns (Martiny et al., 2015).

Metagenomic and whole-genome sequencing approaches enable researchers to evaluate phylogenetic relationships using comparative and consensus-based strategies (recently review by Xu, 2020). This has clarified mis-assignment of fungal species with dual naming conventions based on morphological differences between anamorph/teleomorph stages (Taylor, 2011), and improved discrimination of cryptic species by providing robust molecular data that accounts for stochastic processes like gene flow (Kobmoo et al., 2019). Additionally, whole-genome datasets may be evaluated with population genomic approaches to build coalescent phylogenies from a suite of nuclear loci, which improves the accuracy of cryptic species delimitation (Sato et al., 2020). Similarly, metatranscriptomic data captures a large subset of genes from the active members of microbial communities that may be used to build coalescent phylogenies capable of detecting strain-level differences. Marcelino et al. (2019) showed how metatranscriptomics was able to correctly classify fungal taxa living in mixed communities and further identify strain-level variation within the cryptic species complex *Cryptococcus*. The evaluation of the secretome–or the secreted proteins differentiated from proteomic data–is an emerging technique that enables the detection of pathogenic fungi associated with human and animal hosts (Varona et al., 2020). The expansion of multi-"omics" techniques–using a combination of genomic and phenomic data–has further enabled the application of

phylogenomic coalescent approaches to improve delineation of fungal species living in diverse environments (Reich and Labes, 2017; Libkind et al., 2020; Wibberg et al., 2020). Despite barriers posed by cryptic speciation, horizontal gene transfer between species, and genome plasticity, contemporary genomic and phenomic methods are able to resolve ambiguity in species boundaries. In particular, coupling genome-based techniques with exploration of the phenome, including expression profiles (e.g., transcripts, proteins, metabolites), morphological characters, and biogeographical ranges, offers new insight into the ecological and evolutionary processes that drive speciation in fungi (Martiny et al., 2015; Reich and Labes, 2017). These techniques highlight how integrative taxonomy approaches may provide deeper insight into drivers of evolution that shape species boundaries (**Figures 2A,B**). These findings can be better conceptualized using the eco-evolutionary framework to address how evolutionary trajectories are modified by and in turn influence fungal ecology.

Beyond emerging laboratory and measurement techniques, advancements in modeling and applied mathematics also offer additional insight into identification of speciation in fungi (**Figure 2C**). For example, Bayesian inference and maximum likelihood phylogenetic reconstruction are useful for building coalescent trees to resolve species boundaries, particularly when used in conjunction with multi-locus molecular data (Widhelm et al., 2016; Wyss et al., 2016). Further, metaphenomic modeling (Chowdhury et al., 2019) evaluates microbial phenomics using a reaction network graphical modeling approach to predict the impact of environmental variables on microbial community transcriptional and metabolic expression. Identifying how patterns of microbial signaling and expression change along environmental and ecological conditions provides important insight into potential drivers of microbial diversity and evolution. An additional emerging modeling technique uses human-like decision making, rather than standard 0 or 1 (i.e., "true" or "false") binary systems. This paradigm, termed fuzzy logic, has been used by Yusof et al. (2013) in an automated wood species recognition system to pre-classify tropical wood species in timber industries, improving accuracy by over 4%. Overall, this may better capture evolutionary trajectories which do not follow binary, deterministic processes. These and other modeling techniques may offer important insight into delineation of fungal species, particularly when used in combination with eco-evolutionary theory to contextualize modeling outputs.

Bioinformatics and Computational Biology Approaches

The emergence of novel data-analysis and bioinformatic approaches have made statistical investigation into phylogenomics and population genetics more feasible (**Figure 2D**). For example, random forest classification of fungal ITS barcodes using the ITS2VEC software program makes species identification easier for large datasets as it relies on dimensionality reduction algorithms (Wang et al., 2020). Similarly, in the study of human-associated fungal communities, the software program HumanMycobiomeScan
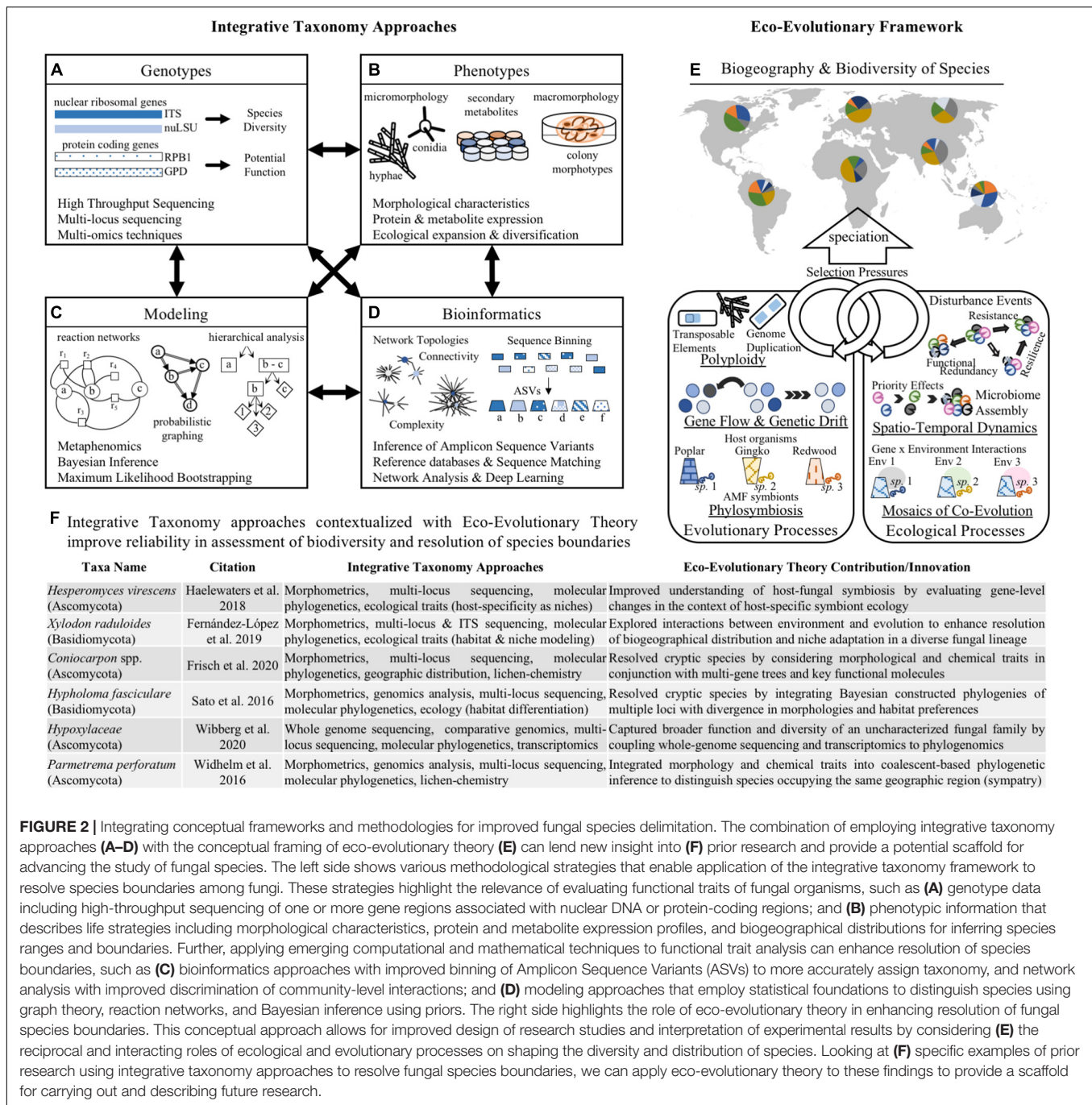
FIGURE 2 | Integrating conceptual frameworks and methodologies for improved fungal species delimitation. The combination of employing integrative taxonomy approaches **(A–D)** with the conceptual framing of eco-evolutionary theory **(E)** can lend new insight into **(F)** prior research and provide a potential scaffold for advancing the study of fungal species. The left side shows various methodological strategies that enable application of the integrative taxonomy framework to resolve species boundaries among fungi. These strategies highlight the relevance of evaluating functional traits of fungal organisms, such as **(A)** genotype data including high-throughput sequencing of one or more gene regions associated with nuclear DNA or protein-coding regions; and **(B)** phenotypic information that describes life strategies including morphological characteristics, protein and metabolite expression profiles, and biogeographical distributions for inferring species ranges and boundaries. Further, applying emerging computational and mathematical techniques to functional trait analysis can enhance resolution of species boundaries, such as **(C)** bioinformatics approaches with improved binning of Amplicon Sequence Variants (ASVs) to more accurately assign taxonomy, and network analysis with improved discrimination of community-level interactions; and **(D)** modeling approaches that employ statistical foundations to distinguish species using graph theory, reaction networks, and Bayesian inference using priors. The right side highlights the role of eco-evolutionary theory in enhancing resolution of fungal species boundaries. This conceptual approach allows for improved design of research studies and interpretation of experimental results by considering **(E)** the reciprocal and interacting roles of ecological and evolutionary processes on shaping the diversity and distribution of species. Looking at **(F)** specific examples of prior research using integrative taxonomy approaches to resolve fungal species boundaries, we can apply eco-evolutionary theory to these findings to provide a scaffold for carrying out and describing future research.

(Soverini et al., 2019) enables direct analysis of metagenomic reads without any pre-processing steps to detect and extract fungal sequences. Improved access to computational techniques allows life science researchers to apply complex mathematics and bioinformatic approaches without extensive knowledge of computational coding.

Further, deep learning and network-based approaches are being employed to classify and cluster fungal sequences from metagenomic data (Vu et al., 2020). These approaches, developed with the implementation of convolutional neural networks,

quickly and accurately assign taxonomic ranks, outperforming more traditional nucleotide read identification methods such as BLAST (Altschul et al., 1990) and RDP (Wang et al., 2007). Improved clustering methods contributes to more reliable characterization of fungal biodiversity across diverse systems. A recent analysis of bioinformatics software showed that the choice of data analysis pipeline influences the accuracy and reliability of taxonomy assignment in plant and soil associated fungal communities (Pauvert et al., 2019). The authors of that study conclude that detection of Amplicon

Sequence Variants (ASVs) using DADA2 software (Callahan et al., 2016) outperforms more than 350 other software parameter combinations, yielding higher species richness and more detailed community composition data for fungi (Pauvert et al., 2019). New bioinformatics tools are providing novel and innovative strategies for classifying organisms from complex communities, thereby improving understanding of fungal biodiversity and species variability. Improved insight into biodiversity facilitates the development of experiments to evaluate the functional consequences, ecological drivers, and evolutionary trajectories of fungal species from a diverse range natural and agricultural of systems.

DISCUSSION: WAYS FORWARD

Distinguishing fungal species boundaries is an important practical concern for researchers seeking to address questions related to biodiversity, species interactions, biogeography, ecological processes, and evolutionary dynamics. Conceptually, eco-evolutionary theory recognizes the interplay between ecological processes and evolutionary dynamics in shaping genetic, phenotypic, and functional traits (Brunner et al., 2019; McDonald et al., 2020). Similarly, integrative taxonomy combines technical evidence from many types of observations (e.g., ecological, phylogenetic, morphological, etc.) in order to build coalescent or holistic conclusions (Fujita et al., 2012; Haelewaters et al., 2018; Fernández-López et al., 2019). Together, these two approaches may improve delimitation of species by acknowledging the influence of eco-evolutionary relationships in shaping spatio-temporal variability among and between species (**Figure 2**). In practice, this may involve using a combination of molecular and morphological techniques to elucidate phylogenetic relationships among cryptic species (Zhou et al., 2014), or to discern causal agents of plant (Kusai et al., 2016) and animal (Haelewaters et al., 2018) diseases. Furthermore, combining multi-locus sequencing with morphological (e.g., spore or conidia measurements) and functional traits (proteome, biochemical compounds, etc.) is improving resolution of diverse fungal groups (Widhelm et al., 2016; Fernández-López et al., 2019; Frisch et al., 2020). Bioinformatics tools, such as amplicon sequence inference (Callahan et al., 2016), and modeling approaches, such as Bayesian hierarchical analysis (Wyss et al., 2016), maximum likelihood probabilistic graphing (Widhelm et al., 2016), and fuzzy logic (Yusof et al., 2013), provide a new toolkit for life scientists to explore the concept of species in fungi. The application of integrative taxonomy strategies to evaluate multiple genotypic and phenotypic factors supports drawing more reliable, robust, and reproducible distinctions between fungal species (Widhelm et al., 2016; Haelewaters et al., 2018; Fernández-López et al., 2019). These approaches are more closely aligned with eco-evolutionary theory, wherein abiotic, biotic, ecological, and environmental interactions contribute to evolutionary trajectories that shape fungal lineages.

Moving forward, we suggest that integrative taxonomy approaches and eco-evolutionary theory will improve research and interpretation of fungal evolution across diverse systems (**Figure 2E**). Specifically, we propose the use of integrative

taxonomy approaches to improve surveys of fungal biodiversity across natural and agricultural systems (Bickford et al., 2007; Möller and Stukenbrock, 2017). This may include exploring spatio-temporal variability among fungal communities in response to environmental stressors (Hawkes and Kiett, 2015; Brooks et al., 2016; Nilsson et al., 2019) and changing climate (Rudgers et al., 2020); as well as investigation of strain-specific responses to perturbations under controlled conditions (Marcelino et al., 2019; Reiter et al., 2020; Varona et al., 2020). These surveys will be particularly important for furthering the understanding of the continuum of symbiosis with applications for human and agricultural disease management (Han et al., 2001; Cai et al., 2011; Kusai et al., 2016). Secondly, we encourage the application of eco-evolutionary theory to conceptualize experimental manipulation of fungal species interactions. Experimental evolution studies are particularly well-suited to explore how fungal communities may shape host health (Koskella and Bergelson, 2020; Morella et al., 2020) and contribute to microbe-microbe interactions that drive speciation (McDonald et al., 2020). Additionally, designing and interpreting research outcomes with concepts such as phylosymbiosis (Kohl, 2020; McDonald et al., 2020) and the geographic mosaic of co-evolution (Medeiros et al., 2018; Fernandes et al., 2019) can improve understanding of drivers of microbial ecology and evolution. This approach is particularly useful for evaluating core microbiomes associated with specific host organisms (Saikkonen et al., 2016; Lofgren et al., 2018), and improving understanding of the functional consequences arising from microbial community interactions (Wei et al., 2019; Trivedi et al., 2020). While our discussion is focused specifically on fungi, many of the analysis techniques and tools are applicable to other microbial lineages, including bacteria and archaea. Thus, it is our hope that the discussion provided here may garner further interest among microbiologists that seek to elucidate the ecological and evolutionary origins of single organisms as well as eco-evolutionary processes shaping entire communities.

AUTHOR CONTRIBUTIONS

AS and JH were involved in the initial conceptualization of this manuscript. AS led the literature review and writing of the first draft. AS, KS, and AQ were involved in the visualization of concepts. AS, KS, AQ, and JH provided revisions and additional conceptual input into the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

to the discussion of some of the topics presented here and he added his spirited debate to many of our laboratory meetings. We would also like to acknowledge the contributions of Erin Carr and Samantha Reese to the initial discussion of the

history of applying species concepts to microorganisms and fungi in particular. We also want to thank the reviewers who provided feedback and comments on the initial submission of this manuscript.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Avise, J. C., and Wollenberg, K. (1997). Phylogenetics and the origin of species. *Proc. Natl. Acad. Sci.* 94, 7748–7755. doi: 10.1073/pnas.94.15.7748

Berruti, A., Desirò, A., Visentin, S., Zecca, O., and Bonfante, P. (2017). ITS fungal barcoding primers versus 18S AMF−specific primers reveal similar AMF−based diversity patterns in roots and soils of three mountain vineyards. *Environ. Microbiol. Rep.* 9, 658–667. doi: 10.1111/1758-2229.12574

Bérubé, J. A., and Dessureault, M. (1989). Morphological studies of the *Armillaria mellea* complex: two new species, *A. gemina* and *A. calvescens. Mycologia* 81, 216–225.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., et al. (2007). Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* 22, 148–155.

Börstler, B., Raab, P. A., Thiéry, O., Morton, J. B., and Redecker, D. (2008). Genetic diversity of the arbuscular mycorrhizal fungus Glomus intraradices as determined by mitochondrial large subunit rRNA gene sequences is considerably higher than previously expected. *New Phytol.* 180, 452–465. doi: 10.1111/j.1469-8137.2008.02574.x

Brooks, A. W., Kohl, K. D., Brucker, R. M., van Opstal, E. J., and Bordenstein, S. R. (2016). Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biol.* 15:e1002587. doi: 10.1371/journal.pbio.2000225

Brunner, F. S., Deere, J. A., Egas, M., Eizaguirre, C., and Raeymaekers, J. A. (2019). The diversity of eco−evolutionary dynamics: comparing the feedbacks between ecology and evolution across scales. *Funct. Ecol.* 33, 7–12. doi: 10.1111/1365-2435.13268

Bruns, T. D., Corradi, N., Redecker, D., Taylor, J. W., and Öpik, M. (2018). Glomeromycotina: what is a species and why should we care? *New Phytol.* 220, 963–967. doi: 10.1111/nph.14913

Cai, L., Giraud, T., Zhang, N., Begerow, D., Cai, G., and Shivas, R. G. (2011). The evolution of species concepts and species recognition criteria in plant pathogenic fungi. *Fungal Diversity* 50:121. doi: 10.1007/s13225-011-0127-8

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Chowdhury, T. R., Lee, J.-Y., Bottos, E. M., Brislawn, C. J., White, R. A., Bramer, L. M., et al. (2019). Metaphenomic responses of a native prairie soil microbiome to moisture perturbations. *mSystems* 4:e00061-19. doi: 10.1128/msystems.00061-19

Cordovez, V., Dini-Andreote, F., Carrión, V. J., and Raaijmakers, J. M. (2019). Ecology and evolution of plant microbiomes. *Annu. Rev. Microbiol.* 73, 69–88. doi: 10.1146/annurev-micro-090817-062524

Dayrat, B. (2005). Towards integrative taxonomy. *Biol. J. Linnean Soc.* 85, 407–415. doi: 10.1111/j.1095-8312.2005.00503.x

de Queiroz, K. (2005). Ernst Mayr and the modern concept of species. *PNAS* 102, 6600–6607. doi: 10.1073/pnas.0502030102

de Queiroz, K. (2007). Species concepts and species delimitation. *System. Biol.* 56, 879–886. doi: 10.1080/10635150701701083

Dugan, F. M., and Everhart, S. (2016). Cryptic species: a leitmotif of contemporary mycology has challenges and benefits for plant pathologists. *Plant Health Prog.* 17, 250–253. doi: 10.1094/PHP-RV-16-0046

Fernandes, L. D., Lemos-Costa, P., Guimaraþes, P. R., Thompson, J. N., and de Aguiar, M. A. M. (2019). Coevolution creates complex mosaics across large landscapes. *Am. Nat.* 194, 217–229. doi: 10.1086/704157

Fernández-López, J., Telleria, M. T., Dueñas, M., Wilson, A. W., Padamsee, M., Buchanan, P. K., et al. (2019). Addressing the diversity of *Xylodon raduloides*

complex through integrative taxonomy. *IMA Fungus* 10:9. doi: 10.1186/s43008-019-0010-x

Frisch, A., Moen, V. S., Grube, M., and Bendiksby, M. (2020). Integrative taxonomy confirms three species of *Coniocarpon* (Arthoniaceae) in Norway. *Mycokeys* 62, 27–51. doi: 10.3897/mycokeys.62.48480

Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., and Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol. Evol.* 27, 480–488. doi: 10.1016/j.tree.2012.04.012

Haelewaters, D., de Kesel, A., and Pfister, D. H. (2018). Integrative taxonomy reveals hidden species within a common fungal parasite of ladybirds. *Sci. Rep.* 8:15966. doi: 10.1038/s41598-018-34319-5

Han, Y., Liu, X., Benny, U., Kistler, H. C., and Vanetten, H. D. (2001). Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca. Plant J.* 25, 1–11. doi: 10.1046/j.1365-313x.2001.00969.x

Hart, M. M., Aleklett, K., Chagnon, P. L., Egan, C., Ghignone, S., Helgason, T., et al. (2015). Navigating the labyrinth: a guide to sequence−based, community ecology of arbuscular mycorrhizal fungi. *New Phytol.* 207, 235–247. doi: 10.1111/nph.13340

Hawkes, C. V., and Kiett, T. H. (2015). Resilience vs. historical contingency in microbial responses to environmental change. *Ecol. Lett.* 18, 612–625. doi: 10.1111/ele.12451

Herr, J. R., Öpik, M., and Hibbett, D. S. (2015). Towards the unification of sequence-based classification and sequence-based identification of host-associated microorganisms. *New Phytol.* 205, 27–31. doi: 10.1111/nph.13180

Hibbett, D., Abarenkov, K., Kõljalg, U., Öpik, M., Chai, B., Cole, J., et al. (2016). Sequence-based classification and identification of Fungi. *Mycologia* 108, 1049–1068. doi: 10.3852/16-130

Howard, S. J., Harrison, E., Bowyer, P., Varga, J., and Denning, D. W. (2011). Cryptic species and azole resistance in the *Aspergillus niger* complex. *Antimicrob. Agents Chemother.* 55, 4802–4809. doi: 10.1128/AAC.00304-11

Inderbitzin, P., Robbertse, B., and Schoch, C. L. (2020). Species identification in plant-associated prokaryotes and fungi using DNA. *Phytobiomes J.* 4, 103–114. doi: 10.1094/pbiomes-12-19-0067-rvw

Kobmoo, N., Mongkolsamrit, S., Arnamnart, N., Luangsa-ard, J. J., and Giraud, T. (2019). Population genomics revealed cryptic species within host-specific zombie-ant fungi (*Ophiocordyceps unilateralis*). *Mol. Phylogenet. Evol.* 140:106580. doi: 10.1016/j.ympev.2019.106580

Kohl, K. D. (2020). Ecological and evolutionary mechanisms underlying patterns of phylosymbiosis in host-associated microbial communities. *Philos. Trans. Royal Soc. B* 375:20190251. doi: 10.1098/rstb.2019.0251

Koskella, B., and Bergelson, J. (2020). The study of host–microbiome (co) evolution across levels of selection. *Philos. Trans. Royal Soc. B* 375:20190604. doi: 10.1098/rstb.2019.0604

Kusai, N. A., Mior Zakuan Azmi, M., Zulkifly, S., Yusof, M. T., and Mohd Zainudin, N. A. I. (2016). Morphological and molecular characterization of *Curvularia* and related species associated with leaf spot disease of rice in Peninsular Malaysia. *Rendiconti Lincei* 27, 205–214. doi: 10.1007/s12210-015-0458-6

Kuske, C. R., Hesse, C. N., Challacombe, J. F., Cullen, D., Herr, J. R., Mueller, R. C., et al. (2015). Prospects and challenges for fungal metatranscriptomics of complex communities. *Fungal Ecol.* 14, 133–137. doi: 10.1016/j.funeco.2014.12.005

Lamarre, G. P. A., Decaëns, T., Rougerie, R., Barbut, J., Dewaard, J. R., Hebert, P. D. N., et al. (2016). An integrative taxonomy approach unveils unknown and threatened moth species in Amazonian rainforest fragments. *Insect Conserv. Diversity* 9, 475–479. doi: 10.1111/icad.12187

Latouche, G. N., Daniel, H.-M., Lee, O. C., Mitchell, T. G., Sorrell, T. C., and Meyer, W. (1997). Comparison of use of phenotypic and genotypic characteristics for identification of species of the Anamorph Genus *Candida* and related

teleomorph yeast species. *J. Clin. Microbiol.* 35, 3171–3180. doi: 10.1128/jcm. 35.12.3171-3180.1997

Li, A.-H., Yuan, F.-X., Groenewald, M., Bensch, K., Yurkov, A. M., Li, K., et al. (2020). Diversity and phylogeny of basidiomycetous yeasts from plant leaves and soil: proposal of two new orders, three new families, eight new genera and one hundred and seven new species. *Stud. Mycol.* 96, 17–140. doi: 10.1016/j. simyco.2020.01.002

Libkind, D., Čadež, N., Opulente, D. A., Langdon, Q. K., Rosa, C. A., Sampaio, J. P., et al. (2020). Towards yeast taxogenomics: lessons from novel species descriptions based on complete genome sequences. *FEMS Yeast Res.* 20:foaa042. doi: 10.1093/femsyr/foaa042

Lofgren, L. A., LeBlanc, N. R., Certano, A. K., Nachtigall, J., LaBine, K. M., Riddle, J., et al. (2018). *Fusarium graminearum*: pathogen or endophyte of North American grasses? *New Phytol.* 217, 1203–1212. doi: 10.1111/nph.14894

Lücking, R., Aime, M. C., Robbertse, B., Miller, A., Ariyawansa, H., Aoki, T., et al. (2020). Unambiguous identification of fungi: where do we stand and how accurate and precise is fungal DNA barcoding? *IMA Fungus* 11:14. doi: 10.1186/s43008-020-00033-z

Manamgoda, D. S., Cai, L., McKenzie, E. H. C., Crous, P. W., Madrid, H., Chukeatirote, E., et al. (2012). A phylogenetic and taxonomic re-evaluation of the *Bipolaris - Cochliobolus - Curvularia* complex. *Fungal Diversity* 56, 131–144. doi: 10.1007/s13225-012-0189-2

Marcelino, V. R., Irinyi, L., Eden, J. S., Meyer, W., Holmes, E. C., and Sorrell, T. C. (2019). Metatranscriptomics as a tool to identify fungal species and subspecies in mixed communities – A proof of concept under laboratory conditions. *IMA Fungus* 10:12. doi: 10.1186/s43008-019-0012-8

Martiny, J. B., Jones, S. E., Lennon, J. T., and Martiny, A. C. (2015). Microbiomes in light of traits: a phylogenetic perspective. *Science* 350:aac9323. doi: 10.1126/science.aac9323

Matute, D. R., and Sepúlveda, V. E. (2019). Fungal species boundaries in the genomics era. *Fungal Genet. Biol.* 131:103249. doi: 10.1016/j.fgb.2019.103249

Mayr, E. (1996). What is a species, and what is not? *Philos. Sci.* 63, 262–277. doi: 10.1086/289912

McDonald, J. E., Marchesi, J. R., and Koskella, B. (2020). Application of ecological and evolutionary theory to microbiome community dynamics across systems. *Proc. Royal Soc. B* 287:20202886. doi: 10.1098/rspb.2020.2886

Medeiros, L. P., Garcia, G., Thompson, J. N., and Guimarães, P. R. (2018). The geographic mosaic of coevolution in mutualistic networks. *Proc. Natl. Acad. Sci.* 115, 12017–12022. doi: 10.1073/pnas.1809088115

Mehrabi, R., Bahkali, A. H., Abd-Elsalam, K. A., Moslem, M., ben M'Barek, S., Gohari, A. M., et al. (2011). Horizontal gene and chromosome transfer in plant pathogenic fungi affecting host range. *FEMS Microbiol. Rev.* 35, 542–554. doi: 10.1111/j.1574-6976.2010.00263.x

Michener, C. D. (1970). Diverse approaches to systematics. *Evol. Biol.* 4, 1–38. doi: 10.1017/cbo9780511608377.003

Möller, M., and Stukenbrock, E. H. (2017). Evolution and genome architecture in fungal plant pathogens. *Nat. Rev. Microbiol.* 15, 756–771.

Morella, N. M., Weng, F. C. H., Joubert, P. M., Metcalf, C. J. E., Lindow, S., and Koskella, B. (2020). Successive passaging of a plant-associated microbiome reveals robust habitat and host genotype-dependent selection. *Proc. Natl. Acad. Sci.* 117, 1148–1159. doi: 10.1073/pnas.1908600116

Muszewska, A., Steczkiewicz, K., Stepniewska-Dziubinska, M., and Ginalski, K. (2019). Transposable elements contribute to fungal genes and impact fungal lifestyle. *Sci. Rep.* 9:4307. doi: 10.1038/s41598-019-40965-0

Nakasone, K. K. (1996). Morphological and molecular studies on *Auriculariopsis albomellea* and *Phlebia albida* and a reassessment of *A. ampla*. *Mycologia* 88, 762–775. doi: 10.2307/3760971

Nie, Y., Yu, D.-S., Wang, C.-F., Liu, X.-Y., and Huang, B. (2020). A taxonomic revision of the genus *Conidiobolus* (*Ancylistaceae, Entomophthorales*): four clades including three new genera. *Mycokeys* 66, 55–81. doi: 10.3897/mycokeys. 66.46575

Nilsson, R. H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P., and Tedersoo, L. (2019). Mycobiome diversity: high-throughput sequencing and identification of fungi. *Nat. Rev. Microbiol.* 17, 95–109. doi: 10.1038/s41579-018-0116-y

Nixon, K. C., and Wheeler, Q. D. (1990). An amplification of the phylogenetic species concept. *Cladistics* 6, 211–223.

Pauvert, C., Buée, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., et al. (2019). Bioinformatics matters: the accuracy of plant and soil fungal community

data is highly dependent on the metabarcoding pipeline. *Fungal Ecol.* 41, 23–33. doi: 10.1016/j.funeco.2019.03.005

Prata, E. M. B., Sass, C., Rodrigues, D. P., Domingos, F. M. C. B., Specht, C. D., Damasco, G., et al. (2018). Towards integrative taxonomy in Neotropical botany: disentangling the *Pagamea guianensis* species complex (*Rubiaceae*). *Botanical J. Linnean Soc.* 188, 213–231. doi: 10.1093/botlinnean/boy051

Pugedo, M. L., de Andrade Neto, F. R., Pessali, T. C., Birindelli, J. L. O., and Carvalho, D. C. (2016). Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha River Basin. *Genetica* 144, 341–349. doi: 10.1007/s10709-016-9903-4

Reich, M., and Labes, A. (2017). How to boost marine fungal research: a first step towards a multidisciplinary approach by combining molecular fungal ecology and natural products chemistry. *Mar. Genomics* 36, 57–75. doi: 10.1016/j. margen.2017.09.007

Reiter, N., Phillips, R. D., Swarts, N. D., Wright, M., Holmes, G., Sussmilch, F. C., et al. (2020). Specific mycorrhizal associations involving the same fungal taxa in common and threatened *Caladenia* (*Orchidaceae*): implications for conservation. *Ann. Bot.* 126, 943–955. doi: 10.1093/aob/mcaa116

Rudgers, J. A., Afkhami, M. E., Bell-Dereske, L., Chung, Y. A., Crawford, K. M., Kivlin, S. N., et al. (2020). Climate disruption of plant-microbe interactions. *Annual Rev. Ecol. Evol. System.* 51, 561–586. doi: 10.1146/annurev-ecolsys-011720-090819

Saikkonen, K., Nissinen, R., and Helander, M. (2020). Toward comprehensive plant microbiome research. *Front. Ecol. Evol.* 8:61. doi: 10.3389/fevo.2020.00061

Saikkonen, K., Young, C. A., Helander, M., and Schardl, C. L. (2016). Endophytic *Epichloë* species and their grass hosts: from evolution to applications. *Plant Mol. Biol.* 90, 665–675. doi: 10.1007/s11103-015-0399-6

Sato, H., Ohta, R., and Murakami, N. (2020). Molecular prospecting for cryptic species of the *Hypholoma fasciculare* complex: toward the effective and practical delimitation of cryptic macrofungal species. *Sci. Rep.* 10:13224. doi: 10.1038/s41598-020-70166-z

Schlick-Steiner, B. C., Steiner, F. M., Seifert, B., Stauffer, C., Christian, E., and Crozier, R. H. (2010). Integrative taxonomy: a multisource approach to exploring biodiversity. *Annu. Rev. Entomol.* 55, 421–438. doi: 10.1146/annurev-ento-112408-085432

Schmidt, T. S., Matias Rodrigues, J. F., and Von Mering, C. (2014). Ecological consistency of SSU rRNA-based operational taxonomic units at a global scale. *PLoS Comput. Biol.* 10:e1003594. doi: 10.1371/journal.pcbi.1003594

Somma, S., Amatulli, M. T., Masiello, M., Moretti, A., and Logrieco, A. F. (2019). *Alternaria* species associated to wheat black point identified through a multilocus sequence approach. *Int. J. Food Microbiol.* 293, 34–43. doi: 10.1016/j.ijfoodmicro.2019.01.001

Soverini, M., Turroni, S., Biagi, E., Brigidi, P., Candela, M., and Rampelli, S. (2019). HumanMycobiomeScan: a new bioinformatics tool for the characterization of the fungal fraction in metagenomic samples. *BMC Genom.* 20:496. doi: 10.1186/s12864-019-5883-y

Taylor, J. W. (2011). One Fungus = One Name: DNA and fungal nomenclature twenty years after PCR. *IMA Fungus* 2, 113–120. doi: 10.5598/imafungus.2011.02.02.01

Taylor, J. W., Jacobson, D. J., Kroken, S., Kasuga, T., Geiser, D. M., Hibbett, D. S., et al. (2000). Phylogenetic species recognition and species concepts in fungi. *Fungal Genet. Biol.* 31, 21–32. doi: 10.1006/fgbi.2000.1228

Tekpinar, A. D., and Kalmer, A. (2019). Utility of various molecular markers in fungal identification and phylogeny. *Nova Hedwigia* 109, 187–224. doi: 10.1127/nova_hedwigia/2019/0528

Tisserant, E., Malbreil, M., Kuo, A., Kohler, A., Symeonidi, A., Balestrini, R., et al. (2013). Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc. Natl. Acad. Sci.* 110, 20117–20122. doi: 10.1073/pnas.1313452110

Todd, R. T., Forche, A., and Selmecki, A. (2017). Ploidy variation in fungi: polyploidy, aneuploidy, and genome evolution. *Microbiol. Spectr.* 5, 599–618. doi: 10.1128/microbiolspec.funk-0051-2016

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., and Singh, B. K. (2020). Plant–microbiome interactions: from community assembly to plant health. *Nat. Rev. Microbiol.* 18, 607–621. doi: 10.1038/s41579-020-0412-1

Valverde-Bogantes, E., Bianchini, A., Herr, J. R., Rose, D. J., Wegulo, S. N., and Hallen-Adams, H. E. (2020). Recent population changes of *Fusarium* head

blight pathogens: drivers and implications. *Can. J. Plant Pathol.* 42, 315–329. doi: 10.1080/07060661.2019.1680442

Van Valen, L. (1976). Ecological species, multispecies, and oaks. *Taxon* 25, 233–239. doi: 10.2307/1219444

Varona, S., Lavín, J. L., and Oguiza, J. A. (2020). Secretomes of medically important fungi reflect morphological and phylogenetic diversity. *Fungal Biol.* 124, 915–923. doi: 10.1016/j.funbio.2020.07.011

Vu, D., Groenewald, M., and Verkley, G. (2020). Convolutional neural networks improve fungal classification. *Sci. Rep.* 10:12628. doi: 10.1038/s41598-020-69245-y

Wang, C., Zhang, Y., and Han, S. (2020). Its2vec: fungal species identification using sequence embedding and random forest classification. *Biomed Res. Int.* 2020, 1–11. doi: 10.1155/2020/2468789

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wei, Z., Gu, Y., Friman, V.-P., Kowalchuk, G. A., Xu, Y., Shen, Q., et al. (2019). Initial soil microbiome composition and functioning predetermine future plant health. *Sci. Adv.* 5:eaaw0759. doi: 10.1126/sciadv.aaw0759

Wibberg, D., Stadler, M., Lambert, C., Bunk, B., Spröer, C., Rückert, C., et al. (2020). High quality genome sequences of thirteen *Hypoxylaceae* (Ascomycota) strengthen the phylogenetic family backbone and enable the discovery of new taxa. *Fungal Diversity* 106, 7–28. doi: 10.1007/s13225-020-00447-5

Widhelm, T. J., Egan, R. S., Bertoletti, F. R., Asztalos, M. J., Kraichak, E., Leavitt, S. D., et al. (2016). Picking holes in traditional species delimitations: an integrative taxonomic reassessment of the *Parmotrema perforatum* group (*Parmeliaceae*, Ascomycota). *Bot. J. Linnean Soc.* 182, 868–884.

Wyss, T., Masclaux, F. G., Rosikiewicz, P., Pagni, M., and Sanders, I. R. (2016). Population genomics reveals that within-fungus polymorphism is common and maintained in populations of the mycorrhizal fungus *Rhizophagus irregularis*. *ISME J.* 10, 2514–2526. doi: 10.1038/ismej.2016.29

Xu, J. (2020). Fungal species concepts in the genomics era. *Genome* 63, 459–468. doi: 10.1139/gen-2020-0022

Yusof, R., Khalid, M., and Khairuddin, A. S. M. (2013). Fuzzy logic-based pre-classifier for tropical wood species recognition system. *Mach. Vis. Appl.* 24, 1589–1604. doi: 10.1007/s00138-013-0526-9

Zachos, F. E. (2016). *Species Concepts in Biology (Vol. 801)*. Cham: Springer.

Zhou, Y. J., Zhang, J., Wang, X. D., Yang, L., Jiang, D. H., Li, G. Q., et al. (2014). Morphological and phylogenetic identification of *Botrytis sinoviticola*, a novel cryptic species causing gray mold disease of table grapes (*Vitis vinifera*) in China. *Mycologia* 106, 43–56. doi: 10.3852/13-032

Toward Microbial Recycling and Upcycling of Plastics: Prospects and Challenges

Jo-Anne Verschoor[1], Hadiastri Kusumawardhani[2], Arthur F. J. Ram[1] and Johannes H. de Winde[1]*

[1] *Molecular Biotechnology, Institute of Biology, Leiden University, Leiden, Netherlands,* [2] *Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland*

Annually, 400 Mt of plastics are produced of which roughly 40% is discarded within a year. Current plastic waste management approaches focus on applying physical, thermal, and chemical treatments of plastic polymers. However, these methods have severe limitations leading to the loss of valuable materials and resources. Another major drawback is the rapid accumulation of plastics into the environment causing one of the biggest environmental threats of the twenty-first century. Therefore, to complement current plastic management approaches novel routes toward plastic degradation and upcycling need to be developed. Enzymatic degradation and conversion of plastics present a promising approach toward sustainable recycling of plastics and plastics building blocks. However, the quest for novel enzymes that efficiently operate in cost-effective, large-scale plastics degradation poses many challenges. To date, a wide range of experimental set-ups has been reported, in many cases lacking a detailed investigation of microbial species exhibiting plastics degrading properties as well as of their corresponding plastics degrading enzymes. The apparent lack of consistent approaches compromises the necessary discovery of a wide range of novel enzymes. In this review, we discuss prospects and possibilities for efficient enzymatic degradation, recycling, and upcycling of plastics, in correlation with their wide diversity and broad utilization. Current methods for the identification and optimization of plastics degrading enzymes are compared and discussed. We present a framework for a standardized workflow, allowing transparent discovery and optimization of novel enzymes for efficient and sustainable plastics degradation in the future.

Keywords: plastics, enzymes, biorecycling, biodegradation, comprehensive workflow

INTRODUCTION—THE WORLD OF PLASTICS

Plastics are man-made polymers that are used for many applications. Their flexibility, strength and erosion resistance allow plastics to be suitable material for a broad spectrum of applications (Wang et al., 2018). Of the almost 400 Mt of plastics produced, 40 percent is used in single-use applications, leading to a significant amount of waste (PlasticEurope-Association of Plastics Manufacturers, 2020). Current waste management systems mostly consist of accumulation in landfills, incineration for energy recovery, and recycling. Although the fraction of plastic that is

being recycled is increasing, the vast majority ends up in the incinerator or in landfills polluting the environment, rather than being reused (Geyer et al., 2017). Especially in Asia millions of metric tons of plastic waste are managed poorly resulting in a high likelihood of leakage into the environment (d'Ambrières, 2019).

Landfilling, careless dumping and other sources of leakage such as the release of microfibres into the environment, causes the accumulation of plastics in both terrestrial and marine ecosystems (Henry et al., 2019). The leakage of plastics into the environment has been shown to have several negative effects on flora and fauna. Environmentally, plastic pollution has a plethora of effects. The presence of plastic films in soil has been shown to decrease crop yield by 3% meaning that this could have significant effects on food security (Zhang et al., 2020). The fragmentation of larger plastic products, when exposed to abiotic factors, into so-called nano (<100 nm) and/or microplastics (<5 mm) is also a risk that needs to be considered (Science Advice for Policy by European Academies [SAPEA], 2019; Chamas et al., 2020). Even though the SAPEA report from 2019 states that micro and nano plastics do not endanger human health and the environment, the evidence seemed limited at that time (Science Advice for Policy by European Academies [SAPEA], 2019). Recently, several studies and reviews regarding the effects of nano and microplastics on the environment have been published and describe that there is a negative effect of these particles in the environment (Shen et al., 2019; Iqbal et al., 2020; Selonen et al., 2020; Shiu et al., 2020; Tiwari et al., 2020; Guerrera et al., 2021; Kumar et al., 2021; Wang et al., 2022; Zhang et al., 2022). The effects of these particles existent from neurotoxicity to oxidative stress and even lethality (Wang et al., 2021). The further degradation of monomers and oligomers of plastics can also result in negative health effects. Exposure of pregnant mice to bisphenol A and S resulted in the disruption of their placentas (Mao et al., 2020). Finally, the release of harmful volatile organic compounds upon photo-degradation of plastics has been neglected until recently meaning that there might be more aspects of plastic degradation threatening our ecosystems (Lomonaco et al., 2020).

The results so far show that plastic pollution harms our planet, meaning that we should aim to decrease plastic pollution. According to a recent model (Lau et al., 2020), several approaches need to be taken in parallel to decrease the rapid accumulation of plastics in the environment: (1) plastic consumption needs to be reduced by limiting single-use plastic products; (2) reuse should be stimulated; (3) waste collection and recycling should be increased; (4) landfilling should eventually be eliminated.

To increase recycling, more efficient collecting, sorting and recycling methods should be developed. A promising and novel recycling approach is biorecycling, which relies on the enzymatic degradation of plastics. After degradation, the monomers are extracted and used to manufacture the original plastic known as recycling or be used as building blocks for other compounds with a higher value (upcycling), without loss of quality.

Sorting waste materials after collection is an essential step in the implementation of recycling as the main method for domestic plastic waste management. As described by Hahladakis and Iacovidou (2019), the lack of efficient sorting may cause unwanted blending of polymers, ultimately yielding new polymeric materials with unwanted characteristics. In addition, accidental mixing of biobased plastics with petrochemical-based plastics compromises the recyclability of the end products (Hahladakis and Iacovidou, 2019). Currently, automated sorting is being employed based on material or color focused separation. Both methods have clear limitations and contamination cannot be avoided (Hahladakis and Iacovidou, 2019). Automated sorting of plastics using near-infrared spectroscopy may separate various household plastics with an accuracy of 99% (Wu et al., 2020). However, this method still needs further improvement since it is not suitable for dark-colored plastics. Moreover, interference and effects of surface coating and surface contamination still need further characterization (Wu et al., 2020).

Despite the limitations in plastic sorting, various recycling methods are in use. For large scale solid waste, mechanical recycling is the preferred method. In this method, all organic residues are washed off which is followed by shredding, melting, and remolding of the polymeric material. Since household waste consists of a wide variety of polymers, sometimes contaminated with coloring, other plastics or metals, the quality of the reprocessed matter is relatively low (Eriksen et al., 2018; Shahid Kashif et al., 2021). To overcome this low quality, mixing with virgin plastic is required to achieve the desirable characteristics for industrial use (Garcia and Robertson, 2017). Another approach is chemical recycling, yielding higher quality products but having significantly higher costs. Examples of chemical recycling are the lysis of glycols and methanol (Grigore, 2017). However, the chemical structure and diversity as well as the use of various plastics mixtures, compromise the efficiency of mechanical recycling (Garcia and Robertson, 2017), and are the main cause for the fact that most of the solid waste is not being recycled. Since energy recovery and mechanical recycling have their limits, large amounts of plastics continue to be landfilled.

Importantly, current gaps in literature concern the still limited characterization of efficient plastic depolymerizing enzymes which are required for the degradation of plastics into their corresponding monomers, which subsequently can be used for recycling and upcycling. In this review, we addressed the enzymatic degradation of plastic. We divide the overall approach into "upstream" and "downstream" approaches. The upstream approaches are dealing with enzymatic depolymerization of plastics while the downstream approaches are focused on upcycling and merging plastic waste management into the bioeconomy. We focus our discussion on how to identify and characterize novel enzymes, especially for the important aforementioned upstream approaches, and how to assess the suitability and efficiency of enzymatic degradation processes, providing a comprehensive workflow which so far is lacking in literature. Thus, our review is substantially different from other recent reviews in the field of microbial plastic degradation (Wang et al., 2021; Sarkar et al., 2022; Zhou et al., 2022). We refer to other recent reviews in the field of microbial plastic degradation reviewing other topics within this rapidly evolving field (Wang et al., 2021; Sarkar et al., 2022; Zhou et al., 2022).

DIVERSITY OF SYNTHETIC POLYMERS

Currently, over twenty different types of plastics are on the market, the majority developed and synthesized from petrochemical feedstocks (PlasticEurope-Association of Plastics Manufacturers, 2020). Clear differences can be observed in the biodegradability of different plastics. Important determinants for the degradability of plastics are the nature of the chemical bonds constituting the polymeric structure, and the crystallinity of the plastic. Recently, the development of biobased plastics has taken an important leap. Biobased plastics are manufactured using biological feedstocks, like starch or lactic acid. However, this does not entail that they can automatically be regarded as biodegradable. Some biobased plastics, like polylactic acid, are highly recalcitrant and hence, are difficult to degrade.

Different plastics may indeed be structurally different, however, the covalent chemical bonds connecting the constituting monomers can be similar. In **Figure 1**, several examples are given of different plastics containing similar chemical bonds connecting the structurally different monomers. We hypothesize that the biodegradability of these plastics correlates with the occurrence and abundance of the specific chemical bonds in natural substrates. The more likely a bond is to be present in natural substrates, the more likely it may be recognized and broken in the synthetic polymer. It is expected that the enzymes' ability to degrade specific bonds can either be coincidental or has evolved depending on the nature of the enzyme. Determining the prevalent chemical bonds within a plastic polymer of interest provides insights into the family of enzymes needed to degrade the plastic and may give insight into possible modifications needed for these enzymes. Four types of covalent bonds are abundantly present in plastics: ester bonds, urethane bonds and carbon-carbon bonds. In addition, the crystallinity of plastics presents another important determinant for their biodegradability.

Ester Bonds

Ester bonds are abundant in both petrochemical-derived plastics and bio-based polymers. Typical plastics containing these bonds are polyethylene terephthalate (PET), polycaprolactone (PCL), poly-lactic acid (PLA), and many more exist. Interestingly, for all of these plastics degrading enzymes have been identified. This likely correlates with the abundance of ester bonds in natural polymers such as lipid- and phenolic-based barriers present in the outer layer of plant cells. Enzymes able to degrade such natural polymers have in some cases been shown to be able to degrade synthetic compounds such as PET. A common example is presented by cutinase enzymes from *Thermobifida* spp., that degrade plant polymeric cutin and in addition have been shown to degrade PET (Ribitsch et al., 2012). Several common lipases and esterases have also been shown to be able to degrade one or more of the above-mentioned polyesters (Hajighasemi et al., 2016).

Urethane Bonds

Urethane bonds constitute polyurethane plastics which are mostly used for long-term applications in construction and automotive, as well as in foams used e.g., in furniture (PlasticEurope-Association of Plastics Manufacturers, 2020). Urethane bonds resemble covalent peptide bonds between amino acids constituting proteins. Electron density and pulling force within these bonds are clearly different from those in ester bonds and hence, different enzymes are required for degradation (Kjeldsen and Zubarev, 2011). In addition, several enzymes such as the polyester hydrolases were shown to degrade polyurethanes (Howard et al., 1999; Russell et al., 2011; Álvarez-Barragán et al., 2016; Schmidt et al., 2017; do Canto et al., 2019).

The majority of published research on polyurethane degradation is performed on Impranil-DLN (Biffinger et al., 2015; Schmidt et al., 2017; Danso et al., 2019; Espinosa et al., 2020). Impranil is an aqueous polyester-polyurethane emulsion that is mostly used for impregnating textile (Biffinger et al., 2015). This emulsion is easily incorporated in solid growth medium and accessible for bacteria or fungi allowing for a ready to use screening method. However, it is not easily translatable to the degradation of polyurethanes like PUR (polyurethane), since Impranil also contains ester bonds (Molitor et al., 2020).

Ether Bonds

Ether bonds are abundantly found in nature, they are present in wood cell walls (Nishimura et al., 2018), plant active molecules such as saponins (Abed and Aziz, 2019) and many more (Domínguez De María et al., 2010). Some plastics contain combinations of bonds for example PUR contains both urethane bonds as well as ether bonds (**Figure 1**). These foams are highly recalcitrant and difficult to degrade. So far it has been shown that a selected microbial community was able to partially degrade these polyether-polyurethane co-polymers when dispersed in water (Gaytán et al., 2020). Recently, the first case of enzymatic poly(ether urethane) foam degradation has been shown using a laccase-mediated system (Magnin et al., 2021). Butylene succinate copolymers like butylene succinate-co-diethylene glycol succinate (PBS-co-DEGS) and butylene succinate-co-butylene diglycolic acid (PBS-co-BDGA) are other examples of polymers with combined bonds since it consists of ester and ether bonds. The dispersed presence of these ether bonds in these specific co-polymers appears to improve the hydrophobicity of these plastics and hence, improves their biodegradability (Li et al., 2018).

Carbon-Carbon Bonds

Carbon-carbon bonds are most commonly found in petrochemical-derived plastics like polystyrene (PS), polyvinyl chloride (PVC), polyethylene (PE), and polypropylene (PP). These plastics, and especially PE, are commonly used in single-use products (packaging, plastic bags, etc.) and contribute extensively to environmental pollution (PlasticEurope-Association of Plastics Manufacturers, 2020). In contrast to urethane bonds and ester bonds, carbon-carbon bonds are much more recalcitrant, and their degradation usually requires redox enzymes. These enzymes cleave the carbon-carbon bond in four steps: (1) oxidation and rearrangement of unstable oxidized intermediates; (2) collapse and rearrangement of radicals and/or cations; (3) oxygenation to obtain hydrolyzed cleavage products

FIGURE 1 | Plastics and their specific chemical bonds.
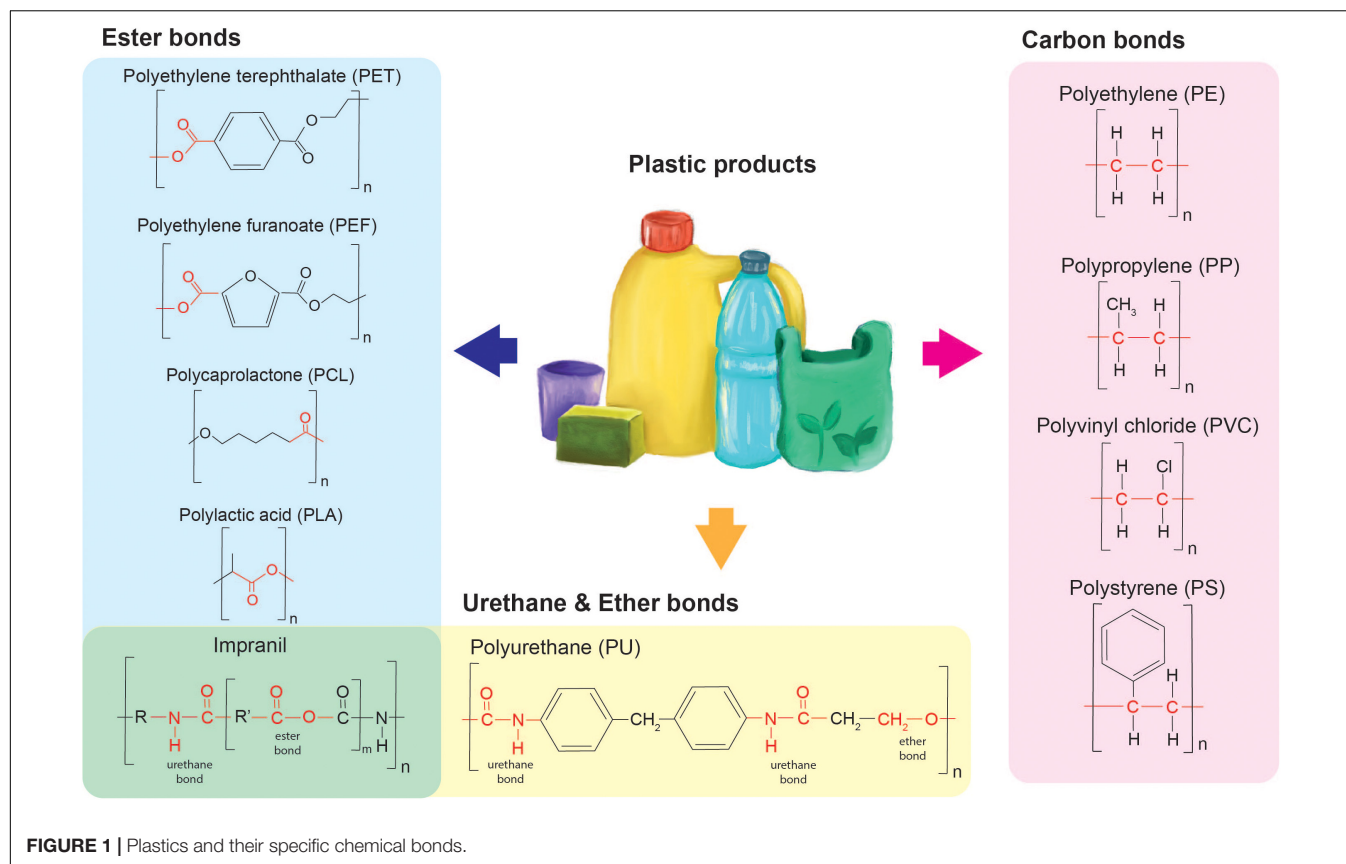
and finally; (4) oxygen activation requiring oxygen for substrate binding. The complex redox enzymes catalyzing these steps require metal ligands (Guengerich and Yoshimoto, 2018).

In contrast to ester bonds, only a few specific enzymes have been identified for the degradation of carbon-carbon bond polymers. Recently, initial steps were taken with the isolation of *P. putida* from the gut of polystyrene-eating super worms (Yang et al., 2020). Kim et al. (2020) isolated *Pseudomonas* sp. DSM50071 from the super worm gut and identified the enzyme responsible for polystyrene degradation. The enzyme was characterized as a serine-hydrolase oxidizing the polystyrene and thereby increasing its hydrophilicity and enabling further degradation of the polymer. For polyethylene (PE), polypropylene (PP), and polyvinyl chloride (PVC) microbial degradation has been observed. For example, PE degrading fungi have been identified using a combination of solid screening on medium containing plastic particles in combination with respirometry assays, Fourier Transform infrared (FTIR) and scanning electron microscopy (Spina et al., 2021). Several Fusarium species, *Aspergillus fructus* and *Purpureocillum lilacinum* were able to damage PE films, however, no specific enzymes were identified (Spina et al., 2021).

Other Structural Determinants

Apart from the chemical bonds connecting the monomers, other physicochemical characteristics may affect the biodegradability of plastics. One important factor is crystallinity which affects the strength of plastic polymers as well as their degradation rate (Pantani and Sorrentino, 2013). Low crystallinity results in better degradation. PET degrading *I. sakaiensis* can degrade amorphous PET but is unable to degrade high crystallinity PET (Wallace et al., 2020). When the crystalline PET was converted to amorphous PET the bacterium again was able to degrade the compound (Wallace et al., 2020). Studies on the degradation of PBS also suggested that crystallinity may be a more important factor influencing degradation than molecular weight (Pan et al., 2018). Moreover, these authors showed that crystallinity of PBS may decrease after several hours of degradation (Pan et al., 2018). Another study showed that the crosslinking of amorphous silk domains with crystalline silk regions protects the amorphous region, making it more resistant against enzymatic degradation (Valente et al., 2020), The degradation of plastics with high crystallinity may be improved by the expansion of the enzyme's binding pocket as shown in the case of PET (Liu et al., 2018; Kitadokoro et al., 2019) or by the pre-treatment of plastics making them less crystalline (Wallace et al., 2020).

The addition of specific biosurfactants may significantly increase or decrease the biodegradability of various plastics. Recently, it was shown that the addition of rhamnolipids increased the microbial colonization of plastics (Taghavi et al., 2021). Whether degradation was indeed increased, appeared dependent on the type of plastic. For PET and PE biodegradation decreased, whereas for PS degradation increased (Taghavi et al., 2021). This is likely to be a rhamnolipid specific feature since

TABLE 1 | Comprehensive overview of enzymes discussed in this review.

Plastic	Chemical bond	Domain and phylum	Species	Method of identification	Enzyme	Enzyme class	References
PET	Ester	Bacteria—Proteobacteria	*Ideonella sakaiensis*	Screening on low-crystallinity PET films	*Is*PETase	Tannase	Yoshida et al., 2016; Liu et al., 2018
PET	Ester	Bacteria—Actinomycetota	*Thermobifida fusca*	Enzyme activity assay on 3PET	*Tf*Cut2	Cutinase	Herrero Acero et al., 2011; Ribitsch et al., 2012, 2015
PET/PEF	Ester	Bacteria—Actinomycetota	*Thermobifida cellulosilytica*	Enzyme activity assay on 3PET Enzymatic assay on PEF powder	*Tc*Cut1	Cutinase	Herrero Acero et al., 2011; Pellis et al., 2016
PET	Ester	Bacteria—Actinomycetota	*Thermobifida alba* AHK119	Screening on 3PET	Est119/*Tha*-Cut1	Cutinase/Esterase	Kitadokoro et al., 2012; Ribitsch et al., 2012
PLA	Ester	Bacteria—Proteobacteria	*Alcanivorax borkumensis*	Screening on various polymer emulsions	*ABO*2449	α/β hydrolase	Hajighasemi et al., 2016
PLA	Ester	Bacteria—Proteobacteria	*Rhodopseudomonas palustris.*	Screening on various polymer emulsions	*RPA*1511	α/β hydrolase	Hajighasemi et al., 2016
PCL	Ester	Bacteria—Proteobacteria	*Alcanivorax borkumensis*	Screening on various polymer emulsions	*ABO*2449	α/β hydrolase	Hajighasemi et al., 2016
PCL	Ester	Bacteria—Proteobacteria	*Rhodopseudomonas palustris.*	Screening on various polymer emulsions	*RPA*1511	α/β hydrolase	Hajighasemi et al., 2016
PCL	Ester	Bacteria—Actinomycetota	*Streptomyces* sp.	*In silico* search, screening on emulsified PCL	*SM*14est	Esterase	Almeida et al., 2019
PS	Carbon-Carbon	Bacteria—Proteobacteria	*Pseudomonas* sp. DSM50071	Grown on PS film	No specific enzyme	Serine hydrolase	Kim et al., 2020
LD-PE	Carbon-carbon	Bacteria—Proteobacteria	*Pseudomonas putida*	Screening on LD-PE films	No specific enzyme	Unknown	Talkad et al., 2014
LD-PE	Carbon-carbon	Fungi—Ascomycota	*Fusarium oxysporum, Fusarium falciforme and Purpureocillum lilacinum*	Initial screening on agar containing pulverized LD-PE	No specific enzyme	Unknown	Spina et al., 2021
LD-PE	Carbon-carbon	Bacteria	Marine organisms	LD-PE powder in medium	No specific enzyme	Unknown	Khandare et al., 2021

another study indicated increased biodegradability of LD-PE upon the addition of biosurfactants (Hussain Ali et al., 2021).

THE SEARCH FOR NEW PLASTICS DEGRADING MICROBES AND ENZYMES

Enzymatic recycling of plastics presents a sustainable and therefore interesting approach for several reasons. These include (*i*) no harsh conditions are required for degradation, (*ii*) costs for bulk enzyme production can be relatively low, and (*iii*) it allows for the retrieval of the monomers which can be polymerized into new plastics or upcycled into new compounds. To develop the enzymatic degradation approach, novel plastics degrading enzymes need to be discovered, characterized and improved. These enzymes can be categorized into two types: (1) plastic depolymerizing enzymes and (2) downstream processing enzymes, ensuring proper processing of the degradation products toward the desired product. For example, the PETase from *Ideonella sakaiensis* (*Is*PETase) can degrade PET into mono-(2-hydroxyethyl) terephthalic acid (MHET) and this enzyme

would be categorized as a PET depolymerizing enzyme. The MHETase that degrades MHET into terephthalic acid and ethylene glycol is regarded as a downstream processing enzyme within the scope of this review (Yoshida et al., 2016). These enzymes were initially identified by Yoshida et al. (2016) and subsequently discussed in a broad spectrum of papers discussing the production, structure and mechanisms of both the PETase and the MHETase (Han et al., 2017; Liu et al., 2018; Ma et al., 2018; Almeida et al., 2019; Palm et al., 2019; Seo et al., 2019; Janatunaim and Fibriani, 2020; Wallace et al., 2020).

The search for novel complex polymer degrading enzymes started decades ago and was extended with plastics depolymerization enzymes in the late 1980s. In, Pometto et al. (1992) showed that *Streptomyces* species can produce extracellular polyethylene degrading enzymes. These enzymes can be optimized to reach a higher efficiency. In addition, *in silico* methods can be used to identify interesting homologs for further investigation. Several strategies to discover new plastic degrading enzymes are discussed in detail below. The enzymes discussed in this paragraph were summarized in **Table 1**.

Experimental Approaches—Environmental Screening for Plastic Degrading Enzymes

Isolating micro-organisms from environments with long-term exposure to rigid natural polymers or plastics has been a fruitful approach and led to the identification of several microbial species able to degrade plastics. This suggests that microbes may adapt toward the use of anthropogenic persistent and complex polymers.

Several bacteria displaying plastic degrading abilities have been isolated from floating debris, garbage soil, landfills, and polluted water. These include *Pseudomonas, Bacillus, Staphylococcus, Streptomyces,* and many more species (El-Shafei et al., 1998; Usha et al., 2011; Bhardwaj et al., 2013; Brunner et al., 2018). Here we describe a few interesting bacterial origins of enzymes. An example is presented by the newly identified bacterium, *Ideonella sakaiensis*, isolated from a PET bottle recycling facility, which uses PET as an energy source (Yoshida et al., 2016). The enzymes required for the depolymerization, and downstream degradation of PET were characterized and shown to act specifically and efficiently (Yoshida et al., 2016; Liu et al., 2018). Another example is presented by the Actinobacteria *Thermobifida* cutinases; these members of the esterase family can hydrolyze the primary and/or secondary ester linkage in cutin, allowing microbes to use the monomers as a carbon source. Cutinases hydrolyze ester bonds and hence, are interesting candidates to screen for the ability to degrade plastics. Since 2010 several studies indicated that cutinases, especially from *Thermobifida* species, can degrade synthetic polymers like PET (Herrero Acero et al., 2011; Ribitsch et al., 2012, 2015). The Cut1 enzyme from *Thermobifida cellulosilytica* has shown to successfully degrade PET and polyethylene furanoate powders (PEF) (Herrero Acero et al., 2011; Pellis et al., 2016).

Another interesting group of microorganisms are the fungi, they are known for their ability to degrade complex polymers such as wood, cellulose and lignin (Bugg et al., 2011; Couturier et al., 2018; Nemli et al., 2018; Zhang et al., 2016; Hou et al., 2020; Qu et al., 2021). Several fungal cutinases may also degrade PET (Suzuki et al., 2014; Sankhla et al., 2020; Anbalagan et al., 2021; Vázquez-Alcántara et al., 2021). In addition, Filamentous fungi isolated from landfill soil have been shown to degrade polyethylene. These include *Fusarium oxysporum, Fusarium falciforme,* and *Purpureocillum lilacinum* (Spina et al., 2021). Various reviews have been published providing an overview of these plastic degrading microorganisms (Pathak and Navneet, 2017; Roohi et al., 2017; Rana, 2019; Jaiswal et al., 2020; Mohanan et al., 2020; Maity et al., 2021; Priya et al., 2022). Due to the abundance of plastic degrading micro-organisms isolated from plastic polluted environments, plastic debris is considered a promising source for the isolation of potentially successful plastic degrading bacteria and fungi.

Computational Approaches

With interesting strains or consortia identified, different approaches can be taken to pinpoint the genes encoding plastic degrading enzymes. The first approach has led to the identification of the *Is*PETase from *I. sakaiensis* in which the sequenced genome was investigated for promising open reading frames containing sequences like already known plastic degrading enzymes (Yoshida et al., 2016). The second approach includes transcriptome analysis and identification of up-and down-regulated genes in the presence and absence of plastics to reveal genes potentially important for PET degradation (Kumari et al., 2021).

Data Mining and *in silico* Screening for Plastic Degrading Enzymes

By performing homology searches using various databases promising enzymes can be revealed. An example is presented by the PETase-like enzyme SM14est found by searching homologs of the *Is*PETase. 30 potential enzymes were found of which one showed PCL degrading activity (Almeida et al., 2019). Surprisingly, to date, only a few papers have been published using this method to identify promising enzymes. Nevertheless, this is a promising technique to narrow down the number of bacteria and fungi for screening. It is, however, important to note that merely looking for homology at the amino acid level is not enough to find interesting enzymes. In addition, understanding and predicting the chemical characteristics of the enzymes and predicting enzyme structures are of utmost importance for the identification of promising new enzymes using *in silico* methods.

Another *in silico* source for enzyme identification is ancestral sequence reconstruction (ASR) By obtaining the possible ancestral sequence, novel enzymes can be identified via homology searches or novel proteins can be expressed by the creation of fusion proteins (Verma et al., 2019; Zitare et al., 2021). In addition, ASR provides insights into the evolutionary development of enzymes (Ruiz-Dueñas et al., 2020). So far, ASR and additional methods have been used to identify novel enzyme families and conserved structures (Voshol et al., 2017). Likewise, it can provide insight into the origin and evolution of plastics degrading enzymes. *In silico* methods indeed, have shown to provide leads aiding the search for novel enzymes and are, in our opinion, useful to gain more insights into promising enzymes.

De novo Enzymes Design

De novo design and development of enzymes, based on well-determined modular structures provides a possible future approach to generate a new generation of plastic degrading properties. The *de novo* design and creation of enzymes is one of the primary goals of synthetic biology and requires computational approaches. Donnelly et al. created *de novo* enzyme Syn-F4 which can hydrolyze the siderophore ferric enterobactin. The expression of Syn-F4 *in vivo* allowed a ferric enterobactin sensitive *E. coli* to survive in the presence of ferric enterobactin (Donnelly et al., 2018). New tools for the *de novo* design of enzymes are emerging rapidly. By using machine learning methods, protein prediction software is being improved drastically. The current prediction programs, however, are still lacking a user-friendly platform for end-users, resulting in limited use of these tools (Abriata and Dal Peraro, 2021). To stimulate the design of novel enzymes using computational synthetic biology approaches is dependent on the prior identification and

characterization of several different plastics degrading enzymes to begin training machine learning software and providing accessible computing tools. In addition, a list of web tools for *in silico* biodegradation studies is available in a recent review (Skariyachan et al., 2022).

Another useful method is the addition of specific protein domains to create better enzymes with altered binding properties. This method has been used e.g., for the optimization of carbohydrate degrading enzymes, where carbohydrate-binding modules (CBMs) have been switched and replaced to optimize carbohydrate degrading abilities as well as thermostability and catalytic efficiency in existing enzymes (Ha et al., 2015; Meng et al., 2015). Likewise, CBMs may be useful tools to increase substrate binding of plastic degrading enzymes, since a recent paper showed that the addition of inactive polysaccharide monooxygenase *Pc*AA14A to the medium increased the efficiency of the *Is*PETase by 27 percent. This inactivated enzyme appears to bind with its CBM to the hydrophobic surface of PET, thus enabling the other enzyme to bind more efficiently (Dai et al., 2021a). Subsequent research by Dai et al. (2021b) shows that the addition of a CBM domain can increase the plastic degrading abilities of the *Is*PETase up to 86 percent, underscoring that enhancing these enzymes with extra binding domains is a promising approach for improved degradation.

ENZYME AND STRAIN OPTIMIZATION FOR PLASTIC DEGRADATION

Existing enzymes can be optimized using various methods to obtain a higher depolymerization efficiency and thermostability. These may be regarded as necessary requirements since plastic depolymerization is currently mostly conducted at elevated temperatures (Nakasaki et al., 2019; Xiong et al., 2020). In addition, *in silico* methods combined with experimental approaches can target specific sites for the catalytic improvement of these enzymes.

Random Mutagenesis

Random mutagenesis and subsequent screening and selection is an effective strategy used to improve the activity of specific strains and enzymes. By exposing plastic degrading microbial strains to mutagens, such as UV, a high mutation frequency can be obtained and mutants with higher enzyme activity can be selected. If required, cycles of mutagenesis and selection can be repeated until a significantly higher efficiency is reached. Talkad et al. (2014) were able to improve the degradation of low-density (low crystallinity) polyethylene by UV and EMS treatment of *P. putida*. Via GWAS studies the SNPs responsible for the improvements can be identified. The same approach has shown to be fruitful in fungi. *Penicillium oxalicum* strain DSYD05-1 was optimized vis UV mutagenesis to increase its PCL degrading ability. The enzymes of this strain showed to have a higher PCL- degrading ability and a wider substrate specify allowing it to additionally degrade poly(β-hydroxybutyrate) (PHB) and poly(butylene succinate) (PBS) (Li et al., 2012).

TABLE 2 | Overview of methods for screening and identification of plastic degradation.

Plastic	Model polymer	Method of incorporation	References
PET	PET	Emulsification	Charnock, 2021
PET	3PET, 2PET, BHET	Dissolving	Herrero Acero et al., 2011; Ma et al., 2018; Palm et al., 2019
PLA	PLA	Emulsification	Pranamuda et al., 1997
PLA	PLA	Film on plate	Kim et al., 2017
PLA	PLA	Spray	Shin et al., 2021
PCL	PCL	Emulsification	Almeida et al., 2019
PUR	Impranil-DLN	Dissolving	Molitor et al., 2020
HD-PE	HD-PE	Films	Awasthi et al., 2017
LD-PE	LD-PE powder	Mixing with medium	Khandare et al., 2021; Spina et al., 2021
PS	PS	Emulsification	Tang et al., 2017
PS	PS	Spray	Shin et al., 2021

Site-Specific Mutagenesis Based on Enzyme Modeling

As a recent development, computer-aided enzyme engineering is increasingly used for the optimization of enzymes. When the crystal structure of an enzyme is known, computer-aided engineering can be used to identify important targets for improvement. Experimental approaches can then induce these changes resulting in optimized enzymes.

For example, based on the crystal structure and modeling of the *Is*PETase, site- directed mutagenesis has been performed on 15 amino acid domains in the first contact shell of the *Is*PETase, to improve its efficiency (Tournier et al., 2020). The addition of disulfide bridges to improve thermostability and mutagenesis of the residues responsible for substrate binding also drastically improved PET-depolymerization, allowing the enzyme to depolymerize 90% of the provided PET in 10 h. The optimized strain was able to degrade 16.7 g of PET per liter per hour (Tournier et al., 2020). The 90% efficiency in 10 h of the enzymatic degradation process comes close to the 98% efficiency in 8 h for chemical PET degradation (Khoonkari et al., 2015). Furthermore, the purified monomers could be used to synthesize new PET without the loss of quality of the material (Tournier et al., 2020).

In addition, Liu et al. (2018) enzyme modeling and experimental data provided insight into the different binding pockets of the PETase from *I. sakaiensis* and of a spectrum of cutinases. Likewise, Kitadokoro et al. (2019) identified the binding pocket residues involved in substrate interaction of a cutinase isolated from *Thermobifida* and how this differs from the *Is*PETase. After comparison with a.o., *Thf*42_Cut1 it was concluded that the success of the *Is*PETase enzyme is caused by the structural features of its binding pocket. Compared to other enzymes, this PETase has a relatively shallow and broad surface, allowing the enzyme to bind to aggregated PET molecules whereas other cutinases are mostly only able to hydrolyze linear PET molecules (Liu et al., 2018).

A similar approach was taken with the *Tf* Cut2 from *Thermobifida fusca*. Furukawa et al. (2019) managed to improve the PET degrading abilities of *Tf* Cut2 by identifying important residues via computer modeling and applying site-directed mutagenesis on the identified substrates thereby increasing the degradation rate of the PET film 12.7-fold.

In another study, the *Thc*_Cut2 cutinase was shown to be significantly less efficient than its close relative *Thc*_Cut1. Previously, Herrero Acero showed that differences in surface properties of these enzymes are responsible for the difference in efficiency. In one of the first site-specific mutagenesis studies on *Thc*_Cut2, Arg29 was mutated into Asn and Ala30 into Val. These changes resulted in significantly higher specific activity on a PET surface (Herrero Acero et al., 2013; Ribitsch et al., 2015). In the future, site-directed mutagenesis combined with the addition of substrate-binding domains (3.3) is promising to further improve the efficiency of plastics degrading.

Adaptive Laboratory Evolution to Improve the Activity of Plastic Degrading Strains

Adaptive laboratory evolution (ALE) is a powerful approach to improve or create certain phenotypes in microbial strains by provoking and stimulating evolutionary adaptation processes (Lee and Kim, 2020). In combination with omics-approaches to characterize the mutations invoked, ALE is a sophisticated and potent strain engineering tool for inducing mutations to improve metabolic pathways and enzymes for rapid growth on a variety of carbon sources and stress tolerances. Various examples of ALE for improved utilization of plastics monomers have emerged, which are important to build plastic-degrading or -upcycling cell factories.

Genome sequencing of *Pseudomonas pseudoalcaligenes* CECT 5344 revealed that this strain could potentially utilize furfurals (monomers of the biobased plastic polyethylene 2,5-furandicarboxylate or PEF), furfuryl alcohol, furfural and furoic acid as carbon sources. However, a long lag period lasting for several days was observed during its growth on furfurals. Growth on furfurals was improved using ALE, following which the adapted strain no longer exhibited any prolonged lag phases (Igeño et al., 2019). Sequencing of this strain revealed a point mutation in an AraC family activator locus (BN5_2307) at the HTH region of the protein (L261R) to be responsible for this improvement by generating an active regulator for the *hmfABCDE* gene cluster located upstream of this gene.

In another example, Li et al. (2019) successfully isolated ALE-derived mutants of *P. putida* KT2440 able to utilize ethylene glycol, a monomeric component of PET besides terephthalate, as its carbon source. Genomic analysis of these mutants revealed missense mutations and a 15 bp deletion on PP_4283 encoding for *gclR*, a transcriptional repressor to the glyoxylate carboligase pathway. In addition, secondary mutations were also found in a transcriptional regulator encoded by PP_2046 and a porin encoded by PP_2662 which further improved the growth of ALE-derived *P. putida* KT2440 on ethylene glycol. These secondary mutations likely maintain flux balances through the initial oxidation of ethylene glycol to glyoxylate.

Although current ALE approaches are mainly focused on the efficient utilization of plastics monomers, ALE has a great potential to evolve and improve plastics depolymerization enzymes. Several members of promiscuous enzyme classes, like esterase, lipase and cutinase, have been identified to be able to depolymerize PET and PLA, although with low specificity and turn-over (Kawai et al., 2019). ALE or directed evolution constitutes a promising approach to acquire new or improved enzymatic activity from promiscuous enzyme classes to develop novel plastic depolymerizing enzymes.

THE CHALLENGES OF EXPERIMENTAL DESIGN AND SET-UP

The diversity of methods available for the identification of micro-organisms and enzymes with plastic depolymerizing abilities complicates the trustworthiness and translatability of the experiments. In any case, the experimental setup is highly determinant to increase the chances of success.

Clearance assays are frequently used in screening for enzyme activities. These rely on a turbid plastic-containing medium. If the plastic in the medium is degraded, a clear halo will occur. Screening for halo formation is a relatively quick method that yields binomial results. Two approaches are frequently used to incorporate plastics in the medium which include emulsification and the addition of semi-water dissolvable plastics. During emulsification, plastics are dissolved in an organic solvent, such as dichloromethane, and mixed with surfactants and growth medium. Subsequently, the solvent is evaporated, resulting in small droplets of plastic in the medium (Pranamuda et al., 1997). This method is used for plastics like polylactic acid, polycaprolactone, and polystyrene. We experienced that emulsification is not an easy-going method, difficult to reproduce and a challenge in keeping the final dispersion stable. Another approach is the addition of plastic granules or pellets to the medium and screen for clearance (Khandare et al., 2021; Spina et al., 2021).

An easier approach is the addition of plastics that are partially dissolvable or plastic simulators such as the aquatic dispersion Impranil-DLN, which is used for the screening of polyester and polyurethane degradation (Molitor et al., 2020). Similarly, bis-(2-hydroxyethyl) terephthalate (BHET) is soluble in water and can be used as a PET-mimicking substrate (Palm et al., 2019). Model compounds are very useful to show e.g., polyester degrading activity but are only partially translatable to the degradation of the actual polymer of interest (**Table 2**; Ribitsch et al., 2011; Almeida et al., 2019; Palm et al., 2019; Molitor et al., 2020). Therefore, one must be aware that results are easily over-interpreted and plastic degrading features may be overstated. However, plate clearance assays are quick, high throughput and important screens for the initial identification of interesting organisms.

Measuring the weight loss of plastics is another method for the identification of plastic degradation. Plastic particles or films are incubated in the presence of microorganisms and regularly

weighed to monitor the possible decline in weight. Major drawbacks of this method are timespan and weight fluctuations (Li et al., 2008; Danso et al., 2019; Przemieniecki et al., 2020). It may take weeks to months to conduct the experiments and weight differences may be caused due to fragmentation by other external or internal factors such as the attachment of microbes to the plastic which is also measured. A better method to examine plastic degradation is provided via respirometry assays, specifically while the plastic is being used as sole carbon. This method is based on measuring carbon dioxide production upon degradation of the polymers, using pH indicators and carbon dioxide calibration (Yang et al., 2020; Spina et al., 2021). This method provides more reliable results than weight loss assays since it is less influenced by other external factors such as the fragmentation of the plastics. Plastics degradation may be examined by combining respirometry assays with microscopy, testing the plastic's physical characteristics and enzymatic activity assays (Janczak et al., 2018; Tournier et al., 2020; Yang et al., 2020; Spina et al., 2021). Interactions of the microorganism with the plastic surface can be visualized using scanning electron microscopy, this does not show enzymatic activity but can provide clues about the physical interaction of the microorganism of interest and the plastics. In addition, Liquid Chromatography-Mass Spectrometry (LC-MS) is a powerful tool for the detection of polymer degradation products and characterization of the degradation pattern (Hajighasemi et al., 2016). An excellent example is the research of Hajighasemi et al. (2016), in this study purified enzymes were used to degrade PLA resulting in a clear chemical degradation pattern.

Another method to observe chemical degradation is FTIR spectroscopy, based on the (changing) absorbance or transmission of specific wavelengths by specific (altered) chemical groups (Sandt et al., 2021). Several studies used FTIR to evaluate microbial PE, PET and polystyrene degradation (Ioakeimidis et al., 2016; Ojha et al., 2017; Canopoli et al., 2020; Yang et al., 2020; Spina et al., 2021). A recently published short communication by Sandt et al. (2021) discusses the use and misuse of FTIR in relation to plastic degradation. Importantly, most studies reported to date are focused on strains displaying plastic depolymerizing abilities instead of the specific enzyme responsible. If enzymes are identified, efficiency is not easily comparable to other enzymes. Herrero Acero et al. (2011) performed several enzymatic assays, PET hydrolysis tests, modeling and docking analyses to be able to compare the Thf42_Cut1 enzyme from *Thermobifida fusca* with two other, previously identified, PET degrading cutinases: Thc_Cut1 and Thc_Cut2 derived from *Thermobifida cellulosilytica* DSM44535. By expressing and purifying these three enzymes in *E. coli* the kinetic properties on PET could be compared. They were able to model the hydrolysis of PET and determine that Thc_Cut1 was most efficient in PET degradation (Herrero Acero et al., 2011). Other methods may include examining enzyme activity on model substrates or actual polymers, as well as examining degradation using chemical analysis or microscopy (Hajighasemi et al., 2016; Austin et al., 2018; Magnin et al., 2019). An interesting approach is the use of dye-containing plastic films, whereupon enzymatic degradation of the films the dye is being released, causing the

appearance of a blue color (Shinozaki et al., 2013). For enzymatic assays to work optimally the suitable conditions for enzyme activity should be determined (Jain et al., 2020).

As discussed by Arnling Bååth et al. (2020), the kinetics of these plastic degrading enzymes is not well understood and there is no framework present to analyze these reactions and therefore score the efficiency of plastic degrading enzymes. There is a need for better substrates to quantify the plastic degrading enzyme activity. Arnling Bååth et al. (2020) showed that the addition of putative attack sites on the surface of PET allowed better comparison of the activity of various enzymes. A faster method is the use of spectrophotometric absorbance to analyze the enzymatic degradation kinetics. With bulk absorbance assays insight may be obtained into the enzyme reaction kinetics and efficiency (Zhong-Johnson et al., 2021). If this method could be used as a standardized method for the comparison of plastic degrading enzymes, such comparison would be simplified drastically.

TOWARD A COMPREHENSIVE WORKFLOW FOR PLASTIC DEGRADING ENZYME IDENTIFICATION

Not only is it important to follow the currently used methods, but it is also important to establish new strategies for finding and optimizing plastic degrading species and enzymes. This will allow to fill current gaps in the identification of plastic degrading enzymes and create efficient pipelines for the identification of such novel enzymes.

The diversity of methods complicates the comparison between results. Therefore, we would like to propose a workflow that would make outcomes more comparable and would result in more experimental support for future experiments (**Figure 2**).

As a quick and reliable method for screening organisms for plastic degrading activities, we suggest using the clearance assays since this method is also suitable for high throughput screening. To screen efficiently, optimal screening conditions need to be determined and possible inducers for enzyme expression should be identified. These screens can be performed using model polymers, or by emulsification, incorporation into the medium or spraying of the original polymers (**Table 2**).

The most favorable strains can be grown in liquid culture containing the above mentioned model plastics, or the actual plastic polymer itself in powder, film, particle or even cube form (Chua et al., 2013; Sepperumal et al., 2013; Pellis et al., 2016; Weinberger et al., 2017; Magnin et al., 2021). By exposing microorganisms to plastic particles in liquid culture, the degradation can be examined providing more insights into the natural degradation processes. Samples from these flasks can be used for several approaches; (1) microscopical analysis, the physical interaction between de microorganism and the plastic can be examined by SEM by making high-resolution images of the interactions between the organism of interest and plastic. A drawback is that only the macrostructure of the film or particle can be observed since it is only a visual technique. (2) Chemical analysis, for example, LC-MS-based analyses to observe

FIGURE 2 | Comprehensive workflow for the identification of plastic degrading enzymes. New plastic degrading enzymes can be found by **(A)** environmental screening of interesting strains (green arrows) and **(B)** the computational methods (magenta arrows). In the case of computational methods, the enzymes need to be expressed in a suitable expression system (2) before high throughput enzyme assays can be performed. Via high throughput enzymatic assays, active strains/enzymes can be pinpointed. For the strains isolated via environmental screening, the specific enzymes need to be identified (2). The enzymes can then be further characterized to identify potential targets for enzyme optimization approaches. The protein depicted in this figure is based on PDB ID 6ANE (Fecker et al., 2018).

and identify the non-metabolized products. In addition, FTIR can detect and indicate the oxidation of the plastics and the creation of new chemical bonds during this process (Sandt et al., 2021). Identification of complete chemical degradation can prove difficult when the degradation products are being metabolized.

This can be visualized using respirometry monitoring carbon dioxide release and provide proof for the use of plastic as a carbon source (Yang et al., 2020; Spina et al., 2021). Another, however, more sophisticated option would be the use of a cell-free system to observe and investigate actual degradation by specific enzymes.

(3) Genomics, transcriptomics and/or proteomics can be used to identify the responsible enzyme(s) and optimized using ALE or site-directed mutagenesis approaches. Parts of this workflow can be used depending on the starting point of the study (**Figure 2**).

THE FUTURE OF PLASTIC DEPOLYMERIZATION—BIORECYCLING AND BIO-UPCYCLING

When plastic degrading strains or enzymes are identified, characterized, and improved, the next step would be to apply the strains or enzyme (cocktails) for the degradation of plastics. We identified two main fields for the application of these strains or enzymes namely: (1) biorecycling and bio-upcycling, (2) bulk degradation and bioremediation. Depending on the application a thoughtful future approach is suggested.

Biorecycling and bio-upcycling are promising solutions to replace the current mechanical recycling methods since less product and quality is lost. Purified enzymes or enzyme cocktails can be used to catalyze the depolymerization process. Some advantages of a cell-free system are the mild reaction conditions and degradation under conditions unsuitable for bacterial or fungal cultures. The major challenge to overcome, still, is the handling of combinations of different plastics since no enzyme is expected to be active on all plastics. Large scale degradation of plastic mixtures can be achieved using enzyme cocktails. Such mixtures can be used to degrade mixes of different polymers or one specific complex polymer consisting of different chemical bonds. Enzyme cocktails are already used in industry for the degradation of natural complex compounds, such as lignocellulose (Lopes et al., 2018). Currently, projects are starting up researching the use of enzyme cocktails for plastic degradation (Ballerstedt et al., 2021).

Another promising method is presented by the industrial degradation of plastics by microbial consortia (Skariyachan et al., 2021). Microbial consortia have shown to degrade PE, PP, and other recalcitrant polymers (Skariyachan et al., 2017; Kang et al., 2020; Blair et al., 2021). These can be carefully structured consortia (Negi et al., 2009; Kang et al., 2020) or specifically isolated consortia from waste (Skariyachan et al., 2016, 2018). No economic models have so far been published about the feasibility of such approaches on an industrial scale. However, if plastic can be used as a sole carbon source for these communities, the use of microbial consortia for plastic degradation would be a cost-efficient approach once the upscaling is successful.

To achieve large scale degradation of these plastics using a cell-free system, the efficiency of the plastic degrading enzymes must be increased, and an efficient waste retrieval system must be established. PET would be an ideal candidate plastic for the establishment of a large-scale enzymatic recycling trial. Many countries already have an established PET retrieval system in place meaning enough PET waste can be supplied. The availability of relatively pure PET waste and PET degrading enzymes to convert PET into its monomers allows efficient recycling. These monomers can then be purified and repolymerized into new plastics (Tournier et al., 2020).

The monomers can also be upcycled toward more valuable compounds. As discussed in the bowtie model by Tiso et al. (2021), bacterial metabolism can be adapted to convert plastic imputes (monomers, oligomers) into a wide variety of compounds such as aromatics, organic alcohols and more. The promise of this has already been shown by the conversion of terephthalic acid to vanillin, coumarin and catechol (Kim et al., 2019; Sadler and Wallace, 2021). The review of Tiso et al. (2021) discusses the promise of engineered microorganisms for the processing of these plastics and the promise of plastic monomers as substitutes for current petrochemical-based materials extensively. The adjustments of these metabolic pathways were extensively described showing the possibilities to further engineer these pathways allowing for microbial upcycling of plastics (Tiso et al., 2021).

There are some challenges to overcome, especially concern plastic depolymerizing enzymes. The main demerits, concerning plastic degrading enzymes, will so far remain enzyme efficiency, thermostability and the degradation of highly crystalline plastics. These aspects need to be dramatically improved to be able to degrade, for example, high-crystallinity PET used for the manufacture of plastic bottles (Yoshida et al., 2016; Kawai et al., 2019; Kitadokoro et al., 2019). Hence, the identification and optimization of (novel) enzymes promise to yield enzymes efficient enough to degrade high crystallinity plastics. Additionally, thermostability of enzymes can be improved using enzyme modeling combined with site-specific mutagenesis (Liu et al., 2018; Tournier et al., 2020). Another option would be the optimization of the degradation conditions, to allow for more efficient degradation. This approach has been taken regarding the microbial production of plastics and could also be an important aspect for its degradation (Thapa et al., 2018; Sabapathy et al., 2019; Mohammed et al., 2020).

Another important aspect that must be considered is where these methods fit within the global development of a sustainable society and circular economy. Implementation of circularity entails that optimally no plastics would be incinerated or lost but all would be reused or recycled (Bucknall, 2020). So far, the circular economy has been implemented only partially since more time and money is needed to create the necessary infrastructure. The requirement to close the loops is highly dependent on the type of plastic; for each type, another approach would be needed (Eriksen et al., 2019). Additionally, several chemical plastic recycling methods have been shown not to be economically viable (Bucknall, 2020). An extensive comparison between all plastic recycling methods has recently been described by Lee and Liew (2021). Currently, a vast knowledge gap still exists considering the economic feasibility of enzymatic and/or microbial recycling in large scale (industrial) applications. Economical models must be established to further investigate the financial feasibility of these methods (Lee and Liew, 2021).

CONCLUDING REMARKS

Promising steps have been taken in the biorecycling, upcycling, and biodegradation of synthetic polymers. By focusing on the

degradation of specific chemical bonds instead of degradation of specific plastics, the search can be targeted to more all-round or generalist plastic degrading enzymes. We hypothesize that the abundance of specific bonds in natural polymers makes degradation of these bonds in synthetic polymers more likely. Further, several strategies to find plastic degrading organisms were mentioned to provide an overview of promising niches or environments to find possible plastic degrading organisms. The discovered strains and enzymes can be optimized using various techniques.

Experimental screening for plastic degradation can be complicated because plastics are intrinsically hard to degrade and difficult to disperse in the medium. Since experiments with plastics are not very straightforward, many different methods are used in literature. To compare studies with each other a standardized workflow was set up and described in this paper to provide guidelines toward the discovery and optimization of plastic degrading enzymes. This workflow can be used to find and optimize plastic degrading

enzymes for biorecycling, bio-upcycling, bulk degradation, or bioremediation.

The utilization of enzymatic cocktails can be an attractive solution considering plastic mixtures, either mixed in the waste stream or products containing several plastics. Therefore, future research should be directed toward the optimization of enzymatic cocktails for plastic degradation. We are convinced that using the microbial diversity, enzyme optimization approaches and optimization methods to produce these enzymes cost-effectively will provide a sustainable method for enzymatic plastic degradation and recycling.

AUTHOR CONTRIBUTIONS

J-AV wrote the majority of the main text of the manuscript. HK contributed to the writing of the manuscript and designed the figures. J-AV, HK, AR, and JW did the conceptualization, reviewing, and editing. All authors contributed to the article and approved the submitted version.

REFERENCES

Abed, M., and Aziz, E. (2019). A Review on Saponins from Medicinal Plants: chemistry, Isolation, and Determination. *J. Nanomed. Res.* 8, 282–288. doi: 10.15406/jnmr.2019.08.00199

Abriata, L. A., and Dal Peraro, M. (2021). State-of-the-art web services for de novo protein structure prediction. *Brief. Bioinform.* 22:bbaa139. doi: 10.1093/bib/bbaa139

Almeida, E. L., Rincón, A. F. C., Jackson, S. A., and Dobson, A. D. W. (2019). In silico Screening and Heterologous Expression of a Polyethylene Terephthalate Hydrolase (PETase)-Like Enzyme (SM14est) With Polycaprolactone (PCL)-Degrading Activity, From the Marine Sponge-Derived Strain Streptomyces sp. SM14. *Front. Microbiol.* 10:2187. doi: 10.3389/fmicb.2019.02187

Álvarez-Barragán, J., Domínguez-Malfavón, L., Vargas-Suárez, M., González-Hernández, R., Aguilar-Osorio, G., and Loza-Tavera, H. (2016). Biodegradative activities of selected environmental fungi on a polyester polyurethane varnish and polyether polyurethane foams. *Appl. Environ. Microbiol.* 82, 5225–5235. doi: 10.1128/AEM.01344-16

Anbalagan, S., Raghava, H., Venkatakrishnan, R., Ravindran, J., Sathyamoorthy, J., Rangabashyam, K. A., et al. (2021). Hydrolytic Degradation of Polyethylene Terephthalate by Cutinase Enzyme Derived from Fungal Biomass–Molecular Characterization. *Biointerface Res. Appl. Chem.* 12, 653–667. doi: 10.33263/briac121.653667

Arnling Bååth, J., Borch, K., Jensen, K., Brask, J., and Westh, P. (2020). Comparative biochemistry of four polyester (PET) hydrolases. *ChemBioChem* 22, 1627–1637. doi: 10.1002/cbic.202000793

Austin, H. P., Allen, M. D., Donohoe, B. S., Rorrer, N. A., Kearns, F. L., Silveira, R. L., et al. (2018). Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proc. Natl. Acad. Sci. U. S. A.* 115, E4350–E4357. doi: 10.1073/pnas.1718804115

Awasthi, S., Srivastava, P., Singh, P., Tiwary, D., and Mishra, P. K. (2017). Biodegradation of thermally treated high-density polyethylene (HDPE) by Klebsiella pneumoniae CH001. *3 Biotech* 7:332. doi: 10.1007/s13205-017-0959-3

Ballerstedt, H., Tiso, T., Wierckx, N., Wei, R., Averous, L., Bornscheuer, U., et al. (2021). MIXed plastics biodegradation and UPcycling using microbial communities: EU Horizon 2020 project MIX-UP started January 2020. *Environ. Sci. Eur.* 33:99. doi: 10.1186/s12302-021-00536-5

Bhardwaj, H., Gupta, R., and Tiwari, A. (2013). Communities of Microbial Enzymes Associated with Biodegradation of Plastics. *J. Polym. Environ.* 21, 575–579. doi: 10.1007/s10924-012-0456-z

Biffinger, J. C., Barlow, D. E., Cockrell, A. L., Cusick, K. D., Hervey, W. J., Fitzgerald, L. A., et al. (2015). The applicability of Impranil® DLN for gauging the biodegradation of polyurethanes. *Polym. Degrad. Stab.* 120, 178–185. doi: 10.1016/j.polymdegradstab.2015.06.020

Blair, E. M., Dickson, K. L., and O'Malley, M. A. (2021). Microbial communities and their enzymes facilitate degradation of recalcitrant polymers in anaerobic digestion. *Curr. Opin. Microbiol.* 64, 100–108. doi: 10.1016/J.MIB.2021.09.008

Brunner, I., Fischer, M., Rüthi, J., Stierli, B., and Frey, B. (2018). Ability of fungi isolated from plastic debris floating in the shoreline of a lake to degrade plastics. *PLoS One* 13:e0202047. doi: 10.1371/journal.pone.0202047

Bucknall, D. G. (2020). Plastics as a materials system in a circular economy. *Philos. Trans. R. Soc. A* 378:20190268. doi: 10.1098/RSTA.2019.0268

Bugg, T. D. H., Ahmad, M., Hardiman, E. M., and Rahmanpour, R. (2011). Pathways for degradation of lignin in bacteria and fungi. *Nat. Prod. Rep.* 28, 1883–1896. doi: 10.1039/c1np00042j

Canopoli, L., Coulon, F., and Wagland, S. T. (2020). Degradation of excavated polyethylene and polypropylene waste from landfill. *Sci. Total Environ.* 698:134125. doi: 10.1016/j.scitotenv.2019.134125

Chamas, A., Moon, H., Zheng, J., Qiu, Y., Tabassum, T., Hee Jang, J., et al. (2020). Degradation Rates of Plastics in the Environment. *ACS Sustain. Chem. Eng* 8:3511. doi: 10.1021/acssuschemeng.9b06635

Charnock, C. (2021). A simple and novel method for the production of polyethylene terephthalate containing agar plates for the growth and detection of bacteria able to hydrolyze this plastic. *J. Microbiol. Methods* 185:106222. doi: 10.1016/j.mimet.2021.106222

Chua, T.-K., Tseng, M., and Yang, M.-K. (2013). Degradation of Poly(ε-caprolactone) by thermophilic Streptomyces thermoviolaceus subsp. thermoviolaceus 76T-2. *AMB Express* 3:8. doi: 10.1186/2191-0855-3-8

Couturier, M., Ladevèze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., Moreau, C., et al. (2018). Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat. Chem. Biol.* 14, 306–310. doi: 10.1038/nchembio.2558

Dai, L., Qu, Y., Hu, Y., Min, J., Yu, X., Chen, C. C., et al. (2021a). Catalytically inactive lytic polysaccharide monooxygenase PcAA14A enhances the enzyme-mediated hydrolysis of polyethylene terephthalate. *Int. J. Biol. Macromol.* 190, 456–462. doi: 10.1016/j.ijbiomac.2021.09.005

Dai, L., Qu, Y., Huang, J. W., Hu, Y., Hu, H., Li, S., et al. (2021b). Enhancing PET hydrolytic enzyme activity by fusion of the cellulose–binding domain of cellobiohydrolase I from Trichoderma reesei. *J. Biotechnol.* 334, 47–50. doi: 10.1016/j.jbiotec.2021.05.006

d'Ambrières, W. (2019). Plastics recycling worldwide: current overview and desirable changes. *OpenEdition J.* 19, 12–21.

Danso, D., Chow, J., and Streita, W. R. (2019). Plastics: environmental and biotechnological perspectives on microbial degradation. *Appl. Environ. Microbiol.* 85, e01095–19. doi: 10.1128/AEM.01095-19

do Canto, V. P., Thompson, C. E., and Netz, P. A. (2019). Polyurethanases: three-dimensional structures and molecular dynamics simulations of enzymes that degrade polyurethane. *J. Mol. Graph. Model.* 89, 82–95. doi: 10.1016/J.JMGM.2019.03.001

Domínguez De María, P., Van Gemert, R. W., Straathof, A. J. J., and Hanefeld, U. (2010). Biosynthesis of ethers: unusual or common natural events?. *Nat. Prod. Rep.* 27, 370–392. doi: 10.1039/B809416K

Donnelly, A. E., Murphy, G. S., Digianantonio, K. M., and Hecht, M. H. (2018). A de novo enzyme catalyzes a life-sustaining reaction in *Escherichia coli. Nat. Chem. Biol.* 14, 253–255. doi: 10.1038/nchembio.2550

El-Shafei, H. A., Abd El-Nasser, N. H., Kansoh, A. L., and Ali, A. M. (1998). Biodegradation of disposable polyethylene by fungi and Streptomyces species. *Polym. Degrad. Stab.* 62, 361–365. doi: 10.1016/S0141-3910(98)00019-6

Eriksen, M. K., Christiansen, J. D., Daugaard, A. E., and Astrup, T. F. (2019). Closing the loop for PET, PE and PP waste from households: influence of material properties and product design for plastic recycling. *Waste Manag.* 96, 75–85. doi: 10.1016/J.WASMAN.2019.07.005

Eriksen, M. K., Pivnenko, K., Olsson, M. E., and Astrup, T. F. (2018). Contamination in plastic recycling: influence of metals on the quality of reprocessed plastic. *Waste Manag.* 79, 595–606. doi: 10.1016/J.WASMAN.2018.08.007

Espinosa, M. J. C., Blanco, A. C., Schmidgall, T., Atanasoff-Kardjalieff, A. K., Kappelmeyer, U., Tischler, D., et al. (2020). Toward Biorecycling: isolation of a Soil Bacterium That Grows on a Polyurethane Oligomer and Monomer. *Front. Microbiol.* 11:404. doi: 10.3389/fmicb.2020.00404

Fecker, T., Galaz-Davison, P., Engelberger, F., Narui, Y., Sotomayor, M., Parra, L. P., et al. (2018). Active site flexibility as a hallmark for efficient PET degradation by I. sakaiensis PETase. *Biophys. J.* 114, 1302–1312. doi: 10.1016/J.BPJ.2018.02.005/ATTACHMENT/B9761E8B-970E-4B58-BC07-305833F16609/MMC1.PDF

Furukawa, M., Kawakami, N., Tomizawa, A., and Miyamoto, K. (2019). Efficient Degradation of Poly(ethylene terephthalate) with Thermobifida fusca Cutinase Exhibiting Improved Catalytic Activity Generated using Mutagenesis and Additive-based Approaches. *Sci. Rep.* 9:16038. doi: 10.1038/s41598-019-52379-z

Garcia, J. M., and Robertson, M. L. (2017). The future of plastics recycling. *Science* 358, 870–872. doi: 10.1126/science.aaq0324

Gaytán, I., Sánchez-Reyes, A., Burelo, M., Vargas-Suárez, M., Liachko, I., Press, M., et al. (2020). Degradation of Recalcitrant Polyurethane and Xenobiotic Additives by a Selected Landfill Microbial Community and Its Biodegradative Potential Revealed by Proximity Ligation-Based Metagenomic Analysis. *Front. Microbiol.* 10:2986. doi: 10.3389/FMICB.2019.02986/BIBTEX

Geyer, R., Jambeck, J. R., and Law, K. L. (2017). Production, use, and fate of all plastics ever made. *Science Adv.* 3:e1700782. doi: 10.1126/sciadv.1700782

Grigore, M. E. (2017). Methods of Recycling, Properties and Applications of Recycled Thermoplastic Polymers. *Recycling* 2:24. doi: 10.3390/recycling2040024

Guengerich, F. P., and Yoshimoto, F. K. (2018). Formation and Cleavage of C-C Bonds by Enzymatic Oxidation-Reduction Reactions. *Chem. Rev.* 118, 6573–6655. doi: 10.1021/acs.chemrev.8b00031

Guerrera, M. C., Aragona, M., Porcino, C., Fazio, F., Laurà, R., Levanti, M., et al. (2021). Micro and nano plastics distribution in fish as model organisms: histopathology, blood response and bioaccumulation in different organs. *Appl. Sci.* 11:5768. doi: 10.3390/app11135768

Ha, J. H., Karchin, J. M., Walker-Kopp, N., Castañeda, C. A., and Loh, S. N. (2015). Engineered Domain Swapping as an On/Off Switch for Protein Function. *Chem. Biol.* 22, 1384–1393. doi: 10.1016/j.chembiol.2015.09.007

Hahladakis, J. N., and Iacovidou, E. (2019). An overview of the challenges and trade-offs in closing the loop of post-consumer plastic waste (PCPW): focus on recycling. *J. Hazard. Mater.* 380:120887. doi: 10.1016/J.JHAZMAT.2019.120887

Hajighasemi, M., Nocek, B. P., Tchigvintsev, A., Brown, G., Flick, R., Xu, X., et al. (2016). Biochemical and Structural Insights into Enzymatic Depolymerization of Polylactic Acid and Other Polyesters by Microbial Carboxylesterases. *Biomacromolecules* 17, 2027–2039. doi: 10.1021/acs.biomac.6b00223

Han, X., Liu, W., Huang, J. W., Ma, J., Zheng, Y., Ko, T. P., et al. (2017). Structural insight into catalytic mechanism of PET hydrolase. *Nat. Commun.* 8:2106. doi: 10.1038/s41467-017-02255-z

Henry, B., Laitala, K., and Klepp, I. G. (2019). Microfibres from apparel and home textiles: prospects for including microplastics in environmental sustainability assessment. *Sci. Total Environ.* 652, 483–494. doi: 10.1016/J.SCITOTENV.2018.10.166

Herrero Acero, E., Ribitsch, D., Dellacher, A., Zitzenbacher, S., Marold, A., Steinkellner, G., et al. (2013). Surface engineering of a cutinase from Thermobifida cellulosilytica for improved polyester hydrolysis. *Biotechnol. Bioeng.* 110, 2581–2590. doi: 10.1002/BIT.24930

Herrero Acero, E., Ribitsch, D., Steinkellner, G., Gruber, K., Greimel, K., Eiteljoerg, I., et al. (2011). Enzymatic surface hydrolysis of PET: effect of structural diversity on kinetic properties of cutinases from Thermobifida. *Macromolecules* 44, 4632–4640. doi: 10.1021/ma200949p

Hou, L., Ji, D., Dong, W., Yuan, L., Zhang, F., Li, Y., et al. (2020). The Synergistic Action of Electro-Fenton and White-Rot Fungi in the Degradation of Lignin. *Front. Bioeng. Biotechnol.* 8:99. doi: 10.3389/fbioe.2020.00099

Howard, G. T., Ruiz, C., and Hilliard, N. P. (1999). Growth of *Pseudomonas chlororaphis* on a polyester-polyurethane and the purification and characterization of a polyurethanase-esterase enzyme. *Int. Biodeterior. Biodegradation* 43, 7–12. doi: 10.1016/S0964-8305(98)00057-2

Hussain Ali, L., Ali, S., Nnaji, C. F., Ogu, E. C., and Akpor, O. B. (2021). Biosurfactants as facilitators in Biodegradation of Low-Density Polyethylene (LDPE). *IOP Conf. Ser. Mater. Sci. Eng.* 1107:012135. doi: 10.1088/1757-899X/1107/1/012135

Igeño, M. I., Macias, D., and Blasco, R. (2019). A Case of Adaptive Laboratory Evolution (ALE): biodegradation of Furfural by *Pseudomonas* pseudoalcaligenes CECT 5344. *Genes* 10:499. doi: 10.3390/genes10070499

Ioakeimidis, C., Fotopoulou, K. N., Karapanagioti, H. K., Geraga, M., Zeri, C., Papathanassiou, E., et al. (2016). The degradation potential of PET bottles in the marine environment: an ATR-FTIR based approach. *Sci. Rep.* 6:23501. doi: 10.1038/srep23501

Iqbal, S., Xu, J., Allen, S. D., Khan, S., Nadir, S., Arif, M. S., et al. (2020). Unraveling consequences of soil micro- and nano-plastic pollution on soil-plant system: implications for nitrogen (N) cycling and soil microbial activity. *Chemosphere* 260:127578. doi: 10.1016/j.chemosphere.2020.127578

Jain, A., Jain, R., and Jain, S. (2020). *Basic techniques in biochemistry, microbiology and Molecular biology principles and Techniques*. New York: Humana Press, 39–51. doi: 10.1007/978-1-4939-9861-6_54

Jaiswal, S., Sharma, B., and Shukla, P. (2020). Integrated approaches in microbial degradation of plastics. *Environ. Technol. Innov.* 17:100567. doi: 10.1016/j.eti.2019.100567

Janatunaim, R. Z., and Fibriani, A. (2020). Construction and cloning of Plastic-Degrading Recombinant Enzymes (MHETase). *Recent Pat. Biotechnol.* 14, 229–234. doi: 10.2174/1872208314666200311104541

Janczak, K., Hrynkiewicz, K., Znajewska, Z., and Dąbrowska, G. (2018). Use of rhizosphere microorganisms in the biodegradation of PLA and PET polymers in compost soil. *Int. Biodeterior. Biodegradation* 130, 65–75. doi: 10.1016/j.ibiod.2018.03.017

Kang, D., Jacquiod, S., Herschend, J., Wei, S., Nesme, J., and Sørensen, S. J. (2020). Construction of Simplified Microbial Consortia to Degrade Recalcitrant Materials Based on Enrichment and Dilution-to-Extinction Cultures. *Front. Microbiol.* 10:3010. doi: 10.3389/FMICB.2019.03010/BIBTEX

Kawai, F., Kawabata, T., and Oda, M. (2019). Current knowledge on enzymatic PET degradation and its possible application to waste stream management and other fields. *Appl. Microbiol. Biotechnol.* 103, 4253–4268. doi: 10.1007/s00253-019-09717-y

Khandare, S. D., Chaudhary, D. R., and Jha, B. (2021). Marine bacterial biodegradation of low-density polyethylene (LDPE) plastic. *Biodegradation* 32, 127–143. doi: 10.1007/s10532-021-09927-0

Khoonkari, M., Haghighi, A. H., Sefidbakht, Y., Shekoohi, K., and Ghaderian, A. (2015). Chemical Recycling of PET Wastes with Different Catalysts. *Int. J. Polym. Sci.* 2015, 1–11. doi: 10.1155/2015/124524

Kim, H. R., Lee, H. M., Yu, H. C., Jeon, E., Lee, S., Li, J., et al. (2020). Biodegradation of Polystyrene by *Pseudomonas* sp. Isolated from the Gut of Superworms (Larvae of Zophobas atratus). *Environ. Sci. Technol.* 54, 6987–6996. doi: 10.1021/acs.est.0c01495/suppl_file/es0c01495_si_001.pdf

Kim, H. T., Kim, J. K., Cha, H. G., Kang, M. J., Lee, H. S., Khang, T. U., et al. (2019). Biological Valorization of Poly(ethylene terephthalate) Monomers for Upcycling Waste PET. *ACS Sustain. Chem. Eng.* 7, 19396–19406. doi: 10.1021/acssuschemeng.9b03908

Kim, M. Y., Kim, C., Moon, J., Heo, J., Jung, S. P., and Kim, J. R. (2017). Polymer Film-Based Screening and Isolation of Polylactic Acid (PLA)-Degrading Microorganisms. *J. Microbiol. Biotechnol.* 27, 342–349. doi: 10.4014/JMB.1610.10015

Kitadokoro, K., Kakara, M., Matsui, S., Osokoshi, R., Thumarat, U., Kawai, F., et al. (2019). Structural insights into the unique polylactate-degrading mechanism of Thermobifida alba cutinase. *FEBS J.* 286, 2087–2098. doi: 10.1111/febs.14781

Kitadokoro, K., Thumarat, U., Nakamura, R., Nishimura, K., Karatani, H., Suzuki, H., et al. (2012). Crystal structure of cutinase Est119 from Thermobifida alba AHK119 that can degrade modified polyethylene terephthalate at 1.76 Å resolution. *Polym. Degrad. Stab.* 97, 771–775. doi: 10.1016/j.polymdegradstab.2012.02.003

Kjeldsen, F., and Zubarev, R. A. (2011). Effects of peptide backbone amide-to-ester bond substitution on the cleavage frequency in electron capture dissociation and collision-activated dissociation. *J. Am. Soc. Mass Spectrom.* 22, 1441–1452. doi: 10.1007/S13361-011-0151-7

Kumar, M., Chen, H., Sarsaiya, S., Qin, S., Liu, H., Awasthi, M. K., et al. (2021). Current research trends on micro- and nano-plastics as an emerging threat to global environment: a review. *J. Hazard. Mater.* 409:124967. doi: 10.1016/j.jhazmat.2020.124967

Kumari, A., Bano, N., Bag, S. K., Chaudhary, D. R., and Jha, B. (2021). Transcriptome-Guided Insights Into Plastic Degradation by the Marine Bacterium. *Front. Microbiol.* 12:2761. doi: 10.3389/fmicb.2021.751571/bibtex

Lau, W., Shiran, Y., Bailey, R. M., Cook, E., Stuchtey, M. R., Koskella, J., et al. (2020). Evaluating scenarios toward zero plastic pollution. *Science* 369, 1455–1461. doi: 10.1126/science.aba9475

Lee, A., and Liew, M. S. (2021). Tertiary recycling of plastics waste: an analysis of feedstock, chemical and biological degradation methods. *J. Mater. Cycles Waste Manag.* 23, 32–43. doi: 10.1007/s10163-020-01106-2

Lee, S. R., and Kim, P. (2020). Current Status and Applications of Adaptive Laboratory Evolution in Industrial Microorganisms. *J. Microbiol. Biotechnol.* 30, 793–803. doi: 10.4014/JMB.2003.03072

Li, C. T., Zhang, M., Weng, Y. X., and Qin, J. X. (2018). Influence of ether linkage on the enzymatic degradation of PBS copolymers: comparative study on poly (butylene succinate-co-diethylene glycol succinate) and poly (butylene succinate-co-butylene diglycolic acid). *Int. J. Biol. Macromol.* 118, 347–356. doi: 10.1016/J.IJBIOMAC.2018.06.062

Li, F., Wang, S., Liu, W., and Chen, G. (2008). Purification and characterization of poly(l-lactic acid)-degrading enzymes from Amycolatopsis orientalis ssp. orientalis. *FEMS Microbiol. Lett.* 282, 52–58. doi: 10.1111/j.1574-6968.2008.01109.x

Li, F., Yu, D., Lin, X., Liu, D., Xia, H., and Chen, S. (2012). Biodegradation of poly(ε-caprolactone) (PCL) by a new Penicillium oxalicum strain DSYD05-1. *World J. Microbiol. Biotechnol.* 28, 2929–2935. doi: 10.1007/S11274-012-1103-5/FIGURES/6

Li, W. J., Jayakody, L. N., Franden, M. A., Wehrmann, M., Daun, T., Hauer, B., et al. (2019). Laboratory evolution reveals the metabolic and regulatory basis of ethylene glycol metabolism by *Pseudomonas putida* KT2440. *Environ. Microbiol.* 21, 3669–3682. doi: 10.1111/1462-2920.14703

Liu, B., He, L., Wang, L., Li, T., Li, C., Liu, H., et al. (2018). Protein crystallography and site-direct mutagenesis analysis of the poly(Ethylene terephthalate) hydrolase PETase from Ideonella sakaiensis. *ChemBioChem* 19, 1471–1475. doi: 10.1002/cbic.201800097

Lomonaco, T., Manco, E., Corti, A., La Nasa, J., Ghimenti, S., Biagini, D., et al. (2020). Release of harmful volatile organic compounds (VOCs) from photo-degraded plastic debris: a neglected source of environmental pollution. *J. Hazard. Mater.* 394:122596. doi: 10.1016/J.JHAZMAT.2020.122596

Lopes, A. M., Ferreira Filho, E. X., and Moreira, L. R. S. (2018). An update on enzymatic cocktails for lignocellulose breakdown. *J. Appl. Microbiol.* 125, 632–645. doi: 10.1111/jam.13923

Ma, Y., Yao, M., Li, B., Ding, M., He, B., Chen, S., et al. (2018). Enhanced Poly(ethylene terephthalate) Hydrolase Activity by Protein Engineering. *Engineering* 4, 888–893. doi: 10.1016/j.eng.2018.09.007

Magnin, A., Entzmann, L., Pollet, E., and Avérous, L. (2021). Breakthrough in polyurethane bio-recycling: an efficient laccase-mediated system for the degradation of different types of polyurethanes. *Waste Manag.* 132, 23–30. doi: 10.1016/J.WASMAN.2021.07.011

Magnin, A., Pollet, E., Perrin, R., Ullmann, C., Persillon, C., Phalip, V., et al. (2019). Enzymatic recycling of thermoplastic polyurethanes: synergistic effect of an esterase and an amidase and recovery of building blocks. *Waste Manag.* 85, 141–150. doi: 10.1016/j.wasman.2018.12.024

Maity, S., Banerjee, S., Biswas, C., Guchhait, R., Chatterjee, A., and Pramanick, K. (2021). Functional interplay between plastic polymers and microbes: a comprehensive review. *Biodegradation* 32, 487–510. doi: 10.1007/s10532-021-09954-x

Mao, J., Jain, A., Denslow, N. D., Nouri, M. Z., Chen, S., Wang, T., et al. (2020). Bisphenol A and bisphenol S disruptions of the mouse placenta and potential effects on the placenta-brain axis. *Proc. Natl. Acad. Sci. U. S. A.* 117, 4642–4652. doi: 10.1073/pnas.1919563117

Meng, D.-D., Ying, Y., Chen, X.-H., Lu, M., Ning, K., Wang, L.-S., et al. (2015). Distinct Roles for Carbohydrate-Binding Modules of Glycoside Hydrolase 10 (GH10) and GH11 Xylanases from Caldicellulosiruptor sp. Strain F32 in Thermostability and Catalytic Efficiency. *Appl. Environ. Microbiol.* 81:2006. doi: 10.1128/aem.03677-14

Mohammed, S., Behera, H. T., Dekebo, A., and Ray, L. (2020). Optimization of the culture conditions for production of Polyhydroxyalkanoate and its characterization from a new Bacillus cereus sp. BNPI-92 strain, isolated from plastic waste dumping yard. *Int. J. Biol. Macromol.* 156, 1064–1080. doi: 10.1016/J.IJBIOMAC.2019.11.138

Mohanan, N., Montazer, Z., Sharma, P. K., and Levin, D. B. (2020). Microbial and Enzymatic Degradation of Synthetic Plastics. *Front. Microbiol.* 11:580709. doi: 10.3389/fmicb.2020.580709

Molitor, R., Bollinger, A., Kubicki, S., Loeschcke, A., Jaeger, K., and Thies, S. (2020). Agar plate-based screening methods for the identification of polyester hydrolysis by *Pseudomonas* species. *Microb. Biotechnol.* 13:274. doi: 10.1111/1751-7915.13418

Nakasaki, K., Hirai, H., Mimoto, H., Quyen, T. N. M., Koyama, M., and Takeda, K. (2019). Succession of microbial community during vigorous organic matter degradation in the primary fermentation stage of food waste composting. *Sci. Total Environ.* 671, 1237–1244. doi: 10.1016/J.SCITOTENV.2019.03.341

Negi, H., Kapri, A., Zaidi, M. G. H., Satlewal, A., and Goel, R. (2009). Comparative in-vitro biodegradation studies of epoxy and its silicone blend by selected microbial consortia. *Int. Biodeterior. Biodegradation* 63, 553–558. doi: 10.1016/J.IBIOD.2009.03.001

Nemli, G., Ayan, E., Ay, N., and Tiryaki, S. (2018). Utilization potential of waste wood subjected to insect and fungi degradation for particleboard manufacturing. *Eur. J. Wood Wood Prod.* 76, 759–766. doi: 10.1007/s00107-017-1224-5

Nishimura, H., Kamiya, A., Nagata, T., Katahira, M., and Watanabe, T. (2018). Direct evidence for α ether linkage between lignin and carbohydrates in wood cell walls. *Sci. Rep.* 8:6538. doi: 10.1038/s41598-018-24328-9

Ojha, N., Pradhan, N., Singh, S., Barla, A., Shrivastava, A., Khatua, P., et al. (2017). Evaluation of HDPE and LDPE degradation by fungus, implemented by statistical optimization. *Sci. Rep.* 7:39515. doi: 10.1038/srep39515

Palm, G. J., Reisky, L., Böttcher, D., Müller, H., Michels, E. A. P., Walczak, M. C., et al. (2019). Structure of the plastic-degrading Ideonella sakaiensis MHETase bound to a substrate. *Nat. Commun.* 10:1717. doi: 10.1038/s41467-019-09326-3

Pan, W., Bai, Z., Su, T., and Wang, Z. (2018). Enzymatic degradation of poly(butylene succinate) with different molecular weights by cutinase. *Int. J. Biol. Macromol.* 111, 1040–1046. doi: 10.1016/J.IJBIOMAC.2018.01.107

Pantani, R., and Sorrentino, A. (2013). Influence of crystallinity on the biodegradation rate of injection-moulded poly(lactic acid) samples in controlled composting conditions. *Polym. Degrad. Stab.* 98, 1089–1096. doi: 10.1016/j.polymdegradstab.2013.01.005

Pathak, V. M., and Navneet. (2017). Review on the current status of polymer degradation: a microbial approach. *Bioresour. Bioprocess.* 4:15. doi: 10.1186/s40643-017-0145-9

Pellis, A., Haernvall, K., Pichler, C. M., Ghazaryan, G., Breinbauer, R., and Guebitz, G. M. (2016). Enzymatic hydrolysis of poly(ethylene furanoate). *J. Biotechnol.* 235, 47–53. doi: 10.1016/j.jbiotec.2016.02.006

PlasticEurope-Association of Plastics Manufacturers (2020). *Plastics – the Facts 2020*. Brussels: PlasticEurope, 1–64.

Pometto, A. L., Lee, B., and Johnson, K. E. (1992). Production of an extracellular polyethylene-degrading enzyme(s) by Streptomyces species. *Appl. Environ. Microbiol.* 58:731. doi: 10.1128/aem.58.2.731-733.1992

Pranamuda, H., Tokiwa, Y., and Tanaka, H. (1997). Polylactide degradation by an Amycolatopsis sp. *Appl. Environ. Microbiol.* 63, 1637–1640. doi: 10.1128/aem.63.4.1637-1640.1997

Priya, A., Dutta, K., and Daverey, A. (2022). A comprehensive biotechnological and molecular insight into plastic degradation by microbial community. *J. Chem. Technol. Biotechnol.* 97, 381–390. doi: 10.1002/jctb.6675

Przemieniecki, S. W., Kosewska, A., Ciesielski, S., and Kosewska, O. (2020). Changes in the gut microbiome and enzymatic profile of Tenebrio molitor larvae biodegrading cellulose, polyethylene and polystyrene waste. *Environ. Pollut.* 256:113265. doi: 10.1016/J.ENVPOL.2019.113265

Qu, Q., Zhang, J., Chen, X., Ravanbakhsh, H., Tang, G., Xiong, R., et al. (2021). Triggered Release from Cellulose Microparticles Inspired by Wood Degradation by Fungi. *ACS Sustain. Chem. Eng.* 9, 387–397. doi: 10.1021/acssuschemeng.0c07514

Rana, K. I. (2019). Usage of Potential Micro-organisms for Degradation of Plastics. *Open J. Environ. Biol.* 4, 7–15. doi: 10.17352/ojeb.000010

Ribitsch, D., Acero, E. H., Greimel, K., Eiteljoerg, I., Trotscha, E., Freddi, G., et al. (2012). Characterization of a new cutinase from Thermobifida alba for PET-surface hydrolysis. *Biocatal. Biotransformation* 30, 2–9. doi: 10.3109/10242422.2012.644435

Ribitsch, D., Acero, E. H., Przylucka, A., Zitzenbacher, S., Marold, A., Gamerith, C., et al. (2015). Enhanced cutinase-catalyzed hydrolysis of polyethylene terephthalate by covalent fusion to hydrophobins. *Appl. Environ. Microbiol.* 81, 3586–3592. doi: 10.1128/AEM.04111-14

Ribitsch, D., Heumann, S., Trotscha, E., Herrero Acero, E., Greimel, K., Leber, R., et al. (2011). Hydrolysis of polyethyleneterephthalate by p-nitrobenzylesterase from Bacillus subtilis. *Biotechnol. Prog.* 27, 951–960. doi: 10.1002/btpr.610

Roohi, Bano, K., Kuddus, M., Zaheer, M. R., Zia, Q., Khan, M. F., Ashraf, G. M., et al. (2017). Microbial Enzymatic Degradation of Biodegradable Plastics. *Curr. Pharm. Biotechnol.* 18, 429–440. doi: 10.2174/1389201018666170523165742

Ruiz-Dueñas, F. J., Barrasa, J. M., Sánchez-García, M., Camarero, S., Miyauchi, S., Serrano, A., et al. (2020). Genomic Analysis Enlightens Agaricales Lifestyle Evolution and Increasing Peroxidase Diversity. *Mol. Biol. Evol.* 38, 1428–1446. doi: 10.1093/molbev/msaa301

Russell, J. R., Huang, J., Anand, P., Kucera, K., Sandoval, A. G., Dantzler, K. W., et al. (2011). Biodegradation of polyester polyurethane by endophytic fungi. *Appl. Environ. Microbiol.* 77, 6076–6084. doi: 10.1128/aem.00521-11

Sabapathy, P. C., Devaraj, S., Parthipan, A., and Kathirvel, P. (2019). Polyhydroxyalkanoate production from statistically optimized media using rice mill effluent as sustainable substrate with an analysis on the biopolymer's degradation potential. *Int. J. Biol. Macromol.* 126, 977–986. doi: 10.1016/J.IJBIOMAC.2019.01.003

Sadler, J. C., and Wallace, S. (2021). Microbial synthesis of vanillin from waste poly(ethylene terephthalate). *Green Chem.* 23, 4665–4672. doi: 10.1039/D1GC00931A

Sandt, C., Waeytens, J., Deniset-Besseau, A., Nielsen-Leroux, C., and Réjasse, A. (2021). Use and misuse of FTIR spectroscopy for studying the bio-oxidation of plastics. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* 258:119841. doi: 10.1016/j.saa.2021.119841

Sankhla, I. S., Sharma, G., and Tak, A. (2020). "Fungal degradation of bioplastics: An overview," in *New and Future Developments in Microbial Biotechnology and Bioengineering*, eds J. Singh and P. Gehlot (Amsterdam: Elsevier). doi: 10.1016/b978-0-12-821007-9.00004-8

Science Advice for Policy by European Academies [SAPEA] (2019). *Micro-Plastics: A Scientific Perspective On In Nature And Society*. Available online at: https://www.sapea.info/topics/microplastics/ (accessed January 26, 2022).

Sarkar, B., Dissanayake, P. D., Bolan, N. S., Dar, J. Y., Kumar, M., Haque, M. N., et al. (2022). Challenges and opportunities in sustainable management of microplastics and nanoplastics in the environment. *Environ. Res.* 207:112179. doi: 10.1016/J.ENVRES.2021.112179

Schmidt, J., Wei, R., Oeser, T., Silva, L. A. D. E. S., Breite, D., Schulze, A., et al. (2017). Degradation of polyester polyurethane by bacterial polyester hydrolases. *Polymers* 9:65. doi: 10.3390/polym9020065

Selonen, S., Dolar, A., Jemec Kokalj, A., Skalar, T., Parramon Dolcet, L., Hurley, R., et al. (2020). Exploring the impacts of plastics in soil – The effects of polyester textile fibers on soil invertebrates. *Sci. Total Environ.* 700:134451. doi: 10.1016/j.scitotenv.2019.134451

Seo, H., Kim, S., Son, H. F., Sagong, H. Y., Joo, S., and Kim, K. J. (2019). Production of extracellular PETase from Ideonella sakaiensis using sec-dependent signal peptides in E. coli. *Biochem. Biophys. Res. Commun.* 508, 250–255. doi: 10.1016/J.BBRC.2018.11.087

Sepperumal, U., Markandan, M., and Palraja, I. (2013). Micromorphological and chemical changes during biodegradation of Polyethylene terephthalate (PET) by Penicillium sp. *J. Microbiol. Biotechnol. Res.* 3, 47–53.

Shahid Kashif, M., Kashif, A., and Choi, Y. (2021). "Advances in the recycling of polymer-based plastic materials," in *Urban Mining for Waste Management and Resource Recovery*, (Boca Raton: CRC Press), 101–110. doi: 10.1201/9781003201076-6

Shen, M., Zhu, Y., Zhang, Y., Zeng, G., Wen, X., Yi, H., et al. (2019). Micro(nano)plastics: unignorable vectors for organisms. *Mar. Pollut. Bull.* 139, 328–331. doi: 10.1016/j.marpolbul.2019.01.004

Shin, G., Park, S. A., Koo, J. M., Kim, M., Lee, M., Jegal, J., et al. (2021). A micro-spray-based high-throughput screening system for bioplastic-degrading microorganisms. *Green Chem.* 23, 5429–5436. doi: 10.1039/d1gc01916c

Shinozaki, Y., Watanabe, T., Nakajima-Kambe, T., and Kitamoto, H. K. (2013). Rapid and simple colorimetric assay for detecting the enzymatic degradation of biodegradable plastic films. *J. Biosci. Bioeng.* 115, 111–114. doi: 10.1016/J.JBIOSC.2012.08.010

Shiu, R. F., Vazquez, C. I., Tsai, Y. Y., Torres, G. V., Chen, C. S., Santschi, P. H., et al. (2020). Nano-plastics induce aquatic particulate organic matter (microgels) formation. *Sci. Total Environ.* 706:135681. doi: 10.1016/j.scitotenv.2019.135681

Skariyachan, S., Manjunatha, V., Sultana, S., Jois, C., Bai, V., and Vasist, K. S. (2016). Novel bacterial consortia isolated from plastic garbage processing areas demonstrated enhanced degradation for low density polyethylene. *Environ. Sci. Pollut. Res. Int.* 23, 18307–18319. doi: 10.1007/S11356-016-7000-Y

Skariyachan, S., Patil, A. A., Shankar, A., Manjunath, M., Bachappanavar, N., and Kiran, S. (2018). Enhanced polymer degradation of polyethylene and polypropylene by novel thermophilic consortia of Brevibacillus sps. and Aneurinibacillus sp. screened from waste management landfills and sewage treatment plants. *Polym. Degrad. Stab.* 149, 52–68. doi: 10.1016/J.POLYMDEGRADSTAB.2018.01.018

Skariyachan, S., Setlur, A. S., Naik, S. Y., Naik, A. A., Usharani, M., and Vasist, K. S. (2017). Enhanced biodegradation of low and high-density polyethylene by novel bacterial consortia formulated from plastic-contaminated cow dung under thermophilic conditions. *Environ. Sci. Pollut. Res.* 24, 8443–8457. doi: 10.1007/S11356-017-8537-0/FIGURES/6

Skariyachan, S., Taskeen, N., Kishore, A. P., and Krishna, B. V. (2022). Recent advances in plastic degradation – From microbial consortia-based methods to data sciences and computational biology driven approaches. *J. Hazard. Mater.* 426:128086. doi: 10.1016/J.JHAZMAT.2021.128086

Skariyachan, S., Taskeen, N., Kishore, A. P., Krishna, B. V., and Naidu, G. (2021). Novel consortia of Enterobacter and Pseudomonas formulated from cow dung exhibited enhanced biodegradation of polyethylene and polypropylene. *J. Environ. Manag.* 284:112030. doi: 10.1016/J.JENVMAN.2021.112030

Spina, F., Tummino, M. L., Poli, A., Prigione, V., Ilieva, V., Cocconcelli, P., et al. (2021). Low density polyethylene degradation by filamentous fungi. *Environ. Pollut.* 274:116548. doi: 10.1016/j.envpol.2021.116548

Suzuki, K., Noguchi, M. T., Shinozaki, Y., Koitabashi, M., Sameshima-Yamashita, Y., Yoshida, S., et al. (2014). Purification, characterization, and cloning of the gene for a biodegradable plastic-degrading enzyme from Paraphoma-related fungal strain B47-9. *Appl. Microbiol. Biotechnol.* 98, 4457–4465. doi: 10.1007/s00253-013-5454-0

Taghavi, N., Zhuang, W.-Q., and Baroutian, S. (2021). Effect of rhamnolipid biosurfactant on biodegradation of untreated and UV-pretreated non-degradable thermoplastics: part 2. *J. Environ. Chem. Eng.* 10:107033.

Talkad, M. S., Chethan, C., Kavya, S., Qudsiya, S. S., Maria, S., Raj, A., et al. (2014). Microbial Degradation of Plastic (LDPE) and domestic waste by induced mutations in *Pseudomonas putida*. *Int. J. Ethics Eng. Manag. Educ.* 1, 2348–4748.

Tang, Z.-L., Kuo, T.-A., and Liu, H.-H. (2017). The Study of the Microbes Degraded Polystyrene. *Adv. Technol. Innov.* 2, 13–17.

Thapa, C., Shakya, P., Shrestha, R., Pal, S., and Manandhar, P. (2018). Isolation of Polyhydroxybutyrate (PHB) Producing Bacteria, Optimization of Culture Conditions for PHB production, Extraction and Characterization of PHB. *Nepal J. Biotechnol.* 6, 62–68. doi: 10.1016/j.colsurfb.2009.07.029

Tiso, T., Winter, B., Wei, R., Hee, J., de Witt, J., Wierckx, N., et al. (2021). The metabolic potential of plastics as biotechnological carbon sources – Review and targets for the future. *Metab. Eng.* doi: 10.1016/j.ymben.2021.12.006

Tiwari, N., Santhiya, D., and Sharma, J. G. (2020). Microbial remediation of micro-nano plastics: current knowledge and future trends. *Environ. Pollut.* 265:115044. doi: 10.1016/j.envpol.2020.115044

Tournier, V., Topham, C. M., Gilles, A., David, B., Folgoas, C., Moya-Leclair, E., et al. (2020). An engineered PET depolymerase to break down and recycle plastic bottles. *Nature* 580, 216–219. doi: 10.1038/s41586-020-2149-4

Usha, R., Sangeetha, T., and Palaniswamy, M. (2011). Screening of polyethylene degrading microorganisms from garbage soil. *Libyan Agric. Res. Cent. J. Int.* 2, 200–204.

Valente, F., Allardyce, B. J., Hepburn, M. S., Wijesinghe, P., Redmond, S. L., Chen, J., et al. (2020). Enhancing Resistance of Silk Fibroin Material to Enzymatic Degradation by Cross-Linking Both Crystalline and Amorphous Domains. *ACS Biomater. Sci. Eng.* 6, 2459–2468. doi: 10.1021/ACSBIOMATERIALS.9B00873/SUPPL_FILE/AB9B00873_SI_001.PDF

Vázquez-Alcántara, L., Oliart-Ros, R. M., García-Bórquez, A., and Peña-Montes, C. (2021). Expression of a Cutinase of Moniliophthora roreri with Polyester and PET-Plastic Residues Degradation Activity. *Microbiol. Spectr.* 9:e0097621. doi: 10.1128/spectrum.00976-21

Verma, S., Kumar, R., and Meghwanshi, G. K. (2019). Identification of new members of alkaliphilic lipases in archaea and metagenome database using reconstruction of ancestral sequences. *3 Biotech* 9:165. doi: 10.1007/s13205-019-1693-9

Voshol, G. P., Vijgenboom, E., and Punt, P. J. (2017). The discovery of novel LPMO families with a new Hidden Markov model. *BMC Res. Notes* 10:105. doi: 10.1186/S13104-017-2429-8

Wallace, N. E., Adams, M. C., Chafin, A. C., Jones, D. D., Tsui, C. L., and Gruber, T. D. (2020). The highly crystalline PET found in plastic water bottles does not support the growth of the PETase-producing bacterium Ideonella sakaiensis. *Environ. Microbiol. Rep.* 12, 578–582. doi: 10.1111/1758-2229.12878

Wang, L., Wu, W. M., Bolan, N. S., Tsang, D. C. W., Li, Y., Qin, M., et al. (2021). Environmental fate, toxicity and risk management strategies of nanoplastics in the environment: current status and future perspectives. *J. Hazard. Mater.* 401:123415. doi: 10.1016/J.JHAZMAT.2020.123415

Wang, M.-L., Chang, R.-Y., and Hsu, C.-H. (2018). *Moulding Simulation: Theory and Practice. Material Properties of Plastics*. Cincinnati: Hanser Publishers.19–54. doi: 10.3139/9781569906200.002

Wang, W., Yuan, W., Xu, E. G., Li, L., Zhang, H., and Yang, Y. (2022). Uptake, translocation, and biological impacts of micro(nano)plastics in terrestrial plants: progress and prospects. *Environ. Res.* 203:111867. doi: 10.1016/j.envres.2021.111867

Weinberger, S., Canadell, J., Quartinello, F., Yeniad, B., Arias, A., Pellis, A., et al. (2017). Enzymatic Degradation of Poly(ethylene 2,5-furanoate) Powders and Amorphous Films. *Catalysts* 7:318. doi: 10.3390/catal7110318

Wu, X., Li, J., Yao, L., and Xu, Z. (2020). Auto-sorting commonly recovered plastics from waste household appliances and electronics using near-infrared

spectroscopy. *J. Clean. Prod.* 246:118732. doi: 10.1016/J.JCLEPRO.2019.118732

Xiong, K., Yan, Z. X., Liu, J. Y., Pei, P. G., Deng, L., Gao, L., et al. (2020). Inter domain interactions influence the substrate affinity and hydrolysis product specificity of xylanase from Streptomyces chartreusis L1105. *Ann. Microbiol.* 70:19. doi: 10.1186/S13213-020-01560-1

Yang, Y., Wang, J., and Xia, M. (2020). Biodegradation and mineralization of polystyrene by plastic-eating super worms Zophobas atratus. *Sci. Total Environ.* 708:135233. doi: 10.1016/j.scitotenv.2019.135233

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., et al. (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* 351, 1196–1199. doi: 10.1126/science.aad6359

Zhang, D., Ng, E. L., Hu, W., Wang, H., Galaviz, P., Yang, H., et al. (2020). Plastic pollution in croplands threatens long-term food security. *Glob. Chang. Biol.* 26, 3356–3367. doi: 10.1111/GCB.15043

Zhang, Y., Li, Y., Su, F., Peng, L., and Liu, D. (2022). The life cycle of micro-nano plastics in domestic sewage. *Sci. Total Environ.* 802:149658. doi: 10.1016/j.scitotenv.2021.149658

Zhang, L., You, T., Zhou, T., Zhang, L., and Xu, F. (2016). Determining lignin degradation in white-rot fungi-treated Sacrau poplar: lignin structural changes and degradation compound analysis. *BioResources* 11, 3972–3986. doi: 10.15376/biores.11.2.3972-3986

Zhong-Johnson, E. Z. L., Voigt, C. A., and Sinskey, A. J. (2021). An absorbance method for analysis of enzymatic degradation kinetics of poly(ethylene terephthalate) films. *Sci. Rep.* 11:928. doi: 10.1038/s41598-020-79031-5

Zhou, Y., Kumar, M., Sarsaiya, S., Sirohi, R., Awasthi, S. K., Sindhu, R., et al. (2022). Challenges and opportunities in bioremediation of micro-nano plastics: a review. *Sci. Total Environ.* 802:149823. doi: 10.1016/J.SCITOTENV.2021.149823

Zitare, U. A., Habib, M. H., Rozeboom, H., Mascotti, M. L., Todorovic, S., and Fraaije, M. W. (2021). Mutational and structural analysis of an ancestral fungal dye-decolorizing peroxidase. *FEBS J.* 288, 3602–3618. doi: 10.1111/FEBS.15687

Organic Phosphorus Scavenging Supports Efficient Growth of Diazotrophic Cyanobacteria Under Phosphate Depletion

Sophie Rabouille[1,2]*, Lauralie Tournier[1], Solange Duhamel[3], Pascal Claquin[4], Olivier Crispi[2], Amélie Talec[1], Angela Landolfi[5,6] and Andreas Oschlies[5]

[1]Laboratoire d'Océanographie de Villefranche (LOV), CNRS, Sorbonne Université, Villefranche-sur-Mer, France, [2]Laboratoire d'Océanographie Microbienne (LOMIC), CNRS, Sorbonne Université, Banyuls-sur-Mer, France, [3]Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ, United States, [4]UMR BOREA (CNRS 8067), MNHN, IRD (207), Normandie Université, Université de Caen Normandie, CREC, Luc-sur-Mer, France, [5]GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, [6]CNR ISMAR, Rome, Italy

Considering the reported significant diazotrophic activities in open-ocean regions where primary production is strongly limited by phosphate, we explored the ability of diazotrophs to use other sources of phosphorus to alleviate the phosphate depletion. We tested the actual efficiency of the open-ocean, N_2-fixer *Crocosphaera watsonii* to grow on organic phosphorus as the sole P source, and observed how the P source affects the cellular C, N, and P composition. We obtained equivalent growth efficiencies on AMP and DL-α-glycerophosphate as compared with identical cultures grown on phosphate, and survival of the population on phytic acid. Our results show that *Crocosphaera* cannot use all phosphomonoesters with the same efficiency, but it can grow without phosphate, provided that usable DOP and sufficient light energy are available. Also, results point out that organic phosphorus uptake is not proportional to alkaline phosphatase activity, demonstrating that the latter is not a suitable proxy to estimate DOP-based growth yields of organisms, whether in culture experiments or in the natural environment. The growth parameters obtained, as a function of the P source, will be critical to improve and calibrate mathematical models of diazotrophic growth and the distribution of nitrogen fixation in the global ocean.

Keywords: *Crocosphaera watsonii*, marine, oligotrophic, dissolved organic phosphorus, phosphonates, nitrogen fixation, alkaline phosphatase

INTRODUCTION

The most important source of P to the marine environment is continental weathering, through riverine transport of particulate and dissolved phases to the coastal ocean (Delaney, 1988), to which anthropogenic sources (including sewage, livestock, or paper manufacturing) add up substantial amounts (Carpenter et al., 1998). Unlike N, P does not have a significant atmospheric gas reservoir. In the open ocean, where the impact of coastal sources is much reduced, P is supplied by aeolian dust deposition like iron (Migon et al., 2001; Mills et al., 2004) and vertical mixing from deeper layers (Galloway et al., 1996). P is thus often much more abundant

in coastal areas than in the open ocean. In oligotrophic oceans, phosphate availability can be as low as a few nanomoles per liter (Wu et al., 2000; Cavender-Bares et al., 2001). Particularly low concentrations are found in the surface waters of the North Pacific and North Atlantic Oceans (Karl et al., 1997) and the Mediterranean Sea (Djaoudi et al., 2018). Brand (1991) conducted a comparative analysis of dissolved inorganic nutrient ratios and pointed out a five times higher Fe: PO_4^{3-} dissolved nutrient ratio in the Atlantic compared to the Pacific, consequent to geochemical fractionation between ocean basins (see Table 1 in Brand, 1991). The limiting character of such low concentrations imply that the dissolved, inorganic phosphorus (DIP) would ultimately control primary production in the most oligotrophic areas of the world ocean, and in particular in the Atlantic (Wu et al., 2000; Ammerman et al., 2003). This hypothesis is also further supported by analyses of dust deposition events in the Northern Atlantic (Moore et al., 2009). The general understanding of the marine P cycle is turning a corner, though, as recent advances brought to the fore a far more complex coupling between the cycles of P, N and metals than previously thought (Duhamel et al., 2021).

N$_2$ fixation is the major entryway for new nitrogen in the open ocean and supports up to half of the new primary production (Karl et al., 1997; Carpenter et al., 2004; Sohm et al., 2011; Zehr, 2011), compensating for losses due to denitrification and anammox (Gruber, 2008). Through their activity of both nitrogen and carbon fixation, diazotrophic cyanobacteria represent a nitrogen source to both the dissolved and particulate pools in the marine environment. Unicellular, nitrogen-fixing cyanobacteria (UCYN) thrive in open oceans and substantially contribute to the oceanic nitrogen budget (Zehr et al., 2001; Montoya et al., 2004; Zehr and Capone, 2020). Autotrophic UCYN are largely confined to the tropics and subtropics, and commonly found in the most oligotrophic areas of the world ocean. The genus *Crocosphaera* is the only cultivated representative of the open-ocean, photo-autotrophic UCYNs and as such, represents a model organism to study the potential for nitrogen fixation in oligotrophic areas.

The ability of N$_2$-fixers to grow on N$_2$ explains their competitiveness in N-depleted regions, while their sensitivity to the phosphorus source is much less understood. Given that diazotrophs are usually poorer competitors than non-diazotrophic phytoplankton (Agawin et al., 2007) they should be out-competed in environments in which the DIP resource is scarce, such as the Northern Atlantic in particular. And yet, very contradicting observations prove that significant diazotrophic activities recurrently occur in the North Atlantic (Luo et al., 2012). A still unexplored route is to consider that diazotrophs may have access to another source of phosphorus to alleviate DIP depletion. Considerations of nutrients balances and the control of N$_2$ fixation by the N:P ratio may, in the literature, have interchangeably used the terms phosphorus and phosphate, implicitly assuming that DIP was the unique source of phosphorus for diazotrophic growth. Phytoplanktonic organisms primarily use inorganic phosphorus (phosphate, PO_4^{3-}), which only accounts for a small fraction of the total, dissolved phosphorus found in oligotrophic oceans. But when facing low PO_4^{3-} availability,

marine microorganisms may utilize enzymes, such as alkaline phosphatase, to acquire PO_4^{3-} from a wide diversity of P-containing, organic compounds in seawater, which together comprise the pool of dissolved organic phosphorus (DOP). DOP is composed of the following major P bond classes: P-esters [P-O-C bonds, including phosphomono-(P-O-C) and diesters (C-O-P-O-C); 80%–85%], polyphosphates (P-O-P; 8%–13%), and phosphonates (C-P; 5%–10%; Young and Ingall, 2010). Because DOP represents the largest pool of P in the oligotrophic surface ocean–typically ~80% of the total dissolved pool (Karl and Björkman, 2015), an active DOP hydrolysis in DIP limited environments would thus contribute to a local phosphorus supply for primary production. The exploitation of DOP by diazotrophs and other picophytoplankton is thus discussed as a possible explanation for their success in oligotrophic environments, especially those with limited PO_4^{3-} availability (e.g., McLaughlin et al., 2013; Karl, 2014; Benavides and Voss, 2015). It was therefore hypothesized that DOP could participate in the control of N$_2$ fixation, the marine fixed N inventory, and primary productivity (Landolfi et al., 2015, 2018; Letscher and Moore, 2015; Somes and Oschlies, 2015). But unless associated with molecular analyses that reveal which organisms are actually producing these enzymes, the *in situ* observation of such activity cannot reveal which planktonic group is responsible for it. It is also still unclear to what extent it supports primary production and diazotrophic growth. So, in comparison to growth on PO_4^{3-}, the actual ability of diazotrophs to use DOP is still poorly understood and described. The resultant efficiency in their growth, photosynthesis and diazotrophy, as well as the fate of N and P into different cellular biochemicals is thus not yet quantified.

Cyanobacteria possess genes coding for alkaline phosphatases, suggesting an ability to use some organic P-containing molecules. A *phn* gene cluster that encodes for phosphonate transport and hydrolysis proteins is for instance found in *Trichodesmium* (Dyhrman et al., 2006) and *Nodularia* (Teikari et al., 2018) and induced upon PO_4^{3-} stress. *Crocosphaera* strains discriminate from other cyanobacteria, as they do not seem to possess the genetic baggage to hydrolyze phosphonates (Dyhrman et al., 2006). However, besides their high affinity phosphate transporters, they also present genes coding for a putative alkaline phosphatase and have the capacity to hydrolyze phosphomonoesters (Dyhrman and Haley, 2006). Some of these genes remain putative and their function has yet to be demonstrated, as well as the actual efficiency of the related DOP scavenging activity. In the present study, we focus on the latter. In order to provide a better quantification of diazotrophic growth on organic P sources when the inorganic, PO_4 substrate becomes depleted, we investigated the growth efficiency of *Crocosphaera watsonii* on different inorganic and organic P compounds. We thus compared growth rates, but also the cellular C, N, and P stoichiometry as well as metabolic rates related to the C and N metabolism. We expected that growing on an organic source of phosphorus would require additional energy (and nitrogen) expenditure to build up a pool of active alkaline phosphatase sufficient to support phosphorus acquisition. If the acquisition of organic phosphorus is not as efficient as that of phosphate,

indirect costs will add up to the unfavorable energetic budget, such as the impact of phosphorus limitation on cell functions like photosynthesis or the buildup of phosphorus containing molecules. We therefore compared metabolic activities of cultures provided either with inorganic phosphorus (phosphate) or an organic phosphorus source. For phosphorus to be the unique environmental variable to differ between treatments, and therefore prevent any bias in the results, we closely controlled the light conditions within the cultures, operating manual dilutions to keep all cultures at similar abundances, in order to ascertain that all treatments would experience the same irradiance and therefore would be provided with the same amount of light energy.

MATERIALS AND METHODS

Culture Conditions

Experiments were carried out on unialgal cultures of the diazotrophic, marine cyanobacterium *C. watsonii* strain WH8501. Bacterial contamination was checked using flow cytometry. The contamination level remained below 2%. A culture medium devoid of phosphorus was prepared under sterile conditions, using natural, autoclaved seawater. This seawater was sampled at the Point B station (43° 41′ 10″ N and 7° 19′ 00″ E), first filtered onto a 1-μm filter, and aged for at least 6 weeks. Before medium preparation, the water was filtered again through a 0.1-μm filter and autoclaved (120°C, 30 min). All macronutrients but phosphate, as well as vitamins and trace metals were added following the proportions defined by Chen et al. (1996). This common base was always prepared in one, large Nalgene bottle and once ready, was split into four different Schott bottles: a first medium was prepared with phosphate (KH_2PO_4) and used for control cultures grown in phosphate (thereafter PO_4 cultures); in the three other media, an organic phosphorus source was added instead of phosphate, such that final concentrations of P in all culture media were all 50 μmol P L^{-1}. The chosen organic sources were DL-α-glycerophosphate (thereafter Alpha cultures), adenosine mono phosphate (thereafter AMP cultures) and phytic acid.

One liter, triplicate cultures of each treatment were prepared and equipped with tubing mounted through the lid to allow for sampling and air bubbling. The sampling tubing was closed with a clamp to prevent any back-flow into the culture. The air tubing was capped with a 0.2-μm air filter to prevent contamination through the air intake. Cultures were gently bubbled and maintained homogeneous using a magnetic stirrer, in addition to manual, stirring operation prior to each sampling. Cultures were maintained at a temperature of 27°C, previously identified as being optimal for growth (Webb et al., 2009) and exposed to an incident irradiance of 200 μE m^{-2} s^{-1} using fluorescent tubes (white light). The light:dark cycle was set to a square, 12 h light and 12 h dark regime.

Cultures were initially inoculated in their respective medium (i.e., containing either DIP or DOP) using the same seed culture for all treatments, and then operated as a fed-batch mode: each culture flask was refreshed periodically in their respective medium, depending on the increase in population density, to insure (i) the maintenance of cells in an exponential state of growth, (ii) prevent them from reaching too high a cell density, which would have affected the actual irradiance perceived within the cultures, and (iii) maintain similar densities in all the cultures and therefore insure that all the different treatments experience the same irradiance, which is a crucial point of this experiment. The growth rate in between dilution phases was derived from the cell abundance monitoring. Cultures were pre-acclimated to their respective conditions for over 45 days, to ensure that metabolic rates and physiological properties were not the remnant of their previous dynamics on phosphate. Cultures were monitored for cell counts on a daily basis and sampled for biochemical parameters during the exponential phase.

The standard parameters recorded on the cultures included cell numbers, cellular contents in carbon, nitrogen, and phosphorus, soluble reactive phosphorus (SRP), total dissolved phosphorus (TDP), particulate organic phosphorus (POP), dissolved organic nitrogen (DON), ammonium (NH_4), nitrite (NO_2), and nitrate (NO_3). We also monitored alkaline phosphatase activity and the diel dynamics of photosynthetic efficiency and nitrogen fixation activity.

Cell counts were performed daily using a Beckman Multisizer Coulter Counter, on samples diluted 100 times with filtered seawater. The average cell size in the population was derived from the size spectrum provided by the Coulter upon each count of the population.

The following convention was used regarding time: the day of inoculation at the onset of the light cycle was taken as the time zero for each experiment. Therefore, experimental days are synchronized with the light cycle; e.g., day 15.0 is the 15th day into the experiment at the onset of the light and day 15.5 is 12 h into the 15th day, which is the light to dark transition.

Biochemical Analyses

Particulate Organic Phosphorus and Dissolved Phosphorus (SRP and TDP)

Volumes of 22 ml were sampled daily on each culture and filtered on combusted (4.5 h at 450°C) GFF filters. Filters were used to quantify POP and filtrates were used for the analysis of dissolved phosphorus. Samples were processed by combustion, according to Strickland and Parsons (1972). Filtrates were divided in two 11 ml subsamples and stored at −80°C until analysis for SRP and TDP. Their concentrations were determined by spectrophotometry according to Murphy and Riley (1962). A calibration curve was prepared from a 1,000 nmol·l^{-1} solution of KH_2PO_4.

Particulate Organic Carbon (POC) and Nitrogen (PON)

Volumes of 15 ml were sampled daily on each culture and filtered on combusted (4.5 h at 450°C) GFC filters. Filters were used for the analysis of the particulate material and filtrates were used to quantify the dissolved fraction. POC and PON were analyzed using a Flash Elemental Analyzer (Thermo Fisher Scientific), according to the method described in Moran et al. (1999), with the standard acetanilide nicotinamide. Cell contents

were estimated using the cell abundance measured at the time of sampling.

Dissolved Organic and Inorganic (NO₃⁻, NO₂⁻, and NH₄⁺) Forms of Nitrogen

Filtrates from POC-PON samples were filtered a second time on GFF filters to respect the usual convention of a 0.7-μm porosity to discriminate between particulate and dissolved fractions, and then placed in pre-acidified Falcon tubes and stored at −80°C until analysis. Filtrates from replicates #1 were used to quantify DON; they were processed according to Pujo-pay and Raimbault (1994) and analyzed using a Technicon autoanalyzer. Filtrates from replicates #2 were used to quantify NO_3^- and NO_2^-. Samples were measured by colorimetry according to Raimbault et al. (1990), using a Technicon autoanalyzer. For all these dissolved fractions, standard solutions of NO_3^- and NO_2^- at 10 μmol·l⁻¹ and blanks were prepared using a 38 g·l⁻¹ NaCl solution, matching the salinity of the culture medium. Filtrates from replicates #3 were used to quantify NH_4^+ and measured by fluorimetry according to Holmes et al. (1999), using an Alliance analyzer (FUTURA). A calibration curve with concentrations ranging from 0 to 5,000 nmol l⁻¹ was prepared in a 35 g·l⁻¹ solution of NaCl.

Nitrogenase Activity

We measured nitrogenase activity in a preliminary experiment carried out on turbidostats cultures maintained in similar conditions as the present experiments. Nitrogenase activity was recorded following the method described in Staal et al. (2001) and using the same setup as in Dron et al. (2012). Fourty milliliter culture from one culture was sampled at the end of the light phase and filtered onto a GFC filter. The filter was placed into an air-tight incubation chamber, overlying a layer of culture medium to maintain the filter fully moist. This chamber had a temperature controlled at 27°C and a flux of $N_2/O_2/C_2H_2$ gas mixture (70%/20%/10%). The nitrogen and oxygen gas tanks used contained 400 ppm CO_2. The gas flow rate was controlled using Brooks mass flow controllers and set to 1 L·h⁻¹. The GC operated one acquisition every 10 min. Ethylene production rates were converted into nitrogen fixation rates (fmol N_2 fixed cell⁻¹ day⁻¹) considering a conversion factor of four to account for the stoichiometric difference between the reduction of acetylene and that of dinitrogen (Capone, 1993; Staal et al., 2001).

Respiration Rates

Fifteen milliliter culture from one culture was sampled and transferred to an air-tight incubation beaker equipped with a Fibox (PreSens) sensor spot, removing any headspace to prevent exchanges between the liquid and gas phases. A Fibox® 4 fiberoptics was connected to the beaker and this setup was immersed in the dark in a temperature-controlled water bath at 27°C. An automated monitoring of oxygen concentrations was immediately started. Acquisitions were done every 5 s from 20 to 30 min. Due to the availability

of a single incubation beaker, oxygen measurements were repeated in time, every 20 to 30 min, over the entire dark phase.

Photosynthetic Efficiency

Photosynthetic activity was monitored over a day at 1 h intervals using a Multicolor PAM (Walz, Effeltrich, Germany). The maximum quantum efficiency of photosystem II (PSII) charge separation (Fv/Fm) was measured without dark adaptation to prevent state transitions from occurring, following the protocol described in Rabouille and Claquin (2016). The sample was excited by a weak blue light (1 μmol photons m⁻² s⁻¹, 470 nm, frequency 0.6 kHz) to record minimum fluorescence (F_0). Maximum fluorescence (F_m) was measured following a saturating light pulse (0.6 s, 1,700 μmol photons m⁻² s⁻¹, 470 nm). Light-response (PI) curves were recorded on each sample. These were exposed to 20 irradiances (I) for 1 min each, covering a range from 0 up to about 1,900 μmol photons m⁻² s⁻¹. The equations proposed by Genty et al. (1989) were used to calculate the effective quantum efficiency of PSII in the dark (F_v/F_m) and at each irradiance ($F_v^{'}/F_m^{'}$). These fluorescence parameters were used to derive the relative electron transport rate (rETR, arbitrary unit) at each irradiance:

$$rETR = F_v^{'}/F_m^{'} \cdot I$$

Results were fitted using the mathematical formulation proposed by Bernard and Rémond (2012) to derive the optimal rETR value (rETR_opt) observed in each PI curve:

$$rETR(I) = rETR_{max} \cdot \cfrac{I}{I + \cfrac{rETR_{max}}{\alpha} \cdot \left(\cfrac{I}{I_{opt}} - 1\right)^2}$$

Alkaline Phosphatase Activity

Alkaline phosphatase activity was monitored according to Jauzein et al. (2010) and Duhamel et al. (2011). Five milliliter samples were taken from each culture replicate; 3,920 μl were transferred into a test tube, to which 80 μl of a 500 μmol·l⁻¹ MUF-P reagent were added, to obtain a final volume of 4 ml. Fluorescence was measured using a fluorimeter (CARY Elipse Varian) with a 355-nm excitation wavelength and a 453 nm emission wavelength. The instrument blank was calibrated using milli-Q water. Blank measurements were obtained by reading the samples before adding the MUF-P reagent. A first reading was performed immediately after adding the MUF-P (t_0), then at $t_0 + 30$ min, $t_0 + 1$ h, and $t_0 + 3$ h. The standard curve was obtained using a MUF solution with concentrations ranging from 0 to 1,000 μmol·l⁻¹.

Statistical Analyses

In the following, we present measured cell sizes, rates, and concentrations, and we compare the average values obtained among the three treatments. The statistical significance was tested using an ANOVA test for equal means with a

post-hoc, Tukey test run over the three data sets; we test the rejection of the null hypothesis at the significance level of 5%.
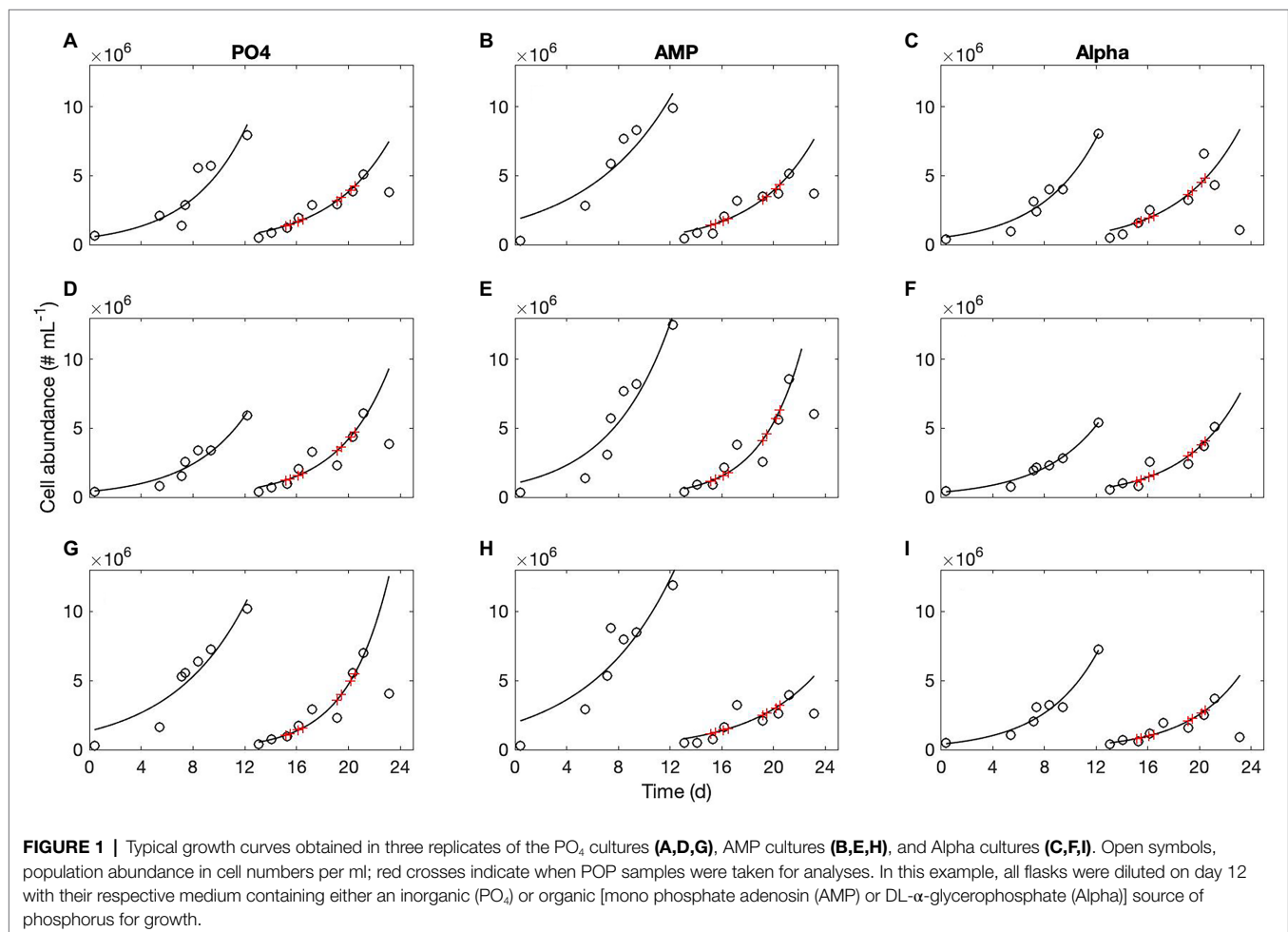
RESULTS

We observed the growth rate and biomass composition of cultured populations of *C. watsonii* grown in media that only differ by their P source. Each triplicate culture was provided with one of the following: phosphate (taken as a reference) or DL-α-glycerophosphate or adenosine mono phosphate or phytic acid.

Concentrations of biochemical species are provided with standard deviations; note that these do not correspond to analytical triplicates but to the measurements taken in the three culture replicates, which explains why variability may be large. A very small difference in the initial concentrations of cells, for instance, may lead to notable discrepancies in the final population abundance and so in the bulk elemental concentrations. This variability does not affect the population dynamics and growth rates as the growing cultures remain non limited throughout the sampling period. We recorded $n = 9$

growth curves on the PO and AMP cultures, and 11 on the Alpha cultures. An example of population monitoring is shown in **Figure 1**. The corresponding cell sizes are shown in the **Supplementary Material**.

Population Dynamics

Population dynamics was monitored throughout the entire experiment and the exponential growth rates determined over several, consecutive growth phases, in order to obtain several estimations of the growth rate, replicated in time. The mean exponential growth rate estimated in the PO$_4$ culture is $\mu_{PO_4} = 0.23 \pm 0.05$ day^{-1} ($n = 9$), with a maximum of 0.31 ± 0.06 day^{-1}. The mean, exponential growth rate estimated in the AMP culture is $\mu_{AMP} = 0.21 \pm 0.05$ day^{-1} ($n = 9$) with a maximum of 0.32 ± 0.08 day^{-1}. The mean exponential growth rate estimated in the Alpha culture is $\mu_{Alpha} = 0.22 \pm 0.03$ day^{-1} ($n = 11$) with an observed maximum of 0.23 ± 0.007 day^{-1}. Cultures grown in phytic acid survived, showing a roughly constant cell abundance over time (it was not possible to fit a growth curve), indicating that the gross growth rate is slightly positive and just sufficient to compensate for the population mortality, leading to a zero net growth rate. The slow growth in the culture did not generate sufficient biomass density for biochemical

FIGURE 1 | Typical growth curves obtained in three replicates of the PO$_4$ cultures **(A,D,G)**, AMP cultures **(B,E,H)**, and Alpha cultures **(C,F,I)**. Open symbols, population abundance in cell numbers per ml; red crosses indicate when POP samples were taken for analyses. In this example, all flasks were diluted on day 12 with their respective medium containing either an inorganic (PO$_4$) or organic [mono phosphate adenosin (AMP) or DL-α-glycerophosphate (Alpha)] source of phosphorus for growth.

analyses. Therefore, in the following, only results for the PO_4, AMP, and Alpha cultures are shown. The exponential growth rates observed in these three treatments do not significantly differ from each other at the significance level of 5% ($p = 0.443$). That is, while *C. watsonii* barely survives on phytic acid, it grows equally well on two other sources of organic phosphorus (AMP and DL-α-glycerophosphate) compared to phosphate.

TABLE 1 | Average cell sizes and their respective standard deviation measured in exponentially growing cultures of *Crocosphaera watsonii* provided with either an inorganic (KH_2PO_4) or organic [mono phosphate adenosin (AMP), DL-α-glycerophosphate (Alpha) or phytic acid] source of phosphorus for growth.

	Cell size (µm ± sd)
PO_4	3.37 ± 0.10 ($n = 45$)
AMP	3.46 ± 0.13 ($n = 46$)
Alpha	3.45 ± 0.11 ($n = 43$)
Phytic acid	3.83 ± 0.55

The value in between bracket is the number of samples used to estimate this average.

TABLE 2 | Mean N (fmol N cell^{-1}) and C (fmol C cell^{-1}) content in cultures of *Crocosphaera watsonii* provided with either an inorganic [phosphate (PO_4)] or organic [mono phosphate adenosin (AMP), DL-α-glycerophosphate (Alpha)] source of phosphorus for growth.

Culture	PO_4	AMP	Alpha
Mean N content (# data points)			
Early light phase	81.2 ± 18.2 (18)	81.0 ± 34.6 (17)	67.4 ± 27.0 (18)
Late light phase	55.8 ± 24.2 (18)	69.8 ± 30.9 (16)	51.7 ± 24.2 (18)
Mean ΔN incorporated per night	25.4 ± 33.1 (18)	12.2 ± 42.3 (16)	15.8 ± 34.2 (18)
Mean C content (# data points)			
Early light phase	463.6 ± 163.4 (18)	497.8 ± 181.7 (17)	440.7 ± 137.8 (18)
Late light phase	499.1 ± 145.0 (16)	646.1 ± 317.7 (17)	529.0 ± 269.2 (18)
Mean ΔC incorporated per day	56.4 ± 206.0 (16)	148.3 ± 325.0 (17)	88.31 ± 240.24 (18)

N and C contents were measured in the early and late light phase. The value in between bracket is the number of samples used to estimate this average.

TABLE 3 | Mean N (fmol N µm^{-3}) and C (fmol C µm^{-3}) content in cultures of *Crocosphaera watsonii* provided with either an inorganic [phosphate (PO_4)] or organic [mono phosphate adenosin (AMP), DL-α-glycerophosphate (Alpha)] source of phosphorus for growth. N and C contents were measured in the early and late light phase.

| Culture	PO_4	AMP	Alpha
Mean N content (# data points)			
Early light phase	4.2 ± 0.9 (18)	4.1 ± 1.7 (17)	3.3 ± 1.3 (18)
Late light phase	2.9 ± 1.3 (18)	3.5 ± 1.6 (16)	2.5 ± 1.2 (18)
Mean ΔN incorporated per night	1.3 ± 1.7 (18)	0.6 ± 2.1 (16)	0.8 ± 1.7 (18)
Mean C content (# data points)			
Early light phase	24.0 ± 8.5 (18)	25.0 ± 9.1 (17)	21.4 ± 6.7 (18)
Late light phase	25.8 ± 7.5 (16)	32.5 ± 16.0 (17)	25.7 ± 13.1 (18)
Mean ΔC incorporated per day	2.9 ± 10.7 (16)	7.5 ± 16.3 (17)	4.3 ± 11.7 (18)

The average cell size measured during the experiments is presented in **Table 1**. Cultures grown on AMP and Alpha are significantly larger than the PO_4 cultures ($p = 8.55 \cdot 10^{-4}$).

Cellular Pools

Particulate Organic Carbon and Nitrogen Content

The particulate carbon and nitrogen contents in the cell biomass are shown in **Table 2**. The overall N content per cell is higher in the AMP cultures, with statistical significance ($p = 0.0035$), while the C content presents no significant difference between cultures. As samples were taken in the beginning and at the end of the light phase, an absence of statistical significance may be related to the wide fluctuations in cell stoichiometry observed in this strain over the diel cycle (Dron et al., 2012), leading to large deviations around the mean. We distinguished the C and N contents between these two sampling times. We find neither statistical difference at the 5% level in the N or C content among the samples taken in the early light phase, nor among those taken in the late-light phase. The N and C contents in the early (resp. the late) light phase are thus statistically similar. We also compared the difference between the amounts incorporated in cells during the light cycle (**Table 2**). The average amount of nitrogen (mean ΔN) acquired over the dark phase appears larger in the PO_4 culture, while the average amount of carbon (mean ΔC) incorporated during the light phase is larger in the AMP culture. However, due to the rather large standard deviations, the differences between the ΔN or the ΔC are not statistically significant. We also estimated the C and N contents per unit of cell volume (fmol C µm^{-3} and fmol N µm^{-3}, shown in **Table 3**) to avoid a possible bias due to differing cell sizes. A comparison of the volumetric contents still returns that differences between the ΔN or the ΔC expressed per volume unit are not statistically significant.

The average C:N ratio in all treatments presents a typical fluctuation (Gallon, 1992; Sherman et al., 1998; Dron et al., 2012, 2013) with lower values in the early-light (at the end of the N_2 fixation period) and higher values in the late-light phase (**Figure 2**). The PO_4 culture observed fluctuations with
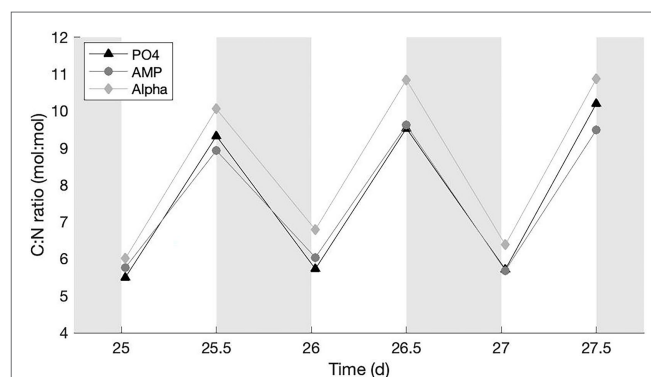
FIGURE 2 | C:N ratios in cultures provided with either an inorganic (PO_4; triangles) or organic [mono phosphate adenosin (AMP; circles) or DL-α-glycerophosphate (Alpha; diamonds)] source of phosphorus for growth. The grey, shaded area represents the dark phase.

TABLE 4 | Particulate phosphorus content normalized per cell, measured in exponentially growing cultures of *Crocosphaera watsonii* provided with either an inorganic (PO_4) or organic [mono phosphate adenosin (AMP) or DL-α-glycerophosphate (Alpha)] source of phosphorus for growth.

	PO_4	AMP	Alpha
Average cell abundance (10^6 cells ml^{-1}) over the sampling period	2.81	2.74	2.86
Average P content (all samples pooled; fmol P cell^{-1})	7.17 ± 1.33 ($n =21$)	9.43 ± 2.37 ($n =21$)	5.98 ± 2.22 ($n =21$)
Average P content in the early light phase (fmol P cell^{-1}; 1st morning is an outlier)	8.25 ± 1.14 ($n =9$)	10.23 ± 2.35 ($n =9$)	6.94 ± 2.22 ($n =9$)
Average P content in the late light phase (fmol P cell^{-1})	6.37 ± 0.77 ($n =12$)	8.84 ± 2.30 ($n =12$)	5.27 ± 2.02 ($n =12$)
P decrease in the light (fmol P cell^{-1})	2.03 ± 0.82	2.10 ± 1.72	1.96 ± 1.22
P decrease in the light (% of the early light phase quota)	24.62%	20.53%	28.24%
P increase in the dark (fmol P cell^{-1})	1.88 ± 1.53	1.25 ± 2.00	1.74 ± 1.43
P increase in the dark (% of the late light phase quota)	29.51%	14.14%	33.02%

TABLE 5 | Particulate phosphorus normalized per cell volume, measured in exponentially growing cultures of *Crocosphaera watsonii* provided with either an inorganic (PO_4) or organic [mono phosphate adenosin (AMP) or DL-α-glycerophosphate (Alpha)] source of phosphorus for growth.

	PO_4 (C2)	AMP (C1)	Alpha (C3)
Average cell volume (um^3) over the sampling period	20.04 ± 0.001	21.69 ± 0.001	21.50 ± 0.001
Average P content (all samples pooled; fmol P um^{-3})	0.046 ± 0.010 ($n =24$)	0.055 ± 0.015 ($n =21$)	0.034 ± 0.014 ($n =21$)
Average P content in the early light phase (fmol P um^{-3})	0.0621 ± 0.0180 ($n =12$)	0.0599 ± 0.0147 ($n =9$)	0.0530 ± 0.0244 ($n =12$)
Average P content in the late light phase (fmol P um^{-3})	0.0404 ± 0.0060 ($n =12$)	0.0507 ± 0.0144 ($n =12$)	0.0285 ± 0.0122 ($n =12$)
P decrease in the light (fmol P um^{-3})	0.0217	0.0092	0.0245
P decrease in the light (% of the early light phase quota)	34.94%	15.36%	46.22%
P increase in the dark (fmol P um^{-3})	0.0118 ± 0.0112	0.0077 ± 0.0159	0.0112 ± 0.0104
P increase in the dark (% of the late light phase quota)	29.2%	15.2%	39.3%

an amplitude that is intermediate between the AMP culture (the lowest amplitude) and Alpha (the largest amplitude). All values are above the Redfield C:N ratio, instead of fluctuating around it. We find no statistical difference among the early light, average C:N ratios in the different treatments. Among the late light samples, the average C:N ratio is significantly higher in the Alpha culture than in the AMP; but the C:N ratio in neither culture grown on DOP significantly differs from that in the reference PO_4 culture.

Particulate Organic Phosphorus Content

Since we observed some differences in cell size between treatments, we normalized the cellular phosphorus content both per cell (**Table 4**) and per unit of cell volume (**Table 5**). When considering all samples, the average POP content is 7.17 ± 1.33, 9.43 ± 2.37, and 5.98 ± 2.22 fmol P cell^{-1} in the PO_4, AMP, and Alpha cultures, respectively. Normalizing per unit of cell volume, the POP content is 0.046 ± 0.010, 0.055 ± 0.015, and 0.034 ± 0.014 fmol P μm^{-3} in the PO_4, AMP, and Alpha cultures, respectively. That is, the volumetric phosphorus content in the AMP and Alpha cultures is 120% and 73% of the average volumetric content in the PO_4 cultures. The reference culture thus shows an intermediate phosphorus content compared to the two cultures grown on DOP. The per cell POP content in the AMP culture is significantly higher than the other two at the 5% level, while it is similar in the PO_4 and Alpha cultures. But it can be expected that being bigger, cells in the AMP cultures contain more material. When considering the volumetric contents, cells in the AMP and PO_4 cultures then show similar POP contents while that in Alpha is statistically lower from both AMP and PO_4.

POP contents also show a diel trend with higher values at the onset of the light and lower at the onset of the dark (**Tables 4** and **5**) and this fluctuation may affect the statistical comparison of the different treatments. Within the early light samples, the POP contents expressed per cell are similar in the phosphate and Alpha cultures while the AMP culture differs from the other two at the 5% level. The same differences are observed within the late light samples. When normalizing per volume of cell, none of the POP contents in the early light samples differs from the others; and in late light samples AMP and PO_4 remain similar while the Alpha culture shows lower contents compared to the other two. The differences observed on the daily averaged values are thus also found when separately comparing the early light and early dark samples. We then estimated the POP decrease in the light phase and the POP increase in the dark phase. Note that the latter avoids the bias that cell division introduces during the light phase. Overall, the AMP culture shows the highest P content and the smallest relative increase in the dark, while the Alpha culture shows the lowest P content but a proportionately larger relative increase in the dark (**Tables 4, 5**).

Alkaline Phosphatase Activity

Alkaline phosphatase activity is normalized using cell abundance and the duration of the incubation, and averaged from the triplicate samples. It is thus expressed in pmol (P hydrolyzed) cell^{-1} h^{-1} (**Table 6A**) and shows no significant difference between the treatments. An important fact is that activities were always recorded at the same time on all cultures, but over different

TABLE 6 | Alkaline phosphatase activity measured in exponentially growing cultures of *Crocosphaera watsonii* provided with either an inorganic (PO$_4$) or organic [mono phosphate adenosin (AMP) or DL-α-glycerophosphate (Alpha)] source of phosphorus for growth.

Incub. time (h)	PO$_4$	AMP	Alpha
A	**pmol (P hydrolyzed) cell^{-1} h^{-1}**		
0.5	0	0.066	0.406
0.5	0	0.070	0.263
0.5	0.004	0.044	0.077
1	0.004	0.102	0.505
1	0.010	0.051	0.111
3	0.005	0.096	0.347
3	0.015	0.054	0.123
B	**Relative activity**		
0.5	0	16.26	100
0.5	0	26.62	100
0.5	5.19	57.14	100
1	0.81	19.68	100
1	9.00	45.95	100
3	1.44	27.67	100
3	12.20	52.85	100

(A) *Activity per cell [pmol (P hydrolyzed) cell^{-1} h^{-1}].* **(B)** *The same records but expressed in proportion to the activity measured in the Alpha culture. Note that the PO$_4$ culture is not taken as reference here because of the two null values recorded; so we instead normalized to the culture with the highest activity.*

days and times in the light cycle. This data set indicates that the activity is not linear over a 3-h incubation: the estimated rate is not the same depending on whether we calculate it over 30 min, 1 or 3 h. Further, there also seems to be a diel trend in each culture: the recorded activities tend to be stronger in the early light phase compared to the late light phase. The alkaline phosphatase may thus not be produced monotonically over the light cycle. For all these reasons, these absolute activities cannot be pooled within each treatment. In order to remove any bias due to temporal dynamics in the light cycle or to incubation time, the same data were expressed as relative activities (**Table 6B**). Normalization is performed relatively to Alpha, which always shows the highest activity. Activities in AMP and PO$_4$ at any given time is thus expressed relatively to the activity in Alpha measured at the same moment. This allows comparing all data and in particular records from different incubation times or different sampling hours. Here, the three cultures show significantly different activities at the 5% level. Alkaline phosphatase activity in the PO$_4$ culture is very low while it is much higher in the two DOP-grown cultures. Last, activity in the Alpha culture is also more than twofold that observed in AMP.

Nutrient Concentrations in the Cultures
SRP and TDP Concentrations
We measured SRP and TDP concentrations in the cultures after 19 days of growth. The reference cultures, grown on phosphate, show the highest levels of SRP with a concentration of $23.93 \pm 4.15 \mu mol$ P L^{-1}. The lowest SRP concentration is found in the Alpha cultures ($4.71 \pm 0.90 \mu mol$ P L^{-1}) while the AMP cultures show an intermediate concentration between

the former two ($14.01 \pm 1.75 \mu mol$ P L^{-1}). Nevertheless, SRP levels provide a partial information as they correspond to the inorganic fraction and thus do not indicate how much DOP actually remains in the AMP and Alpha cultures. In the following, we thus compared the levels of the total, dissolved phosphorus in the medium, that inform on how much dissolved phosphorus is still potentially available for growth.

After 19 days of growth, the reference cultures, grown on phosphate, show a TDP concentration of $28.43 \pm 3.38 \mu mol$ P L^{-1}, whose range agrees well with the SRP level. The AMP cultures show a TDP level of $16.98 \pm 2.90 \mu mol$ P L^{-1}, which closely corresponds to the SRP level, suggesting that all the organic phosphorus in that culture was hydrolyzed and converted to phosphate by the alkaline phosphatase. In the Alpha cultures, the TDP concentration is $30.84 \pm 6.07 \mu mol$ P L^{-1}, suggesting that most of the remaining phosphorus is still organic.

Compared Dynamics of the Total Dissolved (TDP) and Particulate (POP) Phosphorus Pools
We compared the TDP dynamics to that of the phosphorus in the cell biomass (POP). **Figure 3** presents the TDP (open symbols) and POP (closed symbols) concentrations in μmol P L^{-1} measured over 5 consecutive days. In all treatments, the biomass increase is matched with a consistent decrease in the TDP in the culture medium, indicative of a clear consumption of the DOP compounds in the AMP and Alpha cultures.

Dissolved Nitrogen Concentrations
Dissolved Inorganic Nitrogen
The culture medium was prepared without any source of dissolved inorganic nitrogen (DIN). Samples were taken to verify that DIN was absent from the cultures. NH$_4^+$ levels are mostly below 2 to 4 μmol L^{-1}. Nitrate and nitrite remain mostly below 2 μmol L^{-1}. As expected, DIN levels remain negligible in proportion to the other nutrients over the course of the experiment.

Dissolved Organic Nitrogen
We found significant amounts of exuded, organic nitrogen in the culture. At the end of the sampling period (on day 27), DON levels are 61, 152, and 36 μmol N L^{-1} in the PO$_4$, AMP, and Alpha cultures, respectively. The levels of bulk nitrogen biomass formed in the cultures over the same period are 231, 299, and 227 μmol PON L^{-1}, respectively. Cells thus partition their gross nitrogen acquisition as follows: in the PO$_4$ cultures, 80% is incorporated in the biomass and 20% exuded; in the AMP cultures, 66% is incorporated in the biomass and 34% exuded; in the Alpha cultures, 86% is incorporated in the biomass and 14% exuded.

Nitrogen Fixation Rates
We recorded nitrogenase activity overnight in the three culture conditions, starting before the onset of the dark and until after the onset of the light the following day (**Figure 4**). A unique record per culture is not sufficient to quantitatively
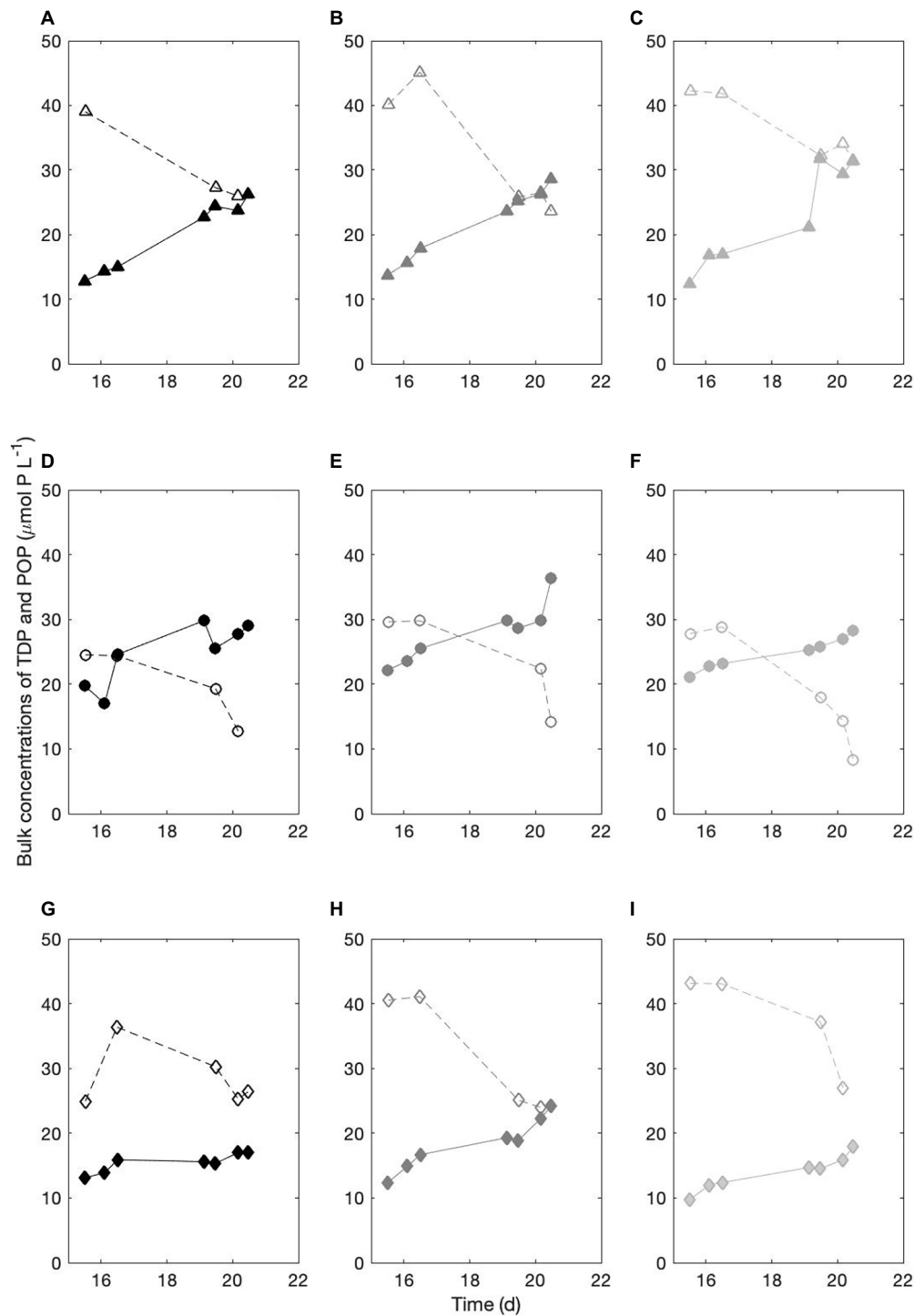
FIGURE 3 | Temporal dynamics of the bulk, total dissolved and particulate phosphorus pools. **(A-C)** three replicates of the PO_4 culture; **(D-F)** three replicates of the AMP culture; **(G-I)** three replicates of the Alpha culture. Open symbols, TDP concentrations in µmol P L^{-1}. Closed symbols, POP concentrations in µmol P L^{-1}.

compare the efficiency and yield of nitrogen fixation between the treatments; but results clearly show that nitrogenase was active in all three culture treatments with rather similar temporal dynamics and maximum efficiency. We concomitantly recorded

the respiration rate in the three treatments over the dark phase. No sensible difference appears between the treatments, which show comparable temporal dynamics in the rate of oxygen consumed (see **Supplementary Material**).
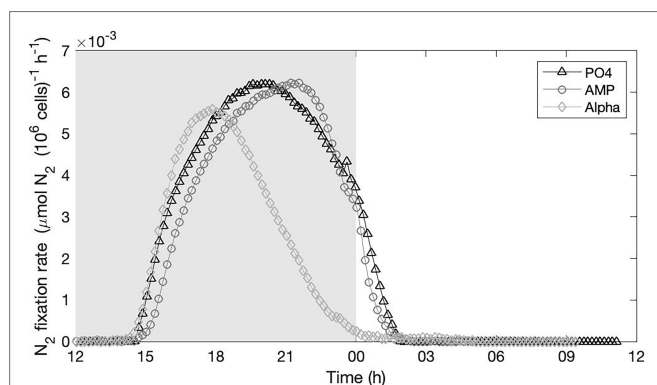
FIGURE 4 | Nitrogenase activity monitored on-line in cultures provided with either an inorganic (PO₄; triangles) or organic [mono phosphate adenosin (AMP; circles) or DL-α-glycerophosphate (Alpha; diamonds)] source of phosphorus for growth. The grey, shaded area represents the dark phase.

Photosynthetic Efficiency

The rETR expresses the instantaneous efficiency of PSII under the current incident irradiance. The changes in time of the optimal rETR ($rETR_{opt}$) observed in each treatment are drawn on **Figure 5**. A clear and very similar temporal dynamics is observed in the PO₄, AMP, and Alpha cultures, with a maximal PSII efficiency around the mid-light phase and relatively comparable maximal efficiencies. Replicated acquisitions over several days would be necessary to assess whether the slightly lower maximal efficiency in the AMP culture has any statistical significance. In contrast, the cultures grown on Phytic acid show a clearly impaired photosynthetic efficiency.

DISCUSSION

It has long been acknowledged that nitrogen fixation being the only biological entryway of new nitrogen into the open ocean, it constitutes a main driver of oceanic primary production (Capone et al., 1997; Karl et al., 1997). Yet, how much N₂ is actually fixed per year is still very uncertain due to the paucity of direct measurements and also to methodological challenges (Großkopf and LaRoche, 2012; Landolfi et al., 2018). Biogeochemical models could supplement the *in situ* measurements for more exhaustive estimations but appropriate calibration data are still missing for the actual, current contribution of nitrogen fixers to the carbon (C), nitrogen (N) and phosphorus (P) cycles to be adequately evaluated in these models. Simulating realistic marine N₂ fixation patterns is a challenge that requires experimental measurements of key processes and key parameters of matter and energy fluxes related to diazotrophic growth. In this regard, culture approaches on isolated strains can provide relevant data on specific, metabolic rates.

A major control of N₂ fixation by PO_4^{3-} availability has been suggested (Sañudo-Wilhelmy et al., 2001) and its excess relative to NO_3^- is often the driver of diazotrophic growth in global biogeochemical models (Deutsch et al., 2007; Weber

and Deutsch, 2012; Landolfi et al., 2013). The consumption of organic phosphorus is an important, yet still overlooked, feature in diazotrophic cyanobacteria that could be essential to their development in oligotrophic oceans. PO_4^{3-} stress induces alternative, organic P (DOP) scavenging strategies in diazotrophs (Mulholland et al., 2002; Dyhrman et al., 2006). DOP use may thus be a strategic adaptation to PO_4^{3-} limiting conditions common to many oligotrophic environments but is not systematic in all marine cyanobacteria and also depends on the compounds structure. Following these pioneering works, genes encoding for alkaline phosphatase were shown to be expressed in populations of both colonial (Orchard et al., 2009) and unicellular diazotrophs (Dyhrman and Haley, 2006), and up-regulated under P starvation. The assimilation of DOP is most probably costly in terms of energy and nitrogen. Therefore, the ability of diazotrophs to fix N₂ might be a necessary advantage to exploit the complex constituents of the DOP pool, bearing in mind that such activity may also increase the cellular carbon demand.

Here, we analyze the capacity of the N₂ fixer *C. watsonii* WH8501 to use dissolved organic phosphorus (DOP) as sole P source, in comparison to phosphate, over sustained multiple generations growth. The PO₄ cultures represent a reference of growth under optimal conditions. We observe whether cultures provided with DOP instead of phosphate perform equally well in terms of growth rate and whether the cellular stoichiometry is altered. We closely monitored cell abundances, which we used to identify when, in the exponential phases, samples had to be taken. The medium used in these experiments initially contained 50 μmol P L⁻¹, whether provided as an inorganic or organic form. Quantification of dissolved phosphorus in the cultures shows that, at the moment of sampling, there were from 25 to 30 μmol P L⁻¹ left in the medium, which is a reasonable indication that cells had not consumed most of the phosphorus and so were not being limited by the availability of this element, ruling out the risk for some of the results to be biased by quantitative P limitation. Also, occurrence of any limitation would have shown as a rupture in the slope of the cell abundance vs. time. **Figure 1** shows that we did sample before the growth curves taper off, indicating that all nutrients (micronutrients and vitamins, for instance) were still abundant; cultures were therefore unlimited by any nutrient at all times of sampling.

Overall Growth Efficiency as a Function of the P Substrate for Growth

Crocosphaera grows on AMP and DL-α-glycerophosphate as also reported by Dyhrman and Haley (2006) but is unable to sustain the same growth rate on phytic acid. This latter result is in contradiction with the observations of Dyhrman and Haley (2006). All our cultures transferred in a medium with phytic acid as sole source of P just survive, indicating that cells are barely able to assimilate this compound and so may use it for growth over the long term, but in a very inefficient way. Our findings indicate that *Crocosphaera* cannot use all phosphomonoesters with the same efficiency, but it can grow
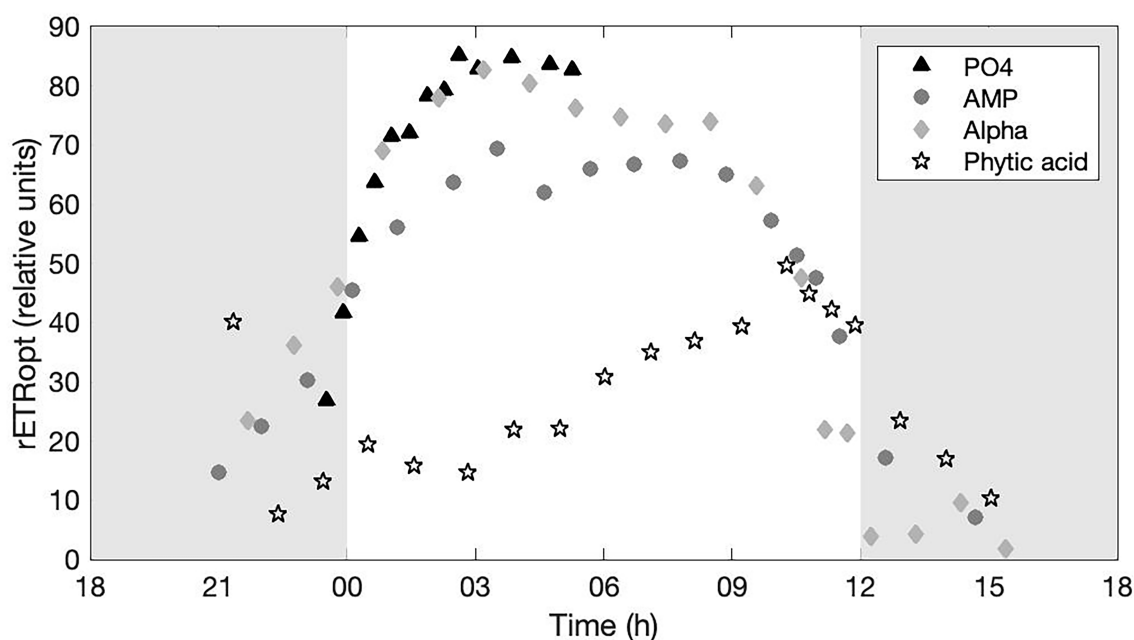
FIGURE 5 | Optimal, relative electron transport rate (rETR$_{opt}$) monitored in cultures provided with either an inorganic (PO$_4$; triangles) or organic [mono phosphate adenosin (AMP; circles), DL-α-glycerophosphate (Alpha; diamonds) or phytic acid (Phytic acid; pentagrams)] source of phosphorus for growth. The grey, shaded area represents the dark phase.

without phosphate, provided that usable DOP is available. We obtain similar growth rates on phosphate or on dissolved organic phosphorus in the form of AMP or DL-α-glycerophosphate. All three showed a division time of about 3 days so each 30 day monitoring indicates that growth was sustained for a minimum of 10 generations. **Figures 2, 3** in Dyhrman and Haley (2006) show that it takes at least 7 days of growth for a difference in population dynamics to be observed between cultures of *Crocosphaera* grown with or without phosphate. This suggests that cells grown without P may have kept growing for about a week using their internal P reserves. In their experiment, Pereira et al. (2019) followed cells for 144 h, which is 6 days. They observed induction of genes encoding high affinity transporters but did not observe the induction of genes that would allow for the synthesis of enzymes required for DOP use. Not knowing the actual magnitude of internal P reserves, we are thus unsure their experiment was carried out over a period long enough for cells to start sensing a severe limitation in their internal stock of Pi and to induce the gene regulation that would allow for DOP scavenging. So we do not rule out that cells grew on their internal reserves in their experiment.

The irradiance level applied here is saturating for this strain (Rabouille et al., 2017). When provided with sufficient light energy and non-limiting nutrients, *Crocosphaera* can thus grow equally well on some organic phosphorus as compared to phosphate. Still, we found a small, yet statistically significant difference in cell size between the treatments, with cells grown on DOP being larger than cells grown on phosphate (**Table 1**). Given that growth dynamics are similar in the three treatments,

one could expect cells to divide at about the same rate, and so to have about the same size range. A bigger cell size may result from a slower growth rate. When cells are exposed to unfavorable conditions, their resource incorporation rate slows down and cells may not be ready at the time of division; they may thus miss the time window for division and, due to the tightly controlled cell cycle, only divide the following day (e.g., Dron et al., 2013). The time between two divisions increases, but then cells also experience a longer light exposure and a longer time to fix nitrogen, which results in a higher total C and N incorporation in between two divisions (e.g., Dron et al., 2013). Cells grown on organic P, because they are slightly bigger, must have somehow incorporated more material between two division phases, but did nevertheless divide at the same pace as the reference cultures grown on phosphate. The small difference observed between our treatments is not unusual as all of them remain within the range reported by Webb et al. (2009) in cultured *C. watsonii* WH8501 (2–4 μm) and also within that reported by Wilson et al. (2017) in natural *Crocosphaera* populations (2.1 ± 0.5 and 5.0 ± 0.8 μm for the small and large populations).

How DOP Use Affects the Cellular Phosphorus Contents

Cultures grown on DOP do not present a systematic difference, in terms of C, N, and P contents, compared to the reference treatment grown on phosphate. Instead, the phosphate-grown cultures show a cell composition that is intermediate between the culture grown on AMP and that grown on DL-α-glycerophosphate. Cellular P contents are highest in the AMP

culture, which is in agreement with the fact that AMP cells were slightly bigger. But the second highest P content is found in the reference culture grown on phosphate, and not in the second, largest-sized cells, meaning that phosphorus content is not the element making the difference in volume between the reference and the Alpha culture. This result also further supports the idea that the different phosphorus-containing molecules are not cleaved and uptaken with the same efficiency.

In all cultures, the phosphorus content, when expressed per cell, shows higher values in the early light phase and lower values in the late light phase (**Tables 4** and **5**). The similar trends in all treatments, with an increase during the dark phase and a decrease in the light suggest that the temporal acquisition dynamics remains unaffected whatever the phosphorus source. The occurrence of cell division in the mid light phase in this strain (Dron et al., 2012) creates a dilution effect as, upon mitosis, the content of a mother cell redistributes into the two daughter cells. We therefore estimated the P increase per cell in the dark, when cell size should not change as much. Not only does the phosphorus composition of cells vary between treatments, but so does the daily amplitude of the incorporation per cell (in the dark): +14% in AMP vs +33% in the Alpha culture (**Table 4**). The response in the PO_4 cultures falls in between the former two (+29%; **Table 4**), suggesting that inorganic phosphate does not necessarily yield the highest biomass when cultures are provided with sufficient energy (i.e., favorable irradiance conditions). Because cell size in both the AMP and Alpha culture is larger than in the reference, PO_4 culture, the difference in P content could be due to a difference in cell size between experiments; we therefore expressed the phosphorus content per volume unit (**Table 5**). When normalized per cubic micrometer of cell, values become independent of a possible fluctuation in cell size with time, and so this avoids, in particular, the bias due to cell mitosis. This reveals an accentuated trend observed in the POP content compared to the per cell contents: The highest P content is observed in the AMP culture, with the lowest night increase (15%), and the lowest P content is observed in the Alpha culture, with the highest night increase (39%). The reference culture still shows intermediate P content and increase (29%).

Assuming the evolution of P content is monotonous during each light or dark phase, P appears to be rather incorporated at night and this daily dynamics is not modified whether P is inorganic or organic. If P is incorporated in the light, the per cell contents indicate that this incorporation is concealed behind the dilution effect due to cell division. Observation of the contents expressed per cell volume further shows that in the light, more phosphorus is lost than gained.

Does the Alkaline Phosphatase Activity Reflect the DOP Use Efficiency?

Alkaline phosphatase activity was concomitantly recorded in all cultures. Activity in the reference PO_4 culture remains very low or undetectable, only a few percent of what is observed in the Alpha culture, which shows the highest rates. The AMP culture shows a high activity although the normalized, recorded

rates are two to six times lower than in the Alpha cultures. Alpha is also the culture with the lowest P content in cells. This could suggest that the activity of the alkaline phosphatase is proportional to the internal depletion in P. Yet, despite this enhanced activity, cells do not incorporate more P. Instead, it takes twice as high an alkaline phosphatase activity on DL-α-glycerophosphate for cells to incorporate equivalent amounts of P as compared to cells grown on AMP. The most plausible hypothesis is that the cleavage efficiency is not the same on all components, being lower on DL-α-glycerophosphate compared to AMP, which points to different efficiencies in scavenging compounds within the class of phosphomonoesters. This hypothesis is also supported by the measured, three-fold lower standing stock of phosphate in the Alpha cultures, which again suggests that alkaline phosphatase is less efficient in cleaving DL-α-glycerophosphate, than it is with AMP.

If the cleavage rate in the Alpha cultures is too low to meet the growth requirements, the liberated phosphate may very likely be consumed as quickly as it is produced and its concentration should remain close to undetectable, or at least at extremely low levels. The fact that soluble, reactive phosphate is clearly detected in the medium proves that the cleavage rate in the Alpha cultures, although lower, is still sufficient to cover the growth requirements. This result also corroborates well with the observed equivalent net growth rates between all three treatments. Note that we did not have a control for abiotic hydrolysis in the Alpha treatment, which, if not null, would also contribute to some (small) P supply. Last, the higher alkaline phosphatase activity in the Alpha culture did not yield a higher growth rate either. This result shows that neither P incorporation nor the growth rate are proportional to alkaline phosphatase activity. The latter is therefore likely a proxy for internal inorganic P depletion in line with PhoB regulation (Pereira et al., 2016). Alkaline phosphatase activity is thus not a suitable proxy to estimate DOP-based growth efficiencies of organisms, whether in culture experiments or in the natural environment.

Impacts of DOP Use on the C and N Metabolism

In addition to testing the ability of *Crocosphaera* to grow on organic phosphorus only, we also aimed to tackle the question of the metabolic costs associated with the use of DOP, which may alter the cell stoichiometry. Ensuring a low contamination level in the cultures is thus crucial in many aspects. Part of the bacterial population would appear in the DOP pool and their growth would thus conceal the decrease in DOP concentration related to cyanobacterial uptake. Bacteria could also potentially consume the DOP and other nutrients, artificially enhancing the DOP decrease. The soundness of our results is further confirmed by the good correspondence between the POC and cell number increases that match the DOP decrease.

Cells perceive phosphate depletion in the medium as a stress that induces a suite of genetic regulations to first produce more of the high affinity phosphate transporters (Pereira et al., 2016) and then induce the production of the alkaline phosphatase

to scavenge available DOP. We hypothesized that this chain of genetic and metabolic reactions would represent an energy cost for the overall metabolism, which may draw on the carbon reserves. That is, if there is a higher carbon cost due to DOP use, then the PO_4 culture should show the lowest expenses and so possibly the highest POC content. But this is not what we observe here. In the present experiments, in which cultures are not limited by light, no significant difference in the C:N ratios appears between the phosphate-grown or DOP-grown cultures, meaning that using DOP did not result in imbalance on the carbon and nitrogen metabolisms. We also observe no significant difference in the absolute C or N contents in cells. These results indicate that if there indeed is a carbon cost to the use of DOP, the gross carbon acquisition efficiency was able to cover for it without visible impact on the net carbon incorporation in cells. Also, if DOP use imposes a carbon cost to cells, there could be competing carbon requirements in cells between the processes of N_2 fixation and DOP use, which could induce a lower N_2 fixation efficiency due to the redirection of carbon towards DOP scavenging. But the same conclusion applies to nitrogen: the genetic and metabolic processes required for the use of DOP were performed here without noticeable impact on the nitrogenase activity and on the net N acquisition in cells. So the overall gross carbon incorporation in all treatments was sufficient to support the carbon requirements of both processes of N_2 fixation and DOP use. Also, contrary to what has been reported in *Trichodesmium* (Orchard et al., 2009), the equivalent nitrogenase dynamics and overall growth rates observed here suggest that nitrogenase synthesis is either not down-regulated when inorganic phosphate is not available, or the influx of phosphate derived from DOP scavenging prevented this downregulation to occur. Note that although they maintain about the same growth rate as the other treatments, the Alpha cultures tend to have a lower carbon content per cell. So their net carbon incorporation in between two divisions may have been slightly lower. This could

be a hint for a higher carbon expenditure, if DL-α-glycerophosphate is not as easy to scavenge as AMP or phosphate.

Overall, the present results lead to the conclusion that *Crocosphaera* can sustainably grow on selected organic phosphorus as a sole phosphorus source at the same rates as on inorganic phosphate.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, and further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SR, SD, PC, AL, and AO designed the study. LT conducted the experiments with contributions of AT, SR, and OC. SR and LT analyzed the data. SR drafted the manuscript and all authors provided input during writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.848647/full#supplementary-material

REFERENCES

Agawin, N. S. R., Rabouille, S., Veldhuis, M. J. W., Servatius, L., Hol, S., van Overzee, H. M. J., et al. (2007). Competition and facilitation between unicellular nitrogen-fixing cyanobacteria and non-nitrogen-fixing phytoplankton species. *Limnol. Oceanogr.* 52, 2233–2248. doi: 10.4319/lo.2007.52.5.2233

Ammerman, J. W., Hood, R. R., Case, D. A., and Cotner, J. B. (2003). Phosphorus deficiency in the Atlantic: an emerging paradigm in oceanography. *Eos. Trans. AGU* 84, 165–170. doi: 10.1029/2003EO180001

Benavides, M., and Voss, M. (2015). Five decades of N_2 fixation research in the North Atlantic Ocean. *Front. Mar. Sci.* 2:40. doi: 10.3389/fmars.2015.00040

Bernard, O., and Rémond, B. (2012). Validation of a simple model accounting for light and temperature effect on microalgal growth. *Bioresour. Technol.* 123, 520–527. doi: 10.1016/j.biortech.2012.07.022

Brand, L. E. (1991). Minimum iron requirements of marine phytoplankton and the implications for the biogeochemical control of new production. *Limnol. Oceanogr.* 36, 1756–1771. doi: 10.4319/lo.1991.36.8.1756

Capone, D. (1993). "Determination of nitrogenase activity in aquatic samples using the acetylene reduction procedure," in *Handbook of Methods in Aquatic Microbial Ecology* (Boca Raton, Fla: Lewis Publishers), 621–631.

Capone, D. G., Zehr, J. P., Paerl, H. W., Bergman, B., and Carpenter, E. J. (1997). *Trichodesmium*, a globally significant marine cyanobacterium. *Science* 276, 1221–1229. doi: 10.1126/science.276.5316.1221

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., and Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* 8, 559–568. doi: 10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2

Carpenter, E. J., Subramaniam, A., and Capone, D. G. (2004). Biomass and primary productivity of the cyanobacterium *Trichodesmium* spp. in the tropical N Atlantic Ocean. *Deep-Sea Res. Pt I.* 51, 173–203. doi: 10.1016/j.dsr.2003.10.006

Cavender-Bares, K. K., Karl, D. M., and Chisholm, S. W. (2001). Nutrient gradients in the western North Atlantic Ocean: relationship to microbial community structure and comparison to patterns in the Pacific Ocean. *Deep-Sea Res. I Oceanogr. Res. Pap.* 48, 2373–2395. doi: 10.1016/S0967-0637(01)00027-9

Chen, Y. B., Zehr, J. P., and Mellon, M. (1996). Growth and nitrogen fixation of the diazotrophic filamentous non heterocystous cyanobacterium *Trichodesmium* sp IMS 101 in defined media: evidence for a circadian rhythm. *J. Phycol.* 32, 916–923. doi: 10.1111/j.0022-3646.1996.00916.x

Delaney, M. (1988). Phosphorus accumulation in marine sediments and the oceanic phosphorus cycle. *Global Biogeochem. Cycles* 12, 563–572. doi: 10.1029/98GB02263

Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N., and Dunne, J. P. (2007). Spatial coupling of nitrogen inputs and losses in the ocean. *Nature* 445, 163–167. doi: 10.1038/nature05392

Djaoudi, K., Van Wambeke, F., Coppola, L., D'Ortenzio, F., Helias-Nunige, S., Raimbault, P., et al. (2018). Sensitive determination of the dissolved phosphate Pool for an improved resolution of its vertical variability in the surface layer: new views in the P-depleted Mediterranean Sea. *Front. Mar. Sci.* 5:234. doi: 10.3389/fmars.2018.00234

Dron, A., Rabouille, S., Claquin, P., Le Roy, B., Talec, A., and Sciandra, A. (2012). Light-dark (12:12) cycle of carbon and nitrogen metabolism in *Crocosphaera watsonii* WH8501: relation to the cell cycle. *Environ. Microbiol.* 14, 967–981. doi: 10.1111/j.1462-2920.2011.02675.x

Dron, A., Rabouille, S., Claquin, P., Talec, A., Raimbault, V., and Sciandra, A. (2013). Photoperiod length paces the temporal orchestration of cell cycle and carbon-nitrogen metabolism in *Crocosphaera watsonii*. *Environ. Microbiol.* 15, 3292–3304. doi: 10.1111/1462-2920.12163

Duhamel, S., Björkman, K. M., Van Wambeke, F., Moutin, T., and Karl, D. M. (2011). Characterization of alkaline phosphatase activity in the north and South Pacific subtropical gyres: implications for phosphorus cycling. *Limnol. Oceanogr.* 56, 1244–1254. doi: 10.4319/lo.2011.56.4.1244

Duhamel, S., Diaz, J. M., Adams, J. C., Djaoudi, K., Steck, V., and Waggoner, E. M. (2021). Phosphorus as an integral component of global marine biogeochemistry. *Nat. Geosci.* 14, 359–368. doi: 10.1038/s41561-021-00755-8

Dyhrman, S. T., Chappell, P. D., Haley, S. T., Moffett, J. W., Orchard, E. D., Waterbury, J. B., et al. (2006). Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* 439, 68–71. doi: 10.1038/nature04203

Dyhrman, S. T., and Haley, S. T. (2006). Phosphorus scavenging in the unicellular marine diazotroph *Crocosphaera watsonii*. *Appl. Environ. Microbiol.* 72, 1452–1458. doi: 10.1128/AEM.72.2.1452-1458.2006

Gallon, J. R. (1992). Reconciling the incompatible-N$_2$ fixation and O$_2$. *New Phytol.* 122, 571–609. doi: 10.1111/j.1469-8137.1992.tb00087.x

Galloway, J. N., Howarth, R. W., Michaels, A. F., Nixon, S. W., Prospero, J. M., and Dentener, F. J. (1996). Nitrogen and phosphorus budgets of the North Atlantic Ocean and its watershed. *Biogeochemistry* 35, 3–25. doi: 10.1007/BF02179823

Genty, B., Briantais, J., and Baker, N. R. (1989). The relationship between the quantum yield of photosynthetic electron transport and of quenching chlorophyll fluorescence. *Biochim. Biophys. Acta* 990, 87–92. doi: 10.1016/S0304-4165(89)80016-9

Großkopf, T., and LaRoche, J. (2012). Direct and indirect costs of dinitrogen fixation in *Crocosphaera watsonii* WH8501 and possible implications for the nitrogen cycle. *Front. Microbiol.* 3:236. doi: 10.3389/fmicb.2012.00236

Gruber, N., (2008). "The marine nitrogen cycle: overview and challenges," in *Nitrogen in the Marine Environment. 2nd Edn.* eds. D. G. Capone, D. A. Bronk, M. R. Mulholland and E. J. Carpenter (San Diego, Calif: Academic Press), 1–50.

Holmes, R. M., Aminot, A., Kérouel, R., Hooker, B. A., and Peterson, B. J. (1999). A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Can. J. Fish. Aquat. Sci.* 56, 1801–1808. doi: 10.1139/f99-128

Jauzein, C., Labry, C., Youenou, A., Quéré, J., Delmas, D., and Collos, Y. (2010). Growth and phosphorus uptake by the toxic dinoflagelate *Alexandrium catenella* (Dinophyceae) in response to phosphate limitation. *J. Phycol.* 46, 926–936. doi: 10.1111/j.1529-8817.2010.00878.x

Karl, D. M. (2014). Microbially mediated transformations of phosphorus in the sea: new views of an old cycle. *Annu. Rev. Mar. Sci.* 6, 279–337. doi: 10.1146/annurev-marine-010213-135046

Karl, D. M., and Björkman, K. M. (2015). "Dynamics of dissolved organic phosphorus," in *Biogeochemistry of Marine Dissolved Organic Matter* (Burlington: Academic Press), 233–334.

Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J., and Hebel, D. (1997). The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* 388, 533–538. doi: 10.1038/41474

Landolfi, A., Dietze, H., Koeve, W., and Oschlies, A. (2013). Overlooked runaway feedback in the marine nitrogen cycle: the vicious cycle. *Biogeosciences* 10, 1351–1363. doi: 10.5194/bg-10-1351-2013

Landolfi, A., Kähler, P., Koeve, W., and Oschlies, A. (2018). Global marine N$_2$ fixation estimates: from observations to models. *Front. Microbiol.* 9:2112. doi: 10.3389/fmicb.2018.02112

Landolfi, A., Koeve, W., Dietze, H., Kähler, P., and Oschlies, A. (2015). A new perspective on environmental controls of marine nitrogen fixation. *Geophys. Res. Lett.* 42, 4482–4489. doi: 10.1002/2015GL063756

Letscher, R. T., and Moore, J. K. (2015). Preferential remineralization of dissolved organic phosphorus and non-Redfield DOM dynamics in the global ocean: impacts on marine productivity, nitrogen fixation, and carbon export. *Global Biogeochem. Cycles* 29, 325–340. doi: 10.1002/2014GB004904

Luo, Y. W., Doney, S. C., Anderson, L. A., Benavides, M., Berman-Frank, I., Bode, A., et al. (2012). Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. *Earth Syst. Sci. Data* 4, 47–73. doi: 10.5194/essd-4-47-2012

McLaughlin, K., Sohm, J. A., Cutter, G. A., Lomas, M. W., and Paytan, A. (2013). Phosphorus cycling in the Sargasso Sea: investigation using the oxygen isotopic composition of phosphate, enzyme-labeled fluorescence, and turnover times. *Glob. Biogeochem. Cycles* 27, 375–387. doi: 10.1002/gbc.20037

Migon, C., Sandroni, V., and Bethoux, J. P. (2001). Atmospheric input of anthropogenic phosphorus to the Northwest Mediterranean under oligotrophic conditions. *Mar. Environ. Res.* 52, 413–426. doi: 10.1016/S0141-1136(01)00095-2

Mills, M. M., Ridame, C., Davey, M., La Roche, J., and Geider, R. J. (2004). Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature* 429, 292–294. doi: 10.1038/nature02550

Montoya, J. P., Holl, C. M., Zehr, J. P., Hansen, A., Villareal, T. A., and Capone, D. G. (2004). High rates of N$_2$ fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. *Nature* 430, 1027–1031. doi: 10.1038/nature02824

Moore, C. M., Mills, M. M., Achterberg, E. P., Geider, R. J., LaRoche, J., Lucas, M. I., et al. (2009). Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. *Nat. Geosci.* 2, 867–871. doi: 10.1038/ngeo667

Moran, S. B., Charette, M. A., Pike, S. M., and Wicklund, C. A. (1999). Differences in seawater particulate organic carbon concentration in samples collected using small-and large-volume methods: the importance of DOC adsorption to the filter blank. *Mar. Chem.* 67, 33–42. doi: 10.1016/S0304-4203(99)00047-X

Mulholland, M. R., Floge, S., Carpenter, E. J., and Capone, D. G. (2002). Phosphorus dynamics in cultures and natural populations of *Trichodesmium* spp. *Mar. Ecol. Prog. Ser.* 239, 45–55. doi: 10.3354/meps239045

Murphy, J., and Riley, J. P. (1962). A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* 27, 31–36. doi: 10.1016/S0003-2670(00)88444-5

Orchard, E. D., Webb, E. A., and Dyhrman, S. T. (2009). Molecular analysis of the phosphorus starvation response in *Trichodesmium* spp. *Environ. Microbiol.* 11, 2400–2411. doi: 10.1111/j.1462-2920.2009.01968.x

Pereira, N., Shilova, I. N., and Zehr, J. P. (2016). Molecular markers define progressing stages of phosphorus limitation in the nitrogen-fixing cyanobacterium, *Crocosphaera*. *J. Phycol.* 52, 274–282. doi: 10.1111/jpy.12396

Pujo-pay, M., and Raimbault, P. (1994). Improvement of the wet-oxidation procedure for simultaneous determination of particulate organic nitrogen and phosphorus collected on filters. *Mar. Ecol. Prog. Ser.* 105, 203–207. doi: 10.3354/meps105203

Rabouille, S., Cabral, G. S., and Pedrotti, M. L. (2017). Towards a carbon budget of the diazotrophic cyanobacterium *Crocosphaera*: effect of irradiance. *Mar. Ecol. Prog. Ser.* 570, 29–40. doi: 10.3354/meps12087

Rabouille, S., and Claquin, P. (2016). Photosystem-II shutdown evolved with nitrogen fixation in the unicellular diazotroph *Crocosphaera watsonii*. *Environ. Microbiol.* 18, 477–485. doi: 10.1111/1462-2920.13157

Raimbault, P., Slawyk, G., Coste, B., and Fry, J. (1990). Feasibility of using an automated colorimetric procedure for the determination of seawater nitrate in the 0 to 100 nM range: examples from field and culture. *Mar. Biol.* 104, 347–351. doi: 10.1007/BF01313277

Sañudo-Wilhelmy, S. A., Kustka, A. B., Gobler, C. J., Hutchins, D. A., Yang, M., Lwiza, K., et al. (2001). Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the Central Atlantic Ocean. *Nature* 411, 66–69. doi: 10.1038/35075041

Sherman, L. A., Meunier, P., and Colon-Lopez, M. S. (1998). Diurnal rhythms in metabolism: a day in the life of a unicellular, diazotrophic cyanobacterium. *Photosynth. Res.* 58, 25–42. doi: 10.1023/A:1006137605802

Sohm, J. A., Webb, E. A., and Capone, D. G. (2011). Emerging patterns of marine nitrogen fixation. *Nat. Rev. Microbiol.* 9, 499–508. doi: 10.1038/nrmicro2594

Somes, C. J., and Oschlies, A. (2015). On the influence of "non-Redfield" dissolved organic nutrient dynamics on the spatial distribution of N$_2$ fixation and the size of the marine fixed nitrogen inventory. *Glob. Biogeochem. Cycles* 29, 973–993. doi: 10.1002/2014gb005050

Staal, M., Lintel-Hekkert, S. T., Harren, F., and Stal, L. J. (2001). Nitrogenase activity in cyanobacteria measured by the acetylene reduction assay: a comparison between batch incubation and on-line monitoring. *Environ. Microbiol.* 3, 343–351. doi: 10.1046/j.1462-2920.2001.00201.x

Strickland, J.D.H., and Parsons, T.R., (1972). *A Practical Handbook of Seawater Analysis, Bulletin (Fisheries Research Board of Canada).* Ottawa: Fisheries Research Board of Canada.

Teikari, J. E., Fewer, D. P., Shrestha, R., Hou, S., Leikoski, N., Mäkelä, M., et al. (2018). Strains of the toxic and bloom-forming *Nodularia spumigena* (cyanobacteria) can degrade methylphosphonate and release methane. *ISME J.* 12, 1619–1630. doi: 10.1038/s41396-018-0056-6

Webb, E. A., Ehrenreich, I. M., Brown, S. L., Valois, F. W., and Waterbury, J. B. (2009). Phenotypic and genotypic characterization of multiple strains of the diazotrophic cyanobacterium, *Crocosphaera watsonii*, isolated from the open ocean. *Environ. Microbiol.* 11, 338–348. doi: 10.1111/j.1462-2920.2008.01771.x

Weber, T., and Deutsch, C. (2012). Oceanic nitrogen reservoir regulated by plankton diversity and ocean circulation. *Nature* 489, 419–422. doi: 10.1038/nature11357

Wilson, S. T., Aylward, F. O., Ribalet, F., Barone, B., Casey, J. R., Connell, P. E., et al. (2017). Coordinated regulation of growth, activity and transcription in natural populations of the unicellular nitrogen-fixing cyanobacterium *Crocosphaera*. *Nat. Microbiol.* 2:17118. doi: 10.1038/nmicrobiol.2017.118

Wu, J., Sunda, W., Boyle, E. A., and Karl, D. M. (2000). Phosphate depletion in the Western North Atlantic Ocean. *Science* 289, 759–762. doi: 10.1126/science.289.5480.759

Young, C. L., and Ingall, E. D. (2010). Marine dissolved organic phosphorus composition: insights from samples recovered using combined electrodialysis/ reverse osmosis. *Aquat. Geochem.* 16, 563–574. doi: 10.1007/s10498-009-9087-y

Zehr, J. P. (2011). Nitrogen fixation by marine cyanobacteria. *Trends Microbiol.* 19, 162–173. doi: 10.1016/j.tim.2010.12.004

Zehr, J. P., and Capone, D. G. (2020). Changing perspectives in marine nitrogen fixation. *Science* 368:eaay9514. doi: 10.1126/science.aay9514

Zehr, J. P., Waterbury, J. B., Turner, P. J., Montoya, J. P., Omoregie, E., Steward, G. F., et al. (2001). Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. *Nature* 412, 635–638. doi: 10.1038/35088063

Check for
updates

Changes in the Gut Metabolic Profile of Gestational Diabetes Mellitus Rats Following Probiotic Supplementation

Qing-Xiang Zheng[1,2†], Hai-Wei Wang[1†], Xiu-Min Jiang[1*], Li Ge[3], Yu-Ting Lai[3], Xin-Yong Jiang[3], Ping-Ping Huang[3], Fan Chen[3] and Xiao-Qian Chen[1,2]

[1] Fujian Maternity and Child Health Hospital, Affiliated to Fujian Medical University, Fuzhou, China, [2] Fujian Obstetrics and Gynecology Hospital, Affiliated to Fujian Medical University, Fuzhou, China, [3] School of Nursing, Fujian University of Traditional Chinese Medicine, Fuzhou, China

The roles of gut microbiota and metabolomics in women with gestational diabetes mellitus (GDM) are not well understood. This study investigated the gut metabolomic profiling of GDM rats and GDM rats treated with probiotic supplements. Associations between gut metabolites and microbiota were also studied in GDM rats. Liquid chromatography–mass spectrometry was used to detect gut metabolites in GDM rats and GDM rats treated with probiotic supplements of 0.5 g (low-dose group) or 1 g (high-dose group) for 15 days. Each gram of probiotic supplement contained 5×10^7 colony-forming units (CFU) of *Lactobacillus rhamnosus LGG* and 1×10^8 CFU of *Bifidobacterium animalis subspecies lactis* Bb12. The association between gut metabolites and microbiota in GDM rats was investigated using Spearman's correlation. Finally, 10 rats in the normal pregnant group, eight rats in the GDM model group, eight GDM rats in the low-dose probiotics group, and nine GDM rats in the high-dose probiotics group were further studied. Serum parameters and pancreatic and colon histology were significantly changed in GDM rats, and these were restored using probiotic supplements. In total, 999 gut metabolites were detected in the feces, and GDM rats were distinguished from normal rats. The levels of 44 metabolites were increased in GDM rats, and they were alleviated using probiotic supplements. Changes in metabolites in GDM rats were associated with amino acids and bile acids metabolism signaling pathways. Furthermore, changes in metabolites after probiotic supplementation were associated with porphyrin and chlorophyll metabolism pathways. We found that the *Allobaculum* genus displayed strong positive correlations, whereas the *Bryobacter* and *Gemmatimonas* genera displayed strong negative correlations with metabolisms of amino acids and bile acids in GDM rats. The *Lactobacillus* and *Bifidobacterium* genera were positively correlated with gut metabolites. Overall, our results showed that metabolism signaling pathways of amino acids and bile acids are associated with the development of GDM. Probiotic supplements alleviate the pathology of GDM through the metabolism pathways of amino acids, bile acids, porphyrin, and chlorophyll.

Keywords: gut metabolites, gestational diabetes mellitus rats, probiotic supplements, glucose and insulin resistance, host metabolism

INTRODUCTION

Gestational diabetes mellitus (GDM) is the most common type of obstetric complication (Veeraswamy et al., 2012) and affects 14.8% of pregnant women in China (Gao et al., 2019). The incidence of GDM has increased globally due to poor dietary structure, unhealthy lifestyle changes, and environmental pollution (Ovesen et al., 2018; Rasmussen et al., 2020). GDM is not only associated with the adverse maternal-infant outcomes in the short term (Kurtzhals et al., 2018) but is also related to metabolic diseases in the woman and offspring in the long term, such as childhood obesity and type 2 diabetes (Kurtzhals et al., 2018; Ponnusamy et al., 2020). Therefore, effective measures to prevent and treat GDM are required.

The gut microbiome plays important roles in modulating host metabolism and immune cell development (He et al., 2020). Dysregulated gut microbiota is associated with diabetes, obesity, and other metabolic diseases (Qin et al., 2012). Metabolites also play important roles in revealing the gut microbiome and maintaining the health status of the host (Liu et al., 2017; Olson et al., 2018; Canfora et al., 2019; He et al., 2020). Abnormalities in gut microbiota and metabolomics in women are important for pathogenesis of GDM. GDM-related differential metabolites are mainly involved in metabolic pathways such as steroid hormone biosynthesis, metabolisms of amino acids, fatty acids, arachidonic acid, and butyric acid, and bile secretion (Li, 2019); however, gut metabolomics of GDM remain unclear.

Probiotics have positive effects on preventing and alleviating diabetes by regulating the intestinal microenvironment and immunity (Liu et al., 2017). Probiotic therapy may be a promising approach to improve glucose control and insulin resistance among pregnant women (de Brito Alves et al., 2019; Pan et al., 2021). Previously, our results showed that *Lactobacillus* and *Bifidobacterium* probiotic supplements could reduce the fasting blood glucose level of GDM rats by restoring the diversity of gut microbiota (Zheng et al., 2021). However, the effects of probiotic supplements on gut metabolomics in GDM are unclear. The associations between gut metabolites and microbiota in GDM rats are also unclear. Therefore, in this study, we analyzed gut metabolomic profiles of GDM rats and GDM rats treated with *Lactobacillus* and *Bifidobacterium* probiotic supplements. Associations between gut metabolites and microbiota in GDM rats were also studied.

MATERIALS AND METHODS

Gestational Diabetes Mellitus Rats Mode and Probiotics Supplement

The procedures of GDM model construction and probiotics supplement can be accessed from our previous study (Zheng et al., 2021). Briefly, after adaptive feeding, 48-week-old female Sprague–Dawley rats were randomly divided into normal pregnant group that fed with ordinary feed and GDM model group that received high-fat and high-sugar feed for 6 weeks.

Then, two female rats were paired with one male rat overnight. Pregnant rats were validated by vaginal secretions and recorded as embryonic day 1. The pregnant rats were further injected with streptozocin (25 mg/kg) to induce the development of GDM. Fasting blood glucose values higher than 7.8 mmol/L at embryonic day 4 were considered indicative of GDM rats. The GDM rats were intragastrically administrated with or without probiotic supplements of 0.5 g (low-dose group) or 1 g (high-dose group) for 15 days. Each gram of probiotic supplement contained 5×10^7 colony-forming units (CFU) of *Lactobacillus rhamnosus LGG* and 1×10^8 CFU of *Bifidobacterium animalis subspecies lactis* Bb12 (Life-space®, Australia). Finally, 10 rats in the normal pregnant group, eight rats in the GDM model group, eight GDM rats in the low-dose probiotics group, and nine GDM rats in the high-dose probiotics group were further studied. All animal procedures were approved by the Fujian University of Traditional Chinese Medicine (certificate number: SYXK 2019-0007; ethics approval number: FJTCM IACUC 2020020).

Enzyme-Linked Immunosorbent Assay of Serum Parameters

Abdominal aortic blood was collected on embryonic day 19. Before blood extraction, a dose of 5 ml/kg of 20% uratan solution was injected intraperitoneally for anesthesia. Abdominal aortic blood was placed at 25°C room temperature for 2 h, centrifuged at a low temperature to obtain the serum sample, and finally stored at −80°C. Concentrations of interleukin-6 (IL-6), tumor necrosis factor–α (TNF-α), nitric oxide (NO), insulin, and low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and lipid factors including total cholesterol (TC) and triglyceride (TG) were measured using enzyme-linked immunosorbent assay (ELISA) kits (Boster Biological Technology Co., Ltd., Wuhan, Hubei, China for IL-6; ProteinTech® Group, Chicago, IL, United States, for insulin and TNF-α; and Nanjing Jiancheng Bioengineering Institute Co., Ltd., Nanjing, Jiangsu, China, for NO, LDL-C, HDL-C, TC, and TG). The value of homeostasis model assessment–estimated insulin resistance (HOMA-IR) was also calculated.

Histopathological Analysis

The pancreatic and colon tissues of rats were quickly collected and placed on ice. The tissues were fixed, embedded in paraffin, and stained with hematoxylin and eosin for histopathological analysis.

Analysis of 16S rRNA Sequencing Data

The procedures and analysis method for 16S rRNA sequencing have been described in a previous study (Zheng et al., 2021). Raw sequence data are available in the Sequence Read Archive under the BioProject accession number PRJNA770477. The 16S rRNA sequencing was performed using Illumina HiSeq platform (Illumina, California, United States). Raw data were cleared using FLASH, TrimMomatic, and UCHIME software. In addition, operational taxonomic units (OTUs) of 16S rRNA sequencing data were clustered using the UCLUST software based on 97%

similarity. The obtained OTUs were then matched with the SILVA database for taxonomic assignment.

Fecal Samples Processing

At embryonic day 19, fecal samples were collected from each rat and stored in a $-80°C$ refrigerator until metabolomic analysis. For fecal sample pretreatment, 50 mg of sample was weighed into an Eppendorf tube, and 1,000 μl of extract solution (acetonitrile:methanol:water = 2:2:1, with an isotopically labeled internal standard mixture) was added. After subjecting to a 30-s vortex, the samples were homogenized at 35 Hz for 4 min and sonicated for 5 min in an ice-water bath. The homogenization and sonication cycles were repeated three times. The samples were then incubated for 1 h at $-40°C$ and centrifuged at 12,000 rpm for 15 min at 4°C. The resulting supernatant was transferred to a fresh glass vial for further analysis. The quality control sample was prepared by mixing an equal aliquot of the supernatant from all samples.

Liquid Chromatography With Tandem Mass Spectrometry Analysis

LC-MS/MS analyses were performed using an ultrahigh-performance liquid chromatography system (Vanquish, Thermo Fisher Scientific) with a UPLC BEH amide column $(2.1 \times 100$ mm, 1.7 μm) coupled to a Q Exactive HF-X mass spectrometer (Orbitrap MS, Thermo Fisher Scientific). The mobile phase consisted of 25 mmol/L of ammonium acetate and 25 mmol/L of ammonium hydroxide in water (pH = 9.75) and acetonitrile. The analysis was carried out with elution gradients as follows: 0–0.5 min, 95% B; 0.5–7.0 min, 95–65% B; 7.0–8.0 min, 65–40% B; 8.0–9.0 min, 40% B; 9.0–9.1 min, 40–95% B; and 9.1–12.0 min, 95% B. The column temperature was maintained at 30°C. The auto-sampler temperature was 4°C, and the injection volume was 3 μl.

The Q Exactive HF-X mass spectrometer was used for its ability to acquire MS/MS spectra in the information-dependent acquisition mode in the control of the acquisition software (Xcalibur, Thermo Fisher Scientific). In this mode, the acquisition software continuously evaluates the full-scan MS spectrum. Electrospray ionization source conditions were set as follows: sheath gas flow rate of 50 arbitrary units, aux gas flow rate of 10 arbitrary units, capillary temperature at 320°C, a full MS resolution of 60,000, MS/MS resolution of 7,500, collision energy of 10/30/60 in the normalized collision energy mode, and spray voltage of 3.5 kV (positive) or -3.2 kV (negative). Meanwhile, raw data were converted to the mzXML format using the ProteoWizard application and processed with an in-house program, which was developed using R and based on the XCMS platform, for peak detection, extraction, alignment, and integration. Then, an in-house MS2 database (BiotreeDB) was used for metabolite annotation. The cutoff for annotation was set at 0.3.

Hierarchical Clustering Analysis

The hierarchical clustering analysis of the gut metabolites was performed using the pheatmap function in R software.

Orthogonal Partial Least-Squares Discrimination Analysis

The Orthogonal Partial Least-Squares Discrimination Analysis (OPLS-DA) was used to analyze the metabolic patterns. The OPLS-DA permutation test was used to test the effect of the model fit. The R^2Y (cum) index of the OPLS-DA permutation test was used to determine the accuracy of the model fit. The Q^2 (cum) index refers to the predictive power of the model, and it is an evaluation parameter that is model-validated to prevent random fitting or overfitting. When the R^2Y (cum) and Q^2 (cum) indexes are closer to 1, the model fits well. In addition, if they are greater than 0.5, then the model is acceptable. On the basis of the threshold criteria of the projection of the first principal component of OPLS-DA > 1, an absolute fold change of < 0.5 or > 2 and a P-value of < 0.05 indicated a significant difference.

Metabolic Signaling Pathways Analysis

Metabolic signaling pathways of gut metabolites were obtained from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database.[1] The signaling pathways included more than four metabolites.

Correlation Analysis and Co-occurrence Network of Target Metabolites and Gut Microbiota

The correlations of metabolites and the relative abundance of gut microbiota were analyzed using the Spearman's correlation test through the psych package in the R software. The co-occurrence network of metabolites and gut microbiota was constructed using the Cytoscape software (version 3.1.1).

Data Analysis

Statistical analysis and plotting of figures were conducted using GraphPad Prism (version 7.0), SPSS (version 25.0), and R programming language software (version 4.4.1). A normality test was performed to determine whether the data obeyed a normal distribution. If data obeyed a normal distribution, then a t-test or *post-hoc* test was performed to analyze the data. Otherwise, a non-parametric test was conducted. Statistical significance was set at $P < 0.05$, and data were presented as mean ± standard deviation (mean ± SD).

RESULTS

Probiotic Supplements Restored Normal Serum Parameters, and Pancreatic and Colon Histology Are Influenced by Gestational Diabetes Mellitus

Previously, we successfully constructed a GDM rat model and demonstrated that *Lactobacillus* and *Bifidobacterium* probiotic supplements can reduce the fasting blood glucose level of GDM rats by restoring the diversity of gut microbiota

---

[1] https://www.kegg.jp/kegg/pathway.html

(Zheng et al., 2021). Furthermore, using the same GDM model, we tested the effects of *Lactobacillus* and *Bifidobacterium* probiotic supplements on the concentrations of IL-6, TNF-α, NO, insulin, LDL-C, HDL-C, TC, and TG and estimated the value of HOMA-IR in the serum of GDM rats. In total, 10 normal pregnant rats, eight GDM rats, eight GDM rats with low-dosage probiotic supplements, and nine GDM rats with high-dosage probiotic supplements were tested. Compared to the normal pregnant rats, the concentrations of TNF-α, NO, insulin, LDL-C, TC, and TG ($P = 0.002$, $P = 0.003$, $P = 0.041$, $P = 0.000$, $P = 0.000$, and $P = 0.000$, respectively) and the value of HOMA-IR were increased in GDM rats ($P = 0.000$). Moreover, high-dose *Lactobacillus* and *Bifidobacterium* probiotic supplements significantly alleviated the concentrations of TNF-α, insulin, and LDL-C ($P = 0.009$, $P = 0.011$, and $P = 0.007$, respectively) and the value of HOMA-IR in GDM rats ($P = 0.000$). However, high-dose *Lactobacillus* and *Bifidobacterium* probiotic supplements did not alleviate the concentrations of NO, TC, and TG in GDM rats ($P = 0.653$, $P = 0.046$, and $P = 0.002$, respectively) (**Figure 1A**). In contrast, the concentrations of HDL-C were decreased in GDM rats ($P = 0.000$) and could not be further increased by the *Lactobacillus* and *Bifidobacterium* probiotic supplements ($P = 0.794$ and $P = 0.358$, respectively) (**Figure 1A**). Furthermore, the concentration of IL-6 was not significantly different among normal pregnant rats, GDM rats, and GDM rats with probiotic supplements ($P < 0.05$) (**Figure 1A**).

Lactobacillus and *Bifidobacterium* probiotic supplements could also influence the affected morphological and structural characteristics of pancreatic tissues in GDM rats. Compared to normal pregnant rats, the pancreatic islands of GDM rats were irregular and cyto-reduced and exhibited vacuolar degeneration. In addition, the capillary vessels of the pancreatic islands had undergone congestive expansion. After probiotic supplementation, the pancreatic islands of the rats were clearer, had less vacuolar degeneration, and showed cellular increase compared to these parameters in GDM rats (**Figure 1B**).

The morphological and structural characteristics of the colon tissues of GDM rats were also analyzed (**Figure 1B**). The colon mucous membrane and cell–cell junction of the epithelial cells of GDM rats were significantly damaged. The cryptostructures of colon tissues partially disappeared, and the beaker cells were reduced. After probiotic supplementation, the colon mucous membrane of rats was more complete and displayed queuing discipline compared to that in the GDM model group, and the numbers of cryptostructures and beaker cells were further increased.

Altered Metabolites in Gestational Diabetes Mellitus Rats

One possible mechanism of *Lactobacillus* and *Bifidobacterium* probiotic supplements restoring the normal serum parameters and pancreatic and colon histology influenced by GDM is the regulation of gut metabolites. Fecal metabolomic studies facilitate the interpretations of the metabolic interactions between the host, diet, and gut microbiota and provide functional data on the microbiome. Therefore, in this study, we further explored the

influence of probiotic supplements on gut metabolites in the fecal samples of GDM rats using LC-MS/MS.

In total, 999 gut metabolites were detected. Normal rats were clearly distinguished from the GDM rats based on the levels of 999 gut metabolites by using the unsupervised clustering heatmap (**Figure 2A**). In addition, in the OPLS-DA score plots, normal pregnant rats and GDM rats were clearly separated from each other (**Figure 2B**). Similarly, the OPLS-DA permutation test indicated that the classifications of normal and GDM rats were robust and exhibited good fitness and prediction [(R^2Y (cum) = (0, 0.71), Q2 (cum) = (0, -1.01)] (**Figure 2C**). Moreover, compared to normal rats, the levels of most gut metabolites were increased in GDM rats (**Figure 2A**).

However, in the unsupervised clustering heatmap, the GDM rats were indistinguishable from the GDM rats treated with low- or high-dose *Lactobacillus* and *Bifidobacterium* probiotic supplements (**Figure 2A**). In addition, in the OPLS-DA score plots and in the OPLS-DA permutation test, GDM rats and GDM rats with low probiotic supplements could not be distinctly separated from each other (**Figures 2B,C**). In contrast, GDM rats with high probiotic supplements represented a separate group in the OPLS-DA score plots and in the OPLS-DA permutation test (**Figures 2B,C**).

Altered Metabolites in Gestational Diabetes Mellitus Rats Treated With Probiotics

Although most of the altered metabolites in GDM rats could not be restored by *Lactobacillus* and *Bifidobacterium* probiotic supplements, partial metabolites were significantly influenced by probiotic supplements. We detected 44 metabolites that were increased in GDM rats, and their levels were alleviated by high-dose *Lactobacillus* and *Bifidobacterium* probiotic supplements (**Figure 3A**).

The levels of N-methyl-α-aminoisobutyric acid, L-proline, L-threonine, isovalerylalanine, 1-aminocyclopropanecarboxylic acid, alanyl-gamma-glutamate, and histidinyl-valine were shown in the box plots. The concentrations of N-methyl-α-aminoisobutyric acid, valyl-proline, L-threonine, isovalerylalanine, 1-aminocyclopropanecarboxylic acid, alanyl-gamma-glutamate, and histidinyl-valine concentrations were increased in GDM rats ($P = 0.008$, $P = 0.002$, $P = 0.000$, $P = 0.004$, $P = 0.015$, $P = 0.000$, and $P = 0.008$, respectively). Moreover, high-dose *Lactobacillus* and *Bifidobacterium* probiotic supplements significantly reduced the concentrations of these metabolites in GDM rats ($P < 0.05$) (**Figure 3B**), suggesting that probiotic supplements restored normal serum parameters and pancreatic and colon histology influenced by GDM through the regulation of these metabolites.

Changes in Metabolism Signaling Pathways of Amino Acids and Bile Acids in Gestational Diabetes Mellitus Rats

The changes in metabolites in the GDM rats were further studied. There were 499 metabolites that significantly changed in GDM rats compared to normal pregnant rats. Furthermore, the top
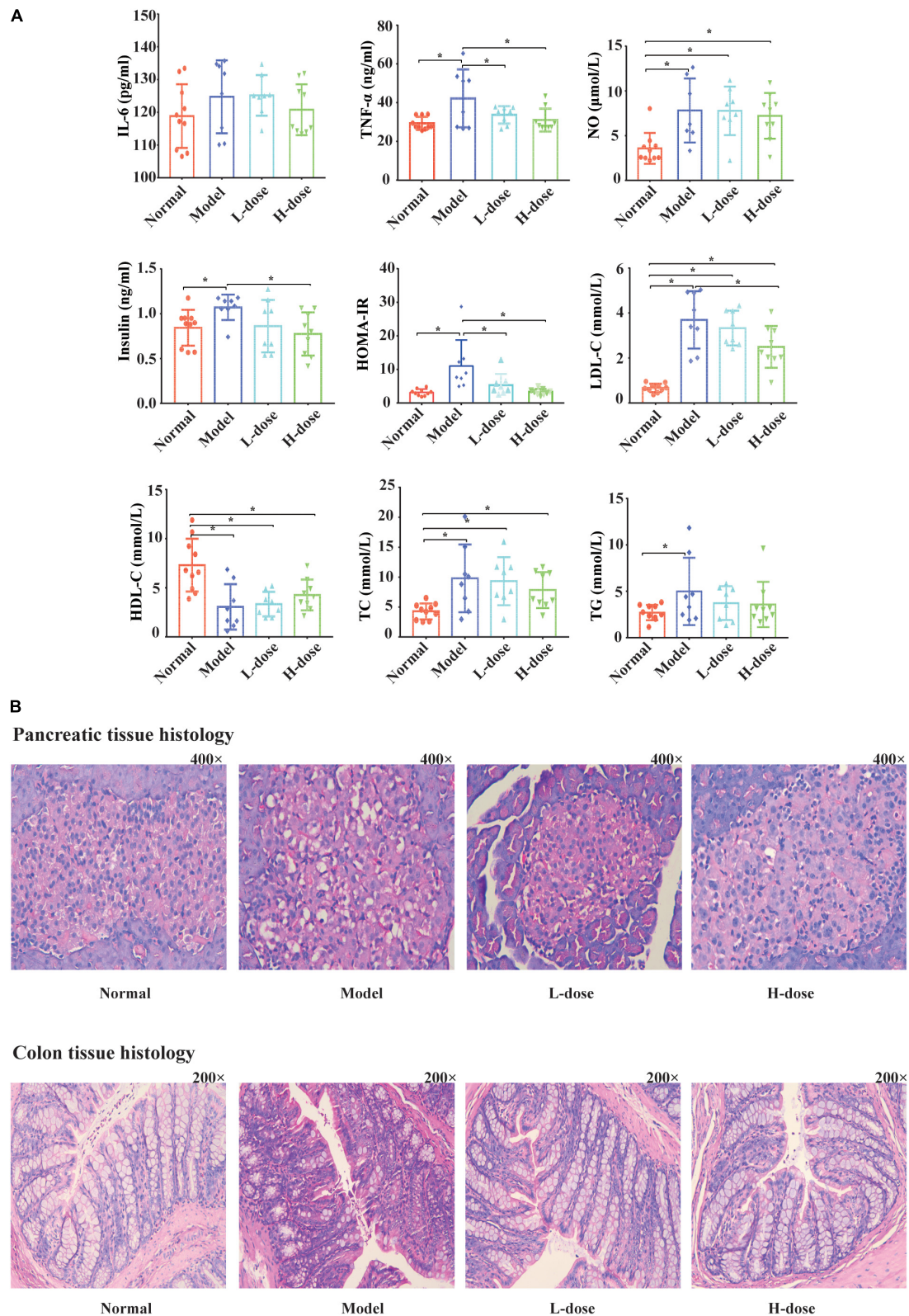
FIGURE 1 | Probiotics supplements restore the normal serum parameters and pancreatic and colon histology influenced by GDM. **(A)** The concentrations of IL-6, TNF-α, NO, insulin, HOMA-IR, LDL-C, HDL, TC, and TG in GDM rats and GDM rats treated with *Lactobacillus* and *Bifidobacterium* probiotics supplements were tested. **(B)** The hematoxylin-eosin staining sections of the pancreas tissue and the colon tissue in GDM rats and GDM rats treated with probiotics supplements. *$P < 0.05$.
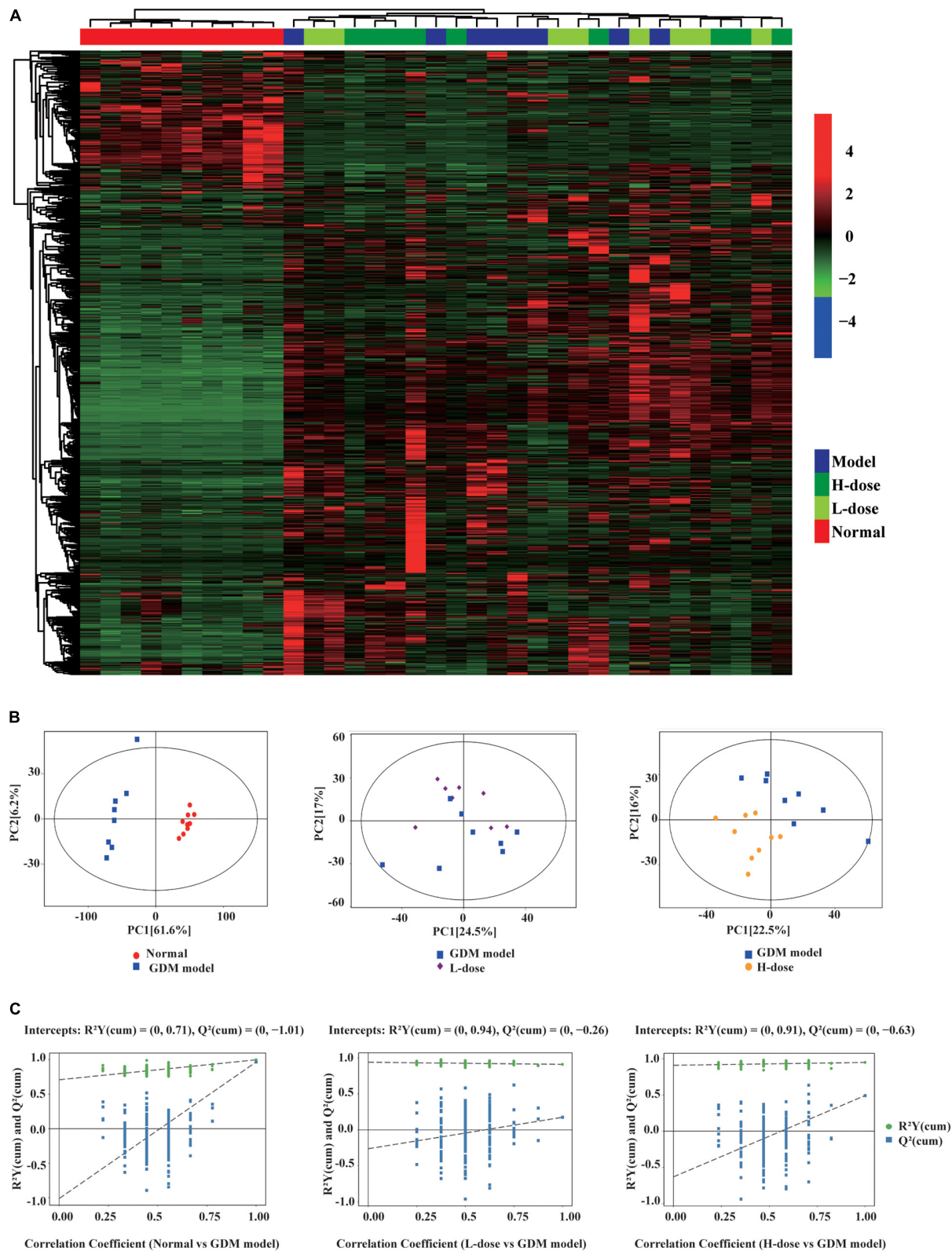
FIGURE 2 | The gut metabolites feature of GDM rats and GDM rats treated with probiotics. **(A)** Hierarchical clustering analysis of 999 metabolites in GDM rats and GDM rats treated with *Lactobacillus* and *Bifidobacterium* probiotics supplements. The OPLS-DA analysis **(B)** and the OPLS-DA permutation test **(C)** of the gut metabolites in GDM rats and GDM rats treated with probiotics supplements.
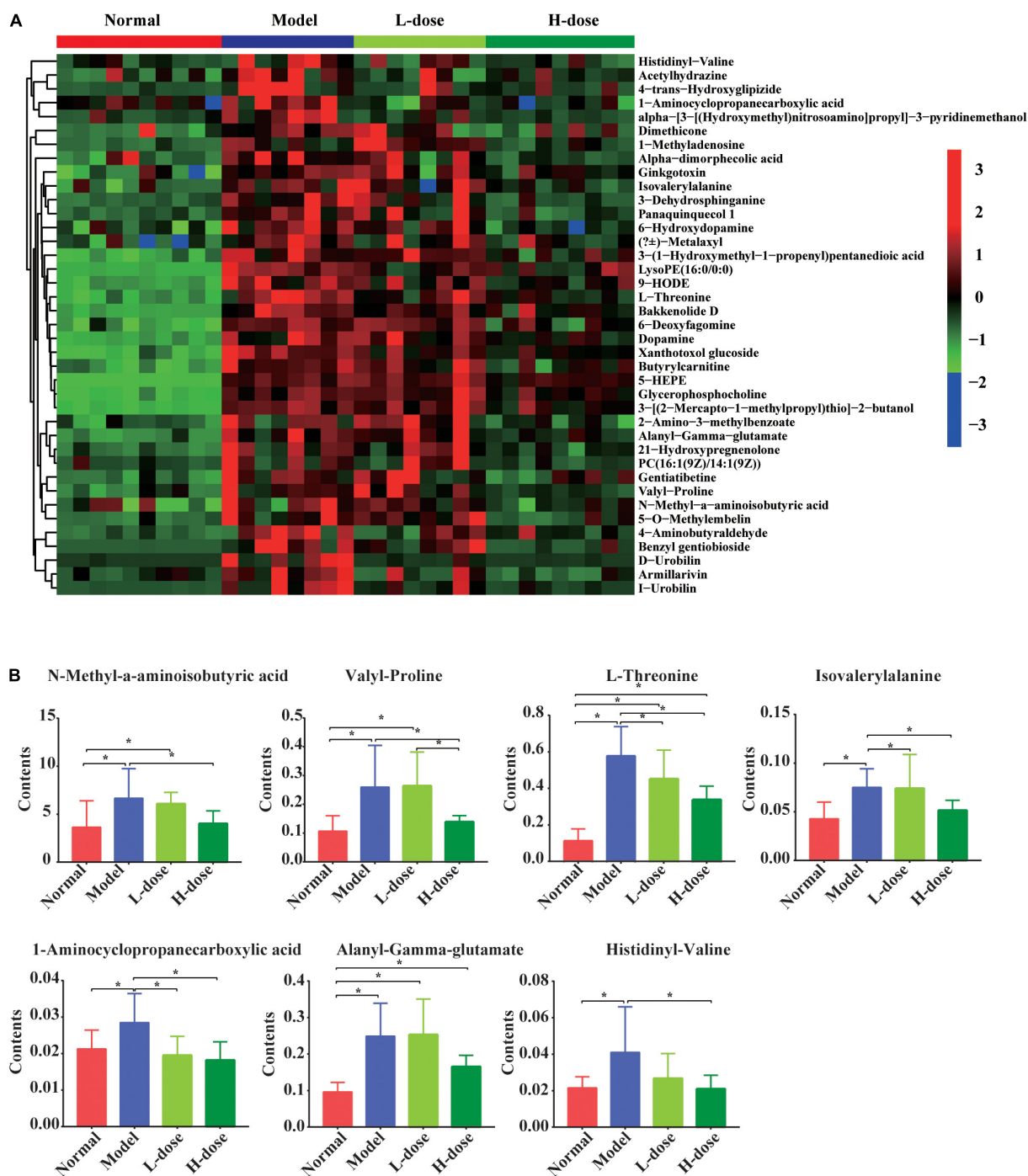
FIGURE 3 | The influenced gut metabolites after *Lactobacillus* and *Bifidobacterium* probiotics supplements in GDM rats. **(A)** Hierarchical clustering analysis of 44 metabolites in GDM rats and GDM rats treated with *Lactobacillus* and *Bifidobacterium* probiotics supplements. **(B)** The comparison of seven amino acids by the ANOVA/Kruskal–Wallis and *post-hoc* test. *$P < 0.05$.

20 most significantly changed metabolites in GDM rats vs. normal pregnant rats were shown in **Table 1**. Those significantly changed metabolites in GDM rats was dihydroisoalantolactone, followed by (E)-2-Butyl-2-octenal and 4-Hydroxy-4-(3-pyridyl)-butanoic acid.

Next, the KEGG signaling pathways associated with the significantly altered metabolites in GDM rats vs. normal pregnant rats were studied. Although these metabolites were diverse, they were mainly concentrated in signaling pathways of amino acids, organic acids, and bile acids (**Figure 4A**).

TABLE 1 | The top 20 most significantly changed metabolites in normal pregnant rats vs. GDM rats.

MS2 name	RT	MZ	VIP	P-value	Fold change
Dihydroisoalantolactone	164.29	235.17	1.55	3.33017E-05	13574.95
(E)-2-Butyl-2-octenal	26.00	183.17	1.55	3.48987E-06	1914.06
4-Hydroxy-4-(3-pyridyl)-butanoic acid	276.66	182.08	1.55	0.000113978	10738.12
Ginsenoyne A linoleate	291.61	521.39	1.55	0.000731182	458.18
Tetraphyllin B	385.17	288.11	1.54	4.58102E-06	42.55
L-Palmitoylcarnitine	191.30	400.34	1.54	1.30394E-05	82.88
18-Nor-4(19),8,11,13-abietatetraene	236.83	255.21	1.54	0.00293506	2975.01
Vaccenyl carnitine	188.27	426.36	1.54	8.58595E-06	80.69
3,4-Dihydroxy-2-hydroxymethyl-1-pyrrolidinepropanamide	422.04	205.12	1.54	2.43937E-06	18.21
(3R,6′Z)-3,4-Dihydro-8-hydroxy-3-(6-pentadecenyl)-1H-2-benzopyran-1-one	213.94	373.27	1.54	0.000121065	47.22
6,7-Dihydro-5-methyl-5H-cyclopenta[b]pyrazine	252.91	135.09	1.51	6.03717E-07	0.18
[10]-Gingerdione	52.02	349.24	1.51	0.000162074	0.03
5-Aminopentanal	257.63	102.09	1.48	5.90244E-07	0.22
Tryptophyl-Valine	307.05	304.16	1.47	5.79417E-07	0.24
Linatine	383.38	260.12	1.47	6.24363E-07	0.24
1,9-Nonanedithiol	289.66	193.11	1.47	2.31553E-05	0.22
3-Methylpyrrolo[1,2-a]pyrazine	144.88	134.07	1.47	3.88429E-06	0.19
2-Methoxycanthin-6-one	307.70	251.08	1.47	4.322E-06	0.27
Calystegin A3	77.68	160.10	1.47	4.32426E-13	0.09
Acetylcholine	262.24	146.12	1.47	1.21646E-07	0.16

RT, retention time; MZ, mass-to-charge ratio; VIP, PLS-DA first principal component variable importance in projection; P-value, t-test significance; the Fold change was < 0.5 or > 2.

Metabolites including amino acids, organic acids, and bile acids and glucuronide metabolites were further depicted in the heatmap (**Figure 4B**). The concentrations of most of the metabolites were increased in GDM rats. For example, pipecolic acid, l-lysine, 5-aminopentanoic acid, L-phenylalanine, L-glutamic acid, and cholic acid from porphyrin and chlorophyll metabolism, biosynthesis of amino acids, and bile secretion signaling pathways were all significantly increased in GDM rats compared to their levels in normal pregnant rats ($P = 0.000$, $P = 0.006$, $P = 0.006$, $P = 0.000$, $P = 0.003$, and $P = 0.015$, respectively) (**Figure 4C**).

Metabolism Pathways of Porphyrin and Chlorophyll Were Associated With the Altered Metabolites Following Probiotic Supplementation
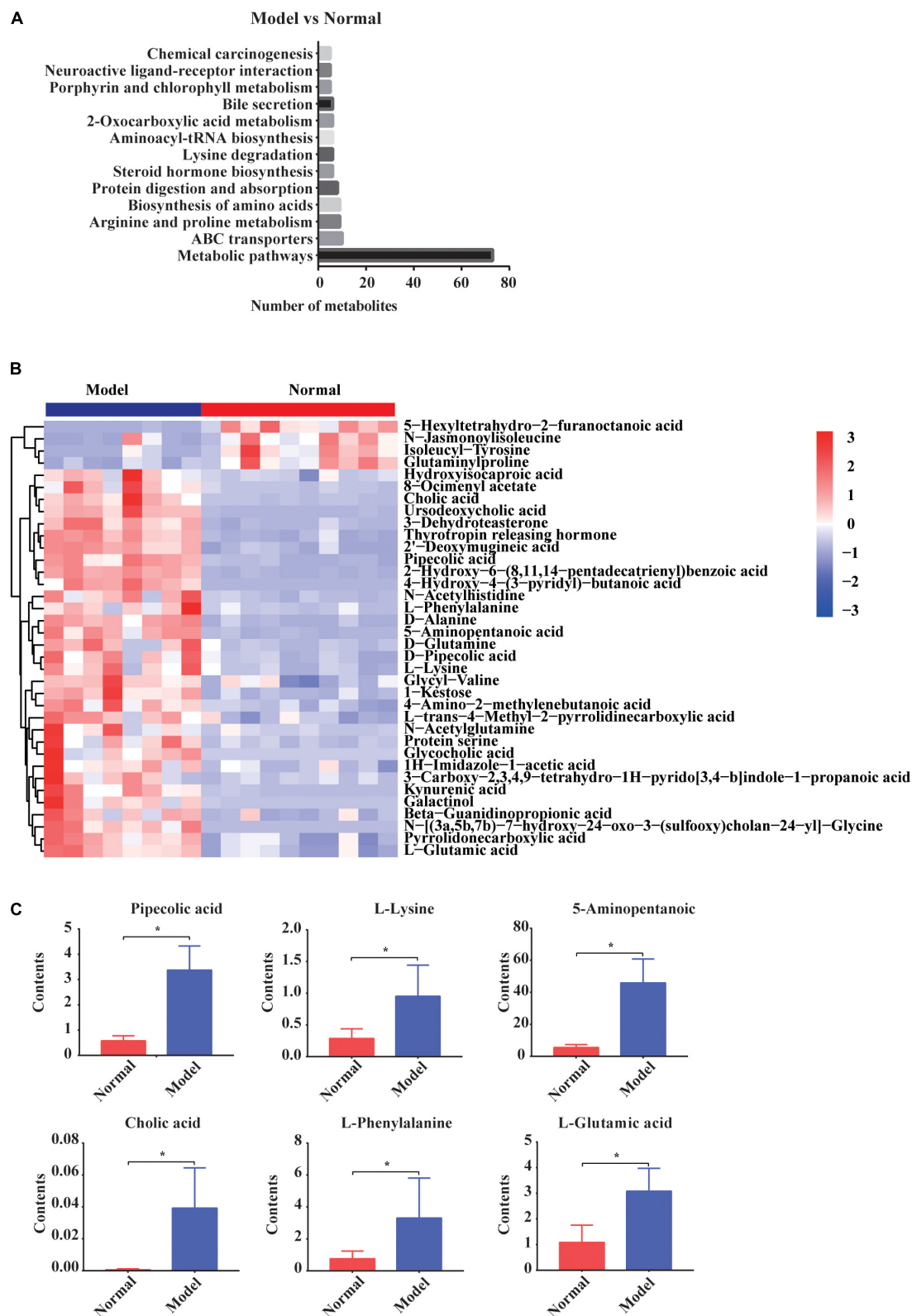
The changes in metabolites in GDM rats treated with *Lactobacillus* and *Bifidobacterium* probiotics were further studied. Compared to the metabolites in the GDM rats, low-dosage probiotic supplements induced the changes of 56

metabolites and high-dosage probiotic supplements induced the changes of 102 metabolites. Moreover, the top 20 most significantly changed metabolites in the GDM rats vs. GDM rats with low-dosage probiotic supplementation and GDM rats vs. GDM rats with high-dose probiotic supplementation are shown in **Tables 2**, **3**, respectively. LysoPE and 2-aminoacrylic acid were significantly changed after low-dose probiotic supplementation, whereas [4]-gingerdiol 3,5-diacetate and propionylcarnitine were significantly changed after high-dose probiotic supplementation.

Next, the KEGG signaling pathways associated with the significantly changed metabolites in GDM rats vs. GDM rats with low-dose probiotic supplementation and GDM rats vs. GDM rats with high-dose probiotic supplementation were analyzed. Compared to GDM rats, the metabolites induced by low- or high-dose probiotic supplementation were concentrated on the metabolism pathways of porphyrin and chlorophyll (**Figures 5A,B**). Interestingly, the metabolism pathways of porphyrin and chlorophyll were also associated with the altered metabolites in GDM rats vs. normal pregnant rats (**Figure 4A**). We further showed that the levels of the metabolites L-threonine, I-urobilin, D-urobilin, and L-urobilin from the metabolism pathways of porphyrin and chlorophyll were decreased by high-dose probiotic supplementation in GDM rats ($P = 0.001$, $P = 0.002$, $P = 0.002$, and $P = 0.005$, respectively) (**Figure 5C**). In addition, D-urobilin and L-urobilin levels were also decreased by low-dose probiotic supplementation in GDM rats ($P = 0.005$ and $P = 0.006$, respectively) (**Figure 5C**), whereas L-glutamic acid level was decreased by low-dosage probiotic administration in GDM rats ($P = 0.026$) (**Figure 5C**).

Changed Metabolites From Metabolism Signaling Pathways of Amino Acids and Bile Acids After Probiotic Supplementation

Although the significantly altered metabolites in GDM rats vs. GDM rats with *Lactobacillus* and *Bifidobacterium* probiotic supplementation were not enriched with metabolism signaling pathways of amino acids and bile acids, some metabolites from these pathways were still influenced by probiotic supplements. We showed that the concentration of 2-aminoacrylic acid was increased in GDM rats following supplementation with low or high dosages *of Lactobacillus* and *Bifidobacterium* probiotics, whereas that of 1-aminocyclopropanecarboxylic acid was decreased in GDM rats following supplementation with low or high dosages of probiotics (**Figures 6A,B**). Moreover, the concentrations of valyllysine, L-glutamic acid, serylvaline, thyrotropin-releasing hormone, clupanodonyl carnitine, and falcarindiol were significantly increased after the administration of low-dose *Lactobacillus* and *Bifidobacterium* probiotic supplements (**Figure 6A**). In contrast, the concentrations of 1-aminocyclopropanecarboxylic acid, benzyl-gentiobioside, L-threonine, isovalerylalanine, aminoadipic acid, alanyl-gamma-glutamate, phenylacetylglycine, L-leucine, and

FIGURE 4 | The changed gut metabolites and signaling pathways in normal pregnant rats vs. GDM rats. **(A)** Gut metabolites' signaling pathways in KEGG Database. **(B)** Hierarchical clustering analysis of 36 metabolites in GDM rats. **(C)** The comparison of amino acids and bile acids that participated in gut metabolites' signaling pathways by independent samples *t*-test or non-parametric test. *$P < 0.05$.

TABLE 2 | The top 20 most significantly changed metabolites in GDM rats treated with lose-dose probiotics supplements vs. GDM rats.

MS2 name	RT	MZ	VIP	P-value	Fold change
LysoPE (0:0/20:5(5Z,8Z,11Z,14Z,17Z))	108.85	500.27	2.45	0.030380947	10.67
2-Aminoacrylic acid	408.26	88.04	2.18	0.020937481	2.91
2,5-Dimethyloxazole	150.27	98.06	2.15	0.02726155	2.82
Kudzusaponin SA2	307.74	945.50	2.04	0.047526679	3.72
Polysorbate 60	257.00	435.29	1.73	0.030919879	2.31
Clupanodonyl carnitine	235.12	474.36	1.72	0.015841942	3.34
Saxitoxin	288.47	300.14	1.60	0.031002897	3.44
PI (16:1(9Z)/18:0)	209.56	837.56	1.37	0.046339361	2.62
N-Hydroxy-L-tyrosine	289.75	198.08	2.44	0.040165452	0.36
(R)-Salsolinol	190.40	180.10	2.43	0.021349123	0.26
L-Urobilin	289.95	595.35	2.32	0.043622447	0.08
Quinoline-4,8-diol	186.32	162.05	2.21	0.041281558	0.38
Octadecanedioic acid	151.91	315.25	2.12	0.032593502	0.12
D-Urobilin	255.62	589.30	2.12	0.038878374	0.12
Dukunolide C	402.78	541.17	2.10	0.015048225	0.22
alpha-[3-[(Hydroxymethyl) nitrosoamino]propyl]-3-pyridinemethanol	296.84	226.12	1.94	0.039346005	0.33
2,3-Dimethyl-5-(2-propenyl) pyrazine	230.60	149.11	1.65	0.04508631	0.48

RT, retention time; MZ, mass-to-charge ratio; VIP, PLS-DA first principal component variable importance in projection; P-value, t-test significance; the Fold change was < 0.5 or > 2.

TABLE 3 | The top 20 most significantly changed metabolites in GDM rats treated with high-dose probiotics supplements vs. GDM rats.

MS2 name	RT	MZ	VIP	P-value	Fold change
[4]-Gingerdiol 3,5-diacetate	255.47	370.22	2.30	0.000782449	3.16
Propionylcarnitine	291.02	218.14	2.29	0.005544317	2.81
LysoPE [0:0/20:5(5Z,8Z,11Z,14Z,17Z)]	108.85	500.27	2.27	0.016205496	4.90
2-Aminoacrylic acid	408.26	88.04	2.24	0.03169583	2.93
Acetamidopropanal	411.68	116.07	2.23	0.009543001	2.22
Carpaine	284.96	479.38	2.20	0.045061999	14.98
Lysyl-Valine	533.40	246.18	2.13	0.045907893	5.69
2,9-Bis (3-methyl-2E-pentenoyl)-2b,9a-dihydroxy-4Z,10(14)-oplopadien-3-one	131.70	443.27	2.13	0.04138836	8.27
Lysyl-Leucine	533.48	260.20	2.06	0.019666532	3.68
Saxitoxin	288.47	300.14	2.04	0.000426356	4.06
L-Leucine	344.89	132.10	2.62	0.00245802	0.21
Aminoadipic acid	82.73	162.08	2.49	0.003544451	0.47
3-Dehydrosphinganine	107.48	300.29	2.46	0.002750615	0.28
L-alpha-Amino-1H-pyrrole-1-hexanoic acid	361.76	197.13	2.42	0.00013354	0.45
Xanthotoxol glucoside	320.46	365.08	2.42	0.000384492	0.59
l-Urobilin	298.02	591.32	2.41	0.018958387	0.14
5-(2-Furanyl)-3,4-dihydro-2H-pyrrole	223.18	136.08	2.40	0.000761589	0.64
(R)-Salsolinol	190.40	180.10	2.40	0.010778167	0.14
Armillarivin	323.14	385.20	2.40	0.005431808	0.20
3-Formyl-6-hydroxyindole	69.03	162.05	2.36	0.039772211	0.13

RT, retention time; MZ, mass-to-charge ratio; VIP, PLS-DA first principal component variable importance in projection; P-value, t-test significance; the Fold change was < 0.5 or > 2.

norvaline were significantly reduced by high-dose probiotic supplements (**Figure 6B**).

Correlations of Gestational Diabetes Mellitus-Related Gut Microbiota and Metabolites

Fecal metabolomics may reflect the functions of the gut microbiome. The altered metabolites in GDM rats may be perturbed by the gut microbiome during GDM pathogenesis. Previously, we studied gut microbiota in GDM rats and GDM rats with *Lactobacillus* and *Bifidobacterium* probiotic supplements (Zheng et al., 2021). Next, we explored the functional correlations between gut microbiota perturbations and metabolite changes. We observed strong correlations between the gut bacterial families and metabolites assigned to the metabolism of amino acids and bile acids in GDM rats, as determined by the Spearman's correlation coefficient. The gut bacterial OTUs belonging to the genus *Allobaculum* displayed strong positive correlations ($P < 0.05$), whereas the genera *Bryobacter* and *Gemmatimonas* displayed strong negative correlations with the metabolism of amino acids and bile acids, such as L-glutamic acid, pipecolic acid, alanyl-gamma glutamate, isovalerylalanine, L-phenylalanine, 5-aminopentanoic acid, L-threonine, glycocholic acid, and lysine ($P < 0.05$) (**Figure 7**). Moreover, the metabolites pipecolic acid, 5-aminopentanoic acid, and cholic acid were associated with

most of the gut bacterial genera in GDM rats ($P < 0.05$) (**Supplementary Figure 1A**).

To further explore the reciprocal synergistic relationships between gut microbiota and metabolites, we constructed a co-occurrence network based on the correlations of the gut bacterial genus and its metabolites in GDM rats (**Figure 7**). We found that metabolites 5-aminopentanoic acid, alanyl-gamma glutamate, benzyl-gentiobioside L-phenylalanine, and the gut bacterial genera *Phascolarctobacterium*, *Barnesiella*, and *Blautia* were at the hub of the co-occurrence network, whereas other gut bacterial genera and their metabolites were connected with those hub factors (**Supplementary Figure 1A**).

Lactobacillus and *Bifidobacterium* Were Positively Correlated With the Gut Metabolites

Lactobacillus *rhamnosus* LGG and *Bifidobacterium lactis* Bb12 were used as probiotic supplements in this study. Therefore, we tested the correlations of the *Lactobacillus* and *Bifidobacterium* genera with gut metabolites. The Spearman's correlation coefficient of the OTU of *Lactobacillus* and *Bifidobacterium* genera and concentrations of 999 detected gut metabolites were determined. On the basis of r-values of > 0.6 and P-values

FIGURE 5 | The changed gut metabolites and pathways in GDM rats treated with probiotics. **(A,B)** The porphyrin and chlorophyll metabolism pathway and metabolic pathway in KEGG Database. **(C)** The comparison of gut metabolites which participated in porphyrin and chlorophyll metabolism pathway and metabolic pathway by the *post-hoc* test or non-parametric test. *$P < 0.05$.
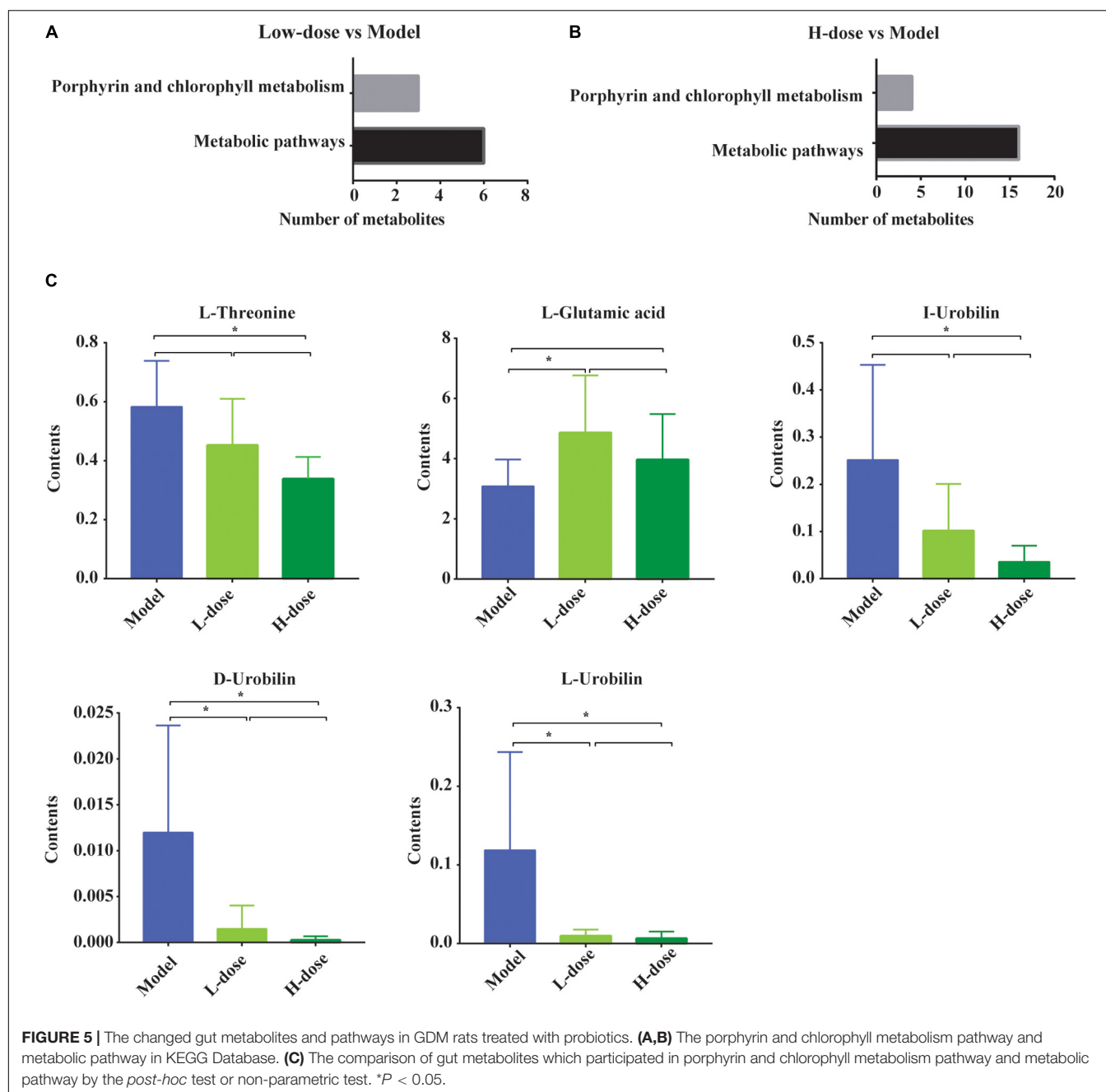
of < 0.01, the *Lactobacillus* genus was positively correlated with 90 gut metabolites and the *Bifidobacterium* genus was positively correlated with 71 gut metabolites. None of the gut metabolite showed a negative correlation with the OTUs of *Lactobacillus* or *Bifidobacterium* genera. The correlations of 26 gut metabolites with *Lactobacillus* and *Bifidobacterium* genera are shown in a heatmap (**Figure 8**); these include the amino acid metabolites arginyl-phenylalanine, L-phenylalanine, D-proline, lysyl-valine, valyl-lysine, lysyl-leucine, L-tyrosine, N-α-acetyl-L-arginine, and tryptophyl-arginine. The co-occurrence network also revealed the connections of *Lactobacillus* and *Bifidobacterium* probiotic supplements with amino acid metabolism (**Supplementary**

Figure 1B). In addition, other metabolites, such as histamine, oleamide, fenpropimorph, dhurrin, and tyramine, were also correlated with the OTUs of the *Lactobacillus* or *Bifidobacterium* genera (**Supplementary Figure 1B**).

DISCUSSION

Probiotics have a mild, stable, and effective influence on controlling blood glucose (Peng et al., 2018; Zhang et al., 2019; Masulli et al., 2020). However, the effects of probiotics on gut metabolites in GDM rats are not clear (Yao et al., 2017). Our

FIGURE 6 | The changed amino acids and bile acids in GDM rats treated with probiotics supplements. **(A)** Hierarchical c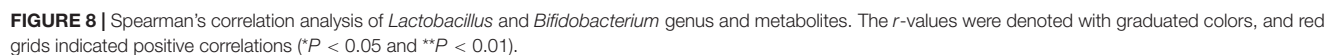lustering analysis of 14 amino acids and bile acids in GDM rats treated with low-dose probiotics supplements vs. GDM rats. **(B)** Hierarchical clustering analysis of 19 amino acids and bile acids in GDM rats treated with high-dose probiotics supplements vs. GDM rats.

results showed that high-dose *Lactobacillus* and *Bifidobacterium* probiotic supplements can decrease the concentrations of TNF-α, insulin, and LDL-C and the value of HOMA-IR in GDM rats. After *Lactobacillus* and *Bifidobacterium* probiotic intervention, the morphological and structural characteristics of pancreatic and colon tissues were similar to those in the normal pregnant group. In addition, after probiotic supplementation, GDM model rats showed significant changes in gut microbiome composition and gut metabolic profiles. Moreover, the perturbed gut microbiota was strongly associated with changes in gut microbiota-related metabolites, suggesting that probiotic supplements can substantially alter the metabolomic profile of the gut microbiome and adjust host metabolite homeostasis.

Our results from a previous study (Zheng et al., 2021) and this study further verified that *L. rhamnosus* LGG and *B. lactis* Bb12 used as probiotic supplements can reduce blood glucose and insulin resistance in GDM rats and also improve the morphological and structural characteristics of their pancreatic and colon tissues. Potential mechanisms of probiotic supplements regulating blood glucose were mainly manifested in regulating the structure of gut microbiota and gut metabolites and then reducing the level of fasting blood glucose and insulin

resistance. In addition, insulin resistance has been regarded as a cause of endothelial dysfunction in diabetes (Zhang et al., 2012).

Changes in gut metabolomics are associated with intestinal dysbacteriosis in the development of obesity and metabolic syndrome (Qin et al., 2012; Karlsson et al., 2013). Moreover, metabolic alterations can further aid in the exploration of the pathological mechanisms of GDM. We found that the fecal metabolic profile of GDM rats was significantly different from that of normal pregnant rats. In total, 999 gut metabolites were detected in the feces. Forty-four metabolites in GDM rats had increased concentrations, which were reduced by probiotic supplements, including amino acids, bile acids, and glucuronide. Changes in metabolites in GDM rats were associated with the metabolism signaling pathways of amino acids and bile acids. Metabolites from metabolism signaling pathways of amino acids and bile acids were influenced by probiotic supplements. Changes in metabolites after probiotic supplementation were associated with the metabolism pathways of porphyrin and chlorophyll.

The fecal metabolic profiles of GDM rats were significantly different from those of normal pregnant rats in the metabolism pathways of amino acids and bile acids. The identified amino acids participating in the development of GDM rats included

FIGURE 7 | Spearman's correlation analysis of metabolites and gut microbiota. The *r*-values were denoted with graduated colors, and red and blue grids indicated positive and negative correlations, respectively (*P < 0.05 and ** P < 0.01).

FIGURE 8 | Spearman's correlation analysis of *Lactobacillus* and *Bifidobacterium* genus and metabolites. The *r*-values were denoted with graduated colors, and red grids indicated positive correlations (*P < 0.05 and **P < 0.01).

1-pyrroline, creatinine, cytosine, 2-hydroxypyridine, choline, 2,6-dimethylpyridine, and quinoline. These results are similar to those of previous studies (Wang et al., 2020; Zhao et al., 2020; Lu et al., 2021) that suggest that fecal gut metabolomics of women with GDM was distinct from that of healthy pregnant women, and the levels of amino acids were disrupted in the former. Moreover, the levels of amino acids in early pregnancy are potential biomarkers for predicting subsequent GDM (Jiang et al., 2020). Consistent with previous reports (Lu et al., 2021), we found that, compared to the metabolic profiles of normal pregnant rats, the bile secretion signaling pathways were significantly increased in GDM rats. Both our results and previous studies suggest that bile acids participate in the development of GDM. Serum bile acids in early pregnancy like deoxycholic acid $\leq$ 0.28 nmol/ml or glycoursodeoxycholic acid $\leq$ 0.07 nmol/ml were significantly associated with increased risk of GDM (Li et al., 2018). Bile acids play important roles in glucose and lipid metabolism (Vítek and Haluzik, 2016). Furthermore, the relationship between bile acid metabolism and the gut microbiome is mutual (Vítek and Haluzik, 2016). On the one hand, gut microbiota can metabolize bile acids. On the other hand, bile acids can affect the composition of gut microbiota (Vítek and Haluzik, 2016). Our results indicated the relationship between gut microbiota and bile acid and the potential influences of their interaction on host glucose and lipid metabolism in GDM.

A balanced structure of gut microbiota and its metabolites are beneficial to the health of the host; however, the poor interaction between the host and its gut microbiota can lead to metabolic dysfunction (Hills et al., 2019). Disturbances in gut microbiota might contribute to GDM pathogenesis by modulating the metabolism signaling pathways of amino acids and bile acids (Wang et al., 2020). For example, gut microbiota perturbations of the *Bacteroides* and *Bifidobacterium* genera displayed strong correlations with amino acid metabolism, and the *Bifidobacterium* genus was also positively associated with bile acid metabolism. *Allobaculum* genus displayed strong positive correlations, whereas the *Bryobacter* and *Gemmatimonas* genera displayed strong negative correlations with amino acid and bile acid metabolism in GDM rats. *Lactobacillus*, *Bifidobacterium*, and *Akkermansia* are widely used to aid in weight loss and ameliorate metabolic diseases such as obesity and diabetes (Park et al., 2019; Yan et al., 2019; Li et al., 2020). *L. rhamnosus* LGG and *B. lactis* Bb12 were used as probiotic supplements to treat GDM model rats in this study. Consistent with the findings from our previous study (Zheng et al., 2021) as well as the present study, a previous randomized controlled trial found that *L. rhamnosus* HN001 supplementation in early pregnancy decreased GDM prevalence and fasting conjugated bile acids, and these bile acids were positively correlated with fasting glucose and fasting insulin, suggesting that they play important roles in improving the glucose metabolism in pregnant women (Chen et al., 2021).

In summary, 16S RNA sequencing and metabolomic analysis conducted in this study revealed the effect of probiotic supplements (*Lactobacillus* and *Bifidobacterium*) on the gut microbiota and gut microbiome of GDM rats and the metabolic status of normal pregnant rats and GDM rats. In addition, this work extended our insights into the relationship among gut microbiota, host metabolism, and GDM rats, pointing to possible future modalities to treat GDM by targeting gut microbiota or gut metabolites. More research is needed to further identify the bacterial species that play a key role in the metabolism of GDM. Overall, this study indicated that probiotic supplementation might be a novel approach to reduce blood glucose and insulin resistance and regulate gut microbiota or gut metabolites in GDM rats. Furthermore, this study provides insights into the applicability of probiotic supplementation in the treatment of GDM.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

All animal procedures were approved by the Fujian University of Traditional Chinese Medicine (Certificate number: SYXK 2019-0007; Ethics approval number: FJTCM IACUC 2020020).

AUTHOR CONTRIBUTIONS

Q-XZ and H-WW analyzed data, plotted figures, drafted, and revised the manuscript. Q-XZ and X-MJ designed this project and performed coordination. LG revised important scholarly content of our manuscript. Y-TL, X-YJ, P-PH, FC, and X-QC performed the rat experiments, collected, and analyzed samples. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.779314/full#supplementary-material

Supplementary Figure 1 | Reciprocal synergistic relationships of metabolites and gut microbiota in GDM rats. **(A)** The co-occurrence network of metabolites and gut microbiota. Yellow points represented gut metabolites, and green points represented gut microbiota. Red lines represented positive correlations, and blue lines represented negative correlations. **(B)** The co-occurrence network of *Lactobacillus* and *Bifidobacterium* genus and metabolites. Yellow points represented gut metabolites, and green points represented gut microbiota. Red lines represented positive correlations.

REFERENCES

Canfora, E. E., Meex, R. C. R., Venema, K., and Blaak, E. E. (2019). Gut microbial metabolites in obesity, NAFLD and T2DM. *Nat. Rev. Endocrinol.* 15, 261–273. doi: 10.1038/s41574-019-0156-z

Chen, Y., Lu, J., Wickens, K., Stanley, T., Maude, R., Stone, P., et al. (2021). Effect of Lactobacillus rhamnosus Probiotic in Early Pregnancy on Plasma Conjugated Bile Acids in a Randomised Controlled Trial. *Nutrients* 13:209. doi: 10.3390/nu13010209

de Brito Alves, J. L., de Oliveira, Y., Carvalho, N. N. C., Cavalcante, R. G. S., Pereira Lira, M. M., Nascimento, L. C. P. D., et al. (2019). Gut microbiota and probiotic intervention as a promising therapeutic for pregnant women with cardiometabolic disorders: Present and future directions. *Pharmacol Res* 145:104252. doi: 10.1016/j.phrs.2019.104252

Gao, C., Sun, X., Lu, L., Liu, F., and Yuan, J. (2019). Prevalence of gestational diabetes mellitus in mainland China: A systematic review and meta-analysis. *J. Diabetes Investig.* 10, 154–162. doi: 10.1111/jdi.12854

He, J., Chan, T., Hong, X., Zheng, F., Zhu, C., Yin, L., et al. (2020). Microbiome and Metabolome Analyses Reveal the Disruption of Lipid Metabolism in Systemic Lupus Erythematosus. *Front. Immunol.* 11:1703. doi: 10.3389/fimmu.2020.01703

Hills, R. D. Jr., Pontefract, A. B., Mishcon, R. H., Black, A. C., Sutton, C. S., and Theberge, R. C. (2019). Gut Microbiome: Profound Implications for Diet and Disease. *Nutrients* 11:1613. doi: 10.3390/nu11071613

Jiang, R., Wu, S., Fang, C., Wang, C., Yang, Y., Liu, C., et al. (2020). Amino acids levels in early pregnancy predict subsequent gestational diabetes. *J. Diabetes* 12, 503–511. doi: 10.1111/1753-0407.13018

Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C. J., Fagerberg, B., et al. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498, 99–103. doi: 10.1038/nature12198

Kurtzhals, L. L., Nørgaard, S. K., Secher, A. L., Nichum, V. L., Ronneby, H., Tabor, A., et al. (2018). The impact of restricted gestational weight gain by dietary intervention on fetal growth in women with gestational diabetes mellitus. *Diabetologia* 61, 2528–2538. doi: 10.1007/s00125-018-4736-6

Li, G. (2019). *Preliminary Study on Intestinal Microflora and Metabolomics in Gestational Diabetes Mellitus.* [Ph.D thesis]. (Zhengzhou, IL: Zhengzhou University).

Li, J., Huo, X., Cao, Y. F., Li, S. N., Du, Z., Shao, P., et al. (2018). Bile acid metabolites in early pregnancy and risk of gestational diabetes in Chinese women: A nested case-control study. *EBioMedicine* 35, 317–324. doi: 10.1016/j.ebiom.2018.08.015

Li, X. P., Huang, Y. M., Song, L. Q., Xiao, Y. C., Lu, S., Xu, J. G., et al. (2020). Lactobacillus plantarum prevents obesity via modulation of gut microbiota and metabolites in high-fat feeding mice. *J. Funct. Foods* 73:104103. doi: 10.1016/j.jff.2020.104103

Liu, R., Hong, J., Xu, X., Feng, Q., Zhang, D., Gu, Y., et al. (2017). Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat. Med.* 23, 859–868. doi: 10.1038/nm.4358

Lu, W., Luo, M., Fang, X., Zhang, R., Li, S., Tang, M., et al. (2021). Discovery of metabolic biomarkers for gestational diabetes mellitus in a Chinese population. *Nutr. Metab.* 18:79. doi: 10.1186/s12986-021-00606-8

Masulli, M., Vitacolonna, E., Fraticelli, F., Della Pepa, G., Mannucci, E., and Monami, M. (2020). Effects of probiotic supplementation during pregnancy on metabolic outcomes: A systematic review and meta-analysis of randomized controlled trials. *Diabetes Res. Clin. Pract.* 162:108111. doi: 10.1016/j.diabres.2020.108111

Olson, C. A., Vuong, H. E., Yano, J. M., Liang, Q. Y., Nusbaum, D. J., and Hsiao, E. Y. (2018). The Gut Microbiota Mediates the Anti-Seizure Effects of the Ketogenic Diet. *Cell* 173, 1728–1741.e13. doi: 10.1016/j.cell.2018.04.027

Ovesen, P. G., Fuglsang, J., Andersen, M. B., Wolff, C., Petersen, O. B., David, et al. (2018). Temporal Trends in Gestational Diabetes Prevalence, Treatment, and Outcomes at Aarhus University Hospital, Skejby, between 2004 and 2016. *J. Diabetes Res.* 2018:5937059. doi: 10.1155/2018/5937059

Pan, Y., Zheng, Q., Jiang, X., Chen, X., Zhang, X., and Wu, J. (2021). Probiotic Supplements Improve Blood Glucose and Insulin Resistance/Sensitivity among Healthy and GDM Pregnant Women: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Evid. Based Complement. Alternat. Med.* 2021:9830200. doi: 10.1155/2021/9830200

Park, S. S., Lee, Y. J., Kang, H., Yang, G., Hong, E. J., Lim, J. Y., et al. (2019). Lactobacillus amylovorus KU4 ameliorates diet-induced obesity in mice by promoting adipose browning through PPARγ signaling. *Sci. Rep.* 9:20152. doi: 10.1038/s41598-019-56817-w

Peng, T. R., Wu, T. W., and Chao, Y. C. (2018). Effect of Probiotics on the Glucose Levels of Pregnant Women: A Meta-Analysis of Randomized Controlled Trials. *Medicina* 54:77. doi: 10.3390/medicina54050077

Ponnusamy, S, Diabetes in Pregnancy Working Group, Maternal Medicine Clinical Study Group, and Royal College of Obstetricians and Gynaecologists (2020). Gestational diabetes: opportunities for improving maternal and child health. *Lancet Diabetes Endocrinol.* 8, 793–800. doi: 10.1016/s2213-8587(20)30161-3

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60. doi: 10.1038/nature11450

Rasmussen, L., Poulsen, C. W., Kampmann, U., Smedegaard, S. B., Ovesen, P. G., and Fuglsang, J. (2020). Diet and Healthy Lifestyle in the Management of Gestational Diabetes Mellitus. *Nutrients* 12:3050. doi: 10.3390/nu12103050

Veeraswamy, S., Vijayam, B., Gupta, V. K., and Kapur, A. (2012). Gestational diabetes: the public health relevance and approach. *Diabetes Res. Clin. Pract.* 97, 350–358. doi: 10.1016/j.diabres.2012.04.024

Vítek, L., and Haluzík, M. (2016). The role of bile acids in metabolic regulation. *J. Endocrinol.* 228, R85–R96. doi: 10.1530/JOE-15-0469

Wang, X., Liu, H., Li, Y., Huang, S., Zhang, L., Cao, C., et al. (2020). Altered gut bacterial and metabolic signatures and their interaction in gestational diabetes mellitus. *Gut. Microbes.* 12, 1–13. doi: 10.1080/19490976.2020.1840765

Yan, F., Li, N., Shi, J., Li, H., Yue, Y., Jiao, W., et al. (2019). Lactobacillus acidophilus alleviates type 2 diabetes by regulating hepatic glucose, lipid metabolism and gut microbiota in mice. *Food Funct.* 10, 5804–5815. doi: 10.1039/c9fo01062a

Yao, K., Zeng, L., He, Q., Wang, W., Lei, J., and Zou, X. (2017). Effect of Probiotics on Glucose and Lipid Metabolism in Type 2 Diabetes Mellitus: A Meta-Analysis of 12 Randomized Controlled Trials. *Med. Sci. Monit.* 23, 3044–3053. doi: 10.12659/msm.902600

Zhang, H., Dellsperger, K. C., and Zhang, C. (2012). The link between metabolic abnormalities and endothelial dysfunction in type 2 diabetes: an update. *Basic Res. Cardiol.* 107:237. doi: 10.1007/s00395-011-0237-1

Zhang, J., Ma, S., Wu, S., Guo, C., Long, S., and Tan, H. (2019). Effects of Probiotic Supplement in Pregnant Women with Gestational Diabetes Mellitus: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *J. Diabetes Res.* 2019:5364730. doi: 10.1155/2019/5364730

Zhao, C., Ge, J., Li, X., Jiao, R., Li, Y., Quan, H., et al. (2020). Integrated metabolome analysis reveals novel connections between maternal fecal metabolome and the neonatal blood metabolome in women with gestational diabetes mellitus. *Sci. Rep.* 10:3660. doi: 10.1038/s41598-020-60540-2

Zheng, Q., Jiang, X., Wang, H., Ge, L., Lai, Y., Jiang, X., et al. (2021). Probiotic supplements alleviate gestational diabetes mellitus by restoring the diversity of gut microbiota: a study based on 16S rRNA sequencing. *J. Microbiol.* 59, 827–839. doi: 10.1007/s12275-021-1094-8

frontiers | Frontiers in Microbiology

Imaging Technologies Build Capacity and Accessibility in Phytoplankton Species Identification Expertise for Research and Monitoring: Lessons Learned During the COVID-19 Pandemic

Sophie Clayton[1]*, Leah Gibala-Smith[1,2], Kathryn Mogatas[1], Chanel Flores-Vargas[1], Kayla Marciniak[3], Maci Wigginton[1] and Margaret R. Mulholland[1]

[1] Department of Ocean and Earth Sciences, Old Dominion University, Norfolk, VA, United States, [2] New Jersey Department of Environmental Protection, Trenton, NJ, United States, [3] Department of Biology, Old Dominion University, Norfolk, VA, United States

As primary producers, phytoplankton play an integral role in global biogeochemical cycles through their production of oxygen and fixation of carbon. They also provide significant ecosystem services, by supporting secondary production and fisheries. Phytoplankton biomass and diversity have been identified by the Global Ocean Observing System (GOOS) as Essential Ocean Variables (EOVs), properties that need to be monitored to better understand and predict the ocean system. Phytoplankton identification and enumeration relies on the skills and expertise of highly trained taxonomic analysts. The training of new taxonomic analysts is intensive and requires months to years of supervised training before an analyst is able to independently and consistently apply identification skills to a sample. During the COVID-19 pandemic, access to laboratories was greatly restricted and social distancing requirements prevented supervised training. However, access to phytoplankton imaging technologies such as the Imaging FlowCytobot (IFCB), FlowCam, and PlanktoScope, combined with open online taxonomic identification platforms such as EcoTaxa, provided a means to continue monitoring, research, and training activities remotely when in-person activities were restricted. Although such technologies can not entirely replace microscopy, they have a great potential for supporting an expansion in taxonomic training, monitoring, surveillance, and research capacity. In this paper we highlight a set of imaging and collaboration tools and describe how they were leveraged during laboratory lockdowns to advance research and monitoring goals. Anecdotally, we found that the use of imaging tools accelerated the training of new taxonomic analysts in our phytoplankton analysis laboratory. Based on these experiences, we outline how these technologies can be used to increase capacity in taxonomic training and expertise, as well as how they can be used more broadly to expand research opportunities and capacity.

Keywords: phytoplankton, taxonomic expertise, harmful algal bloom (HAB), environmental monitoring, phytoplankton ecology, imaging cytometry, species identification, plankton

CHALLENGES IN PHYTOPLANKTON MONITORING AND TRAINING NEW TAXONOMIC ANALYSTS

Expertise in marine phytoplankton taxonomy is largely restricted to laboratories involved in programs geared toward water quality monitoring, harmful algal bloom (HAB) surveillance, and other taxa-specific research and assessments (Stauffer et al., 2019). However, in recent years, it has become apparent that taxa-specific information is crucial to our understanding of carbon flow in aquatic systems and how this may be altered in the future due to climate change and concomitant changes in phytoplankton community structure. Indeed, the Global Ocean Observing System (GOOS) has recognized phytoplankton biomass and diversity as Essential Ocean Variables (EOVs), key variables that need to be monitored to better understand and predict the ocean system (Lindstrom et al., 2012). Surveillance and monitoring for the presence and abundance of HABs has brought many new imaging technologies to the forefront as they generate invaluable data for detecting, tracking and forecasting HAB events (Harred and Campbell, 2014; Stauffer et al., 2019). HABs are forecast to increase in frequency and duration with climate change (Najjar et al., 2010; Gobler et al., 2017; Griffith and Gobler, 2020) highlighting the need to develop better methods for early detection of HAB events and their initiation. Despite the importance of understanding taxon-specific physiological capacity and capabilities to predict how phytoplankton communities are likely to change as a result of climate change and other stressors, there has been a decline in the number of trained taxonomists and analysts (McQuatters-Gollop et al., 2017).

The Need to Increase Capacity in Phytoplankton Species Identification Training

Skilled phytoplankton taxonomic analysts typically undergo months of closely supervised training that is generally biogeographically specific. However, since the onset of the genomics era and satellite ocean color surveillance, there has been a decline in the number of taxonomic analysts, a reduction in the number of new trainees entering the pipeline (Drew, 2011; Pearson et al., 2011; McQuatters-Gollop et al., 2017), and a shift toward molecular detection of species of interest (HARRNESS, 2005; Jewett et al., 2007; HAB RDDTT, 2008). The decline in taxonomic training has been noted by numerous US federal agencies as an impediment to long-term monitoring and tracking of HABs, and tracking changes in phytoplankton phenology and biogeography. Efforts to support training and development of taxonomic skills are urgently needed to ensure effective management, support continued surveillance for HABs, detect shifts in phytoplankton communities that may impact ecosystem structure and function, and the emergence of HAB species and shifts in the biogeography of phytoplankton communities as a result of climate change.

Taxonomic identification and enumerations of phytoplankton are crucial for early warning programs for marine and freshwater HABs, and highly trained experts are an essential part of such monitoring programs. However, classical taxonomic training can be time-consuming and costly as it requires a trained expert hours to analyze a single sample. Further, even expert taxonomists can diverge in their identifications of phytoplankton species and groups. Modern imaging technologies have been deployed on moorings or used to sample underway from ships in some regions in support of research (e.g., Martha's Vineyard Coastal Observatory[1]; North East U.S. Shelf Long-Term Ecological Research[2]) and for early detection of HABs (e.g., Harred and Campbell, 2014) but these technologies are expensive and data are limited to either the sites where, or time periods when instruments are deployed. Although light microscopy can yield a higher taxonomical resolution of phytoplankton community composition and, in some cases, more exact species determinations than imaging methods, imaging methods have the advantage that they can produce a large number of images at higher spatial and temporal frequency, and are more geared toward the archiving and re-use of cell images. Molecular ecology tools have also been employed to identify and quantify target species for which appropriate primers are available, but this approach is generally limited to one or a few species of interest and has limited ability to produce quantitative assessments of community structure. Here we propose an alternative method for training taxonomic analysts that combines the use of images generated by automated imaging systems with more traditional techniques, to accelerate training and expand capacity. Additionally, we highlight how image libraries generated through the use of imaging tools can be used to harmonize species identification across laboratories, and provide a spatially— and temporally resolved data archive of phytoplankton community composition that can be re-interrogated over time as research questions evolve.

Lessons Learned From the COVID Lockdown: An Accidental Experiment Leading to Adaptations to Workflow

The Phytoplankton Analysis Laboratory at Old Dominion University (referred to hereafter as the ODU Phyto Lab) has been in operation since the early 1980s and undertakes a range of taxonomic analyses in support of research projects as well as local, state and federal agencies involved in the Chesapeake Bay restoration, and shellfish and public health monitoring in coastal and freshwater bodies in the Commonwealth of Virginia (VA). Although the ODU Phyto Lab is managed by full-time experts in phytoplankton taxonomic analysis, it also relies on undergraduate and graduate student interns rotating through the laboratory who must be trained in phytoplankton taxonomy and microscopy. Trainees are expected to conduct microscope analysis on samples collected from monitoring stations associated with a long-term monitoring project through the Chesapeake Bay Program and shellfish growing areas within waters of Virginia. This requires species-level understanding, with special attention paid to a suite of seasonal bloomers, including the potentially toxigenic HABs *Pseudo-nitzschia* spp. and *Dinophysis* spp. Up

---

[1]https://ifcb-data.whoi.edu/timeline?dataset=mvco
[2]https://nes-lter.whoi.edu/data/

until the onset of the COVID-19 pandemic, this training involved laboratory staff working at close quarters with trainees and having access to microscopes and slide preparation materials in the laboratory.

An Imaging FlowCytobot (IFCB; Olson and Sosik, 2007) was acquired in 2019, and samples being collected as part of a local time series site (Mulholland et al., 2018) were used to begin to build an image database and prepare the instrument for deployment. When the ODU Phyto Lab's regular taxonomic analysis training model was interrupted by the COVID-19 pandemic, the image library produced by the IFCB was leveraged to implement and test different training and analysis methods and tools. During the COVID-19 pandemic, many laboratories including the ODU Phyto Lab, were shut down and/or required to practice social distancing. As a result, taxonomic training using microscopes with trainer and trainee in close proximity in the laboratory became impossible despite the need for continuity in the monitoring work on which state and federal entities depend. It became necessary to develop new, socially distanced ways to continue the ODU Phyto Lab's monitoring, research, and training activities that had until then relied solely on the ability of staff to work collaboratively on shared microscopes.

Although access to the laboratory during the early part of the pandemic was restricted, it was still possible for a single individual to collect samples in the field, run those samples through an imaging instrument in the laboratory, and process and upload the images to EcoTaxa (Picheral et al., 2017[3]), an open access web application for the taxonomic annotation of images of plankton. The water samples collected and images processed prior to the COVID-19 lockdown also provided a large existing image bank that could be used for training while field sampling was suspended. Using this image bank, students were trained to identify a subset of the taxa present in the phytoplankton community for their individual projects. This approach proved to have many advantages over the more traditional microscope-based training methods which requires working at close quarters in the laboratory, direct supervision by the trainer, and complete identification and enumeration of the phytoplankton community within a sample before it is discarded. Here we outline how this approach might, in conjunction with traditional microscopy methods, modernize taxonomic training and phytoplankton analyses to build capacity in taxonomic expertise for research, speed up the detection of and response to HAB events, expand monitoring capacity by increasing the frequency and spatial coverage of sampling, and create databases that can be re-interrogated as our understanding of phytoplankton communities evolve and climate changes.

SHORT OVERVIEW OF TECHNOLOGIES AND TOOLS USED IN THIS CASE STUDY

Several cell imaging instruments have been developed over the last decade and a handful have matured to the point where they are commercially available and routinely used as part of HAB

and phytoplankton monitoring programs and research. These technologies have been reviewed in detail elsewhere (Lombard et al., 2019; Stauffer et al., 2019), but here we will briefly describe three imaging instruments that have been in use by the ODU Phyto Lab since January 2020, and that form the basis of the recommendations outlined here. The ODU Phyto Lab has access to an IFCB; several PlanktoScopes, low-cost imaging microscopes which have been built in-house (Pollina et al., 2020[4]); and a FlowCam (Sieracki et al., 1998) loaned to the ODU Phyto Lab for 3 months through the manufacturer's grant program. The IFCB was used to image cells in raw seawater samples, whereas the PlanktoScopes and FlowCam were used to generate images from both preserved (generally using Lugol's) and unpreserved water samples.

Images collected using the IFCB, FlowCam and PlanktoScope were uploaded to EcoTaxa, an open access online web application for the taxonomic annotation of images of plankton, and annotated/classified by the ODU Phyto Lab and two undergraduate interns, taxonomic analysts in training. Images collected using the IFCB are processed using a set of publicly available MATLAB-based tools,[5] which segment the images into individual regions of interest (ROIs) and generate a set of geometric parameters based on the ROIs (e.g., major axis length, minor axis length). Images produced using a PlanktoScope are processed in a similar way using MorphoCut, a publicly available Python-based analysis pipeline which can be run directly on the PlanktoScope's onboard Raspberry Pi computer to segment images into ROIs and generate descriptive geometric parameters.[6] FlowCam provides proprietary software called VisualSpreadsheet for analyzing and extracting data from FlowCam images. Although previous versions of VisualSpreadsheet allowed for FlowCam images to be uploaded to EcoTaxa, the new version of the software uses a very different file structure, and so FlowCam images cannot be directly uploaded for analysis, however it is possible to work around this limitation using MorphoCut. After sample analysis, images generated by the IFCB, FlowCam, and PlanktoScope, their geometric characteristics and associated metadata can be uploaded to EcoTaxa for annotation, classification and validation. Although other similar pipelines exist for image annotation, classification and validation, the main advantage of EcoTaxa for use during the pandemic lockdowns was the fact that it is a web-hosted application that can be accessed from anywhere with an internet connection. As a result it was not necessary for laboratory personnel or student interns to work in shared spaces, use laboratory computers or even to log in to institutional servers to work with or train on the image data. EcoTaxa also allows for the tracking of annotations by different users, allowing for a highly iterative approach to both training and sample analysis, as feedback and corrections to annotations could be done asynchronously as needed. Another big advantage of the web-based application was that training could be done synchronously but remotely, with trainers and trainees communicating using

---

[3]http://ecotaxa.obs-vlfr.fr

[4]https://www.planktoscope.org/

[5]https://github.com/hsosik/ifcb-analysis/wiki

[6]https://github.com/morphocut/morphocut

video conferencing tools while accessing the same image bank without needing to be physically co-located. This combination of imaging instruments and an easily accessible online platform are the basis of the training model described below.

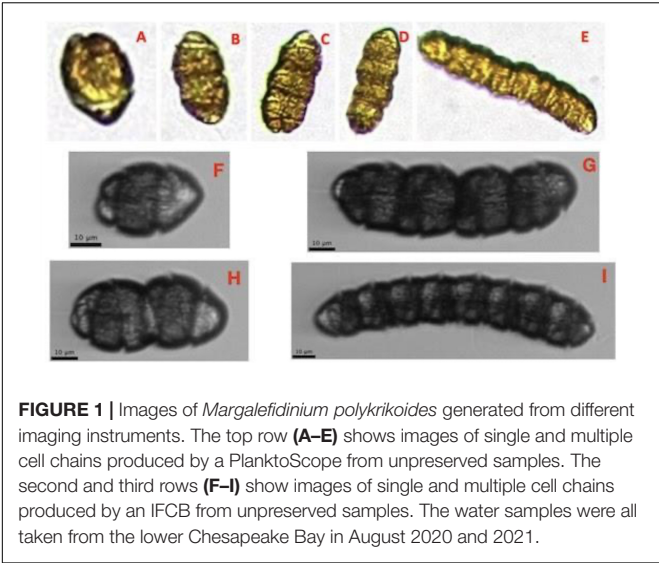# AN IMAGING TECHNOLOGY-SUPPORTED MODEL FOR SPECIES IDENTIFICATION TRAINING

Using images and online sharing tools, we developed a set of guidelines for using technologies to train students in phytoplankton species identification, following a traditional sequence. Experienced analysts within the ODU Phyto Lab identified a set of images from an ongoing (or current) investigation for a select group of taxa that occured in large numbers within the image collection. These categorized images, along with traditional taxonomic reference materials, were provided to trainees to allow them gain familiarity with the morphological features used to identify those specific taxa. Initially trainees were taught to recognize cell shapes, leading to the initial classification of functional groups, e.g., classify cells as dinoflagellates or diatoms. Subsequently trainees were gradually taught to recognize and classify taxa within a given functional group, e.g., centric vs. pennate diatoms. Experienced analysts and trainees, using Zoom and share screen functions with the EcoTaxa project page in view, reviewed targeted taxa structure and how to apply traditional microscope identification skills to reviewing the images in the EcoTaxa project. The trainees were then assigned the task of finding further images of those taxa in the project's image collection. The trainers could then check the trainees' annotations to make sure that they were identifying the cells in the images correctly. When difficulties arose, experienced analysts were able to quickly respond to questions, and teach the trainees how to better differentiate between similar looking groups either by real-time communication *via* Zoom or emailing images for asynchronous responses. In addition, as the trainee would scan the image collection looking for their targeted taxa, they became familiar with other groups of organisms and were able to increase the number of taxa they could accurately and correctly identify. Almost more importantly, trainees were able to quickly identify a cell or group of cells that did not fall into a group they were familiar with, set them aside into a single temporary holding category area, and use their newly acquired skills to independently use the resources available to them to determine their correct taxonomic group. This skill developed much more rapidly with our "COVID experiment" trainees than had ever occurred with traditionally trained students in the 6 years prior, within days or a few weeks as opposed to many weeks or months.

During this time, ODU Phyto Lab, trainees were tasked with identifying particular subgroups of phytoplankton taxa while also contributing to annotating images of other taxa they were familiar with, thereby enhancing both their taxonomic analysis expertise and the skill of machine learning algorithms used for automated classification. For example, after gaining a broad but superficial mastery of the shapes of different phytoplankton functional groups, trainees were then given more in-depth training for only one of the those functional groups, with some trainees focusing on dinoflagellates, others on centric diatoms, and yet others on pennate diatoms. In this way, the identification of phytoplankton in samples was distributed across trainees and laboratory personnel with more senior and experienced taxonomists able to act as a check on trainees' image classifications. Trainees did not need to have a broad taxonomic knowledge to start contributing meaningfully to the identification work because images could be revisited as machine learning algorithms were refined. Anecdotally, we saw shorter training periods when incorporating imaging into our traditional training method, with trainees trained with imaging able to transition to working and analyzing microscopy samples independently after around 3 months, about half the usual 6 month training period.

As outlined above, the use of imaging technology for species identification training is an iterative process that relies on the images being collected, some of them annotated (e.g., identified and labeled with an associated taxonomic classification), analyzed using a machine learning classification algorithm and then archived, allowing for them to be revisited at a later date. When working with samples under a microscope, unless images are collected manually, no visual record remains after analysis. Thanks to the automated saving and archiving of phytoplankton images by imaging technologies, their use allows for some specific and important differences and advantages over light microscopy for training:

1. People can be trained quickly and remotely, as compared with traditional microscopy techniques. This represents a real potential for building capacity in the field. Accuracy can be improved through online resources and image libraries.
2. Training does not need to be "all at once," that is, trainees can learn to identify specific taxa or functional groups without needing to be familiar with all of the taxa that might be found in a sample. Trainees can contribute to a laboratory's analysis work at an earlier stage of their training.
3. Using a web-based application such as EcoTaxa, user annotations and validations are tracked, so feedback between the trainer and trainee can be asynchronous, as the trainer does not need to check a physical sample before it can be discarded.
4. Curated image libraries for reference and training can be built and added to over time (as has been done with data from the IFCB deployment at MVCO, Sosik et al., 2015). Regional and international image libraries can be developed in collaboration with other laboratories around the world to ensure consistency in identification and to help standardize taxonomy regionally and globally. Images from any platform, including light microscopy, can (and should) be included in such libraries.
5. Image libraries will include images of cells from different orientations as well as different life stages (e.g., resting cysts, multiple cell chains, **Figure 1**). This can lead to new information and assist in harmonized species identifications and estimations of biovolumes (Orr et al., 2021).

FIGURE 1 | Images of *Margalefidinium polykrikoides* generated from different imaging instruments. The top row **(A–E)** shows images of single and multiple cell chains produced by a PlanktoScope from unpreserved samples. The second and third rows **(F–I)** show images of single and multiple cell chains produced by an IFCB from unpreserved samples. The water samples were all taken from the lower Chesapeake Bay in August 2020 and 2021.

In summary, with imaging techniques, sample imaging and analysis can be asynchronous, and trainees do not need to learn the entire phytoplankton species composition all at once. They can focus on certain taxa or groups leaving the full dataset to be interrogated iteratively later. Mastering species identification can take years and this method of training allows for students to reliably identify specific constituents of the phytoplankton community rather than the full taxonomic range that might be encountered, thereby allowing trainees to contribute to phytoplankton community composition evaluations of samples in meaningful ways immediately. We also reiterate that, in our admittedly limited case study, we found that including imaging in

training resulted in an appreciable reduction in that time needed for a trainee to become proficient enough in species identification to work independently.

EXPANDING HARMFUL ALGAL BLOOM MONITORING AND PHYTOPLANKTON RESEARCH WITH IMAGING TECHNOLOGIES

Many of the advantages of using imaging for phytoplankton species identification training can also be leveraged to enable advances in phytoplankton ecology research and HAB surveillance and monitoring applications. In both cases, research and monitoring are hampered by a scarcity of taxonomically resolved field data in time and space. Because field samples can be processed and imaged relatively quickly using the imaging platforms described above, live cells from many samples can be imaged and the data stored for analysis at a later date, thereby increasing sample throughput. There is a tradeoff between resolution, which is generally better using light microscopy, and sample throughput, which is higher for imaging instruments. However we do note that the ability to quickly image live cells does have the advantage that taxa such as coccolithophores or naked flagellated species that might degrade or be altered when administered preservatives such as Lugol's solution or formalin (Throndsen, 1978) can be captured by imaging instruments. As a result, the wider adaptation of imaging technologies will greatly increase the volume of data available for a wide range of research and monitoring activities. Thanks to the large number of images produced, machine learning algorithms can be trained and

TABLE 1 | Summary table of the respective advantages of imaging and microscopy for different applications requiring taxonomic identification of phytoplankton.

	APPLICATION		
	Species Identification Training	**HAB Tracking and Monitoring**	**Phytoplankton Ecology**
IMAGING	• Speeds up the training process. • Trainees can work with samples non-destructively. • Trainees can be tasked with identification of a subset of species of interest. • Trainers and trainees can interact and work remotely and asynchronously.	• Higher frequency of sampling in time and space is possible. • Enables inter-calibration between regional and global laboratories. • Allows rapid counts of a small suite of target species without losing information on other species present. • Allows for real time results to inform regulatory actions even if field staff or citizen scientists are not trained taxonomic analysts (i.e., through automated classification of images).	• Can generate large image libraries which enables regional and global inter-comparisons. • Images can be archived for re-examination and analysis at a later date. • Allows for easier estimation of cell size and biovolume, useful for model validation and comparison. • Images cells from multiple orientations allowing more accurate estimates of biovolume.
MICROSCOPY	• Allows for better optical resolution and precise focus to investigate morphological features of a cell that occur on different focal planes. • Allows for the use of specialized methods, such as fluorescent dyes to facilitate identification based on specific morphological features.	• Allows for detection of emerging HAB taxa that may be unexpected in the area or not previously detected. • Allows for better optical resolution and precise focus to investigate different focal planes. • Allows for the use of specialized methods, such as fluorescent dyes.	• Allows for more reliable enumeration of rare species as sample volumes may be larger. • Allows for finer distinction between similarly structured species. • Allows for better resolution of small celled taxa ($<5 \mu m$) that are mostly missed by imaging instruments.

used to automate the identification of phytoplankton cells allowing for a much greater number of samples to be enumerated than by microscopy alone. Further, through the sharing of image databases, taxonomic identifications can be harmonized regionally and globally.

The development of low-cost imaging instruments such as the PlanktoScope which we have described above, as well as other platforms such as the HABscope[7] which has been designed for use by trained volunteers, is opening up the possibility of generating near real-time cell counts directly from field sampling locations. These low-cost instruments can be provided to stakeholders such as agency field staff or shellfish farmers who can run samples in the field and upload the data to a central repository such as EcoTaxa, assuming that they have internet access. The resulting images are then classified using machine learning models trained on locally generated image libraries, and some may also be inspected remotely by expert taxonomists to ensure the performance of the classification algorithm. This could allow for a significant expansion of HAB monitoring, generating higher frequency data in both time and space, and speeding up the time between sample collection and identification. HABscope has already done this for monitoring and detection of *Karenia brevis* in the Gulf of Mexico (Hardison et al., 2019).

Establishing and maintaining image libraries either regionally or globally will be key to capitalizing on many of the novel advantages of plankton imaging technologies. We envisage that such libraries will allow for regional tracking of climate change impacts on specific phytoplankton communities, for example, enabling the detection and tracking of invasive species. Similarly, image libraries could be used to identify, assess, and track temporal or regional physiological differences (e.g., cell major and minor axis lengths, mean cell sizes, biovolume, chain length, etc.) among populations and communities. This will be of particular relevance to the development of ecological and biogeochemical models, whose performance can be difficult to assess due to lack of data on specific taxa of interest (e.g., HAB prediction models), and difficulties in converting cell densities (e.g., cells L^{-1}) reported from microscopic enumeration to biomass concentrations (mmol C m^{-3}) used in ecological models.

CONCLUSION

In this paper we have outlined ways in which imaging tools and online data sharing platforms can be used to increase capacity in phytoplankton species identification expertise, and expand monitoring capacity, summarized in **Table 1**. The experience that trainees obtained using the image library prior to being introduced to the phytoplankton community in samples on the microscope allowed them to reach a level of proficiency rapidly, thus increasing workflow and laboratory productivity. Although anecdotal, our experience has shown that imaging technologies can be used to shorten by up to half the period of training required before trainees can make meaningful contributions

---

[7] https://habscope.gcoos.org/

to generating robust species identifications for monitoring and research purposes. Imaging technologies and open online sharing platforms allowed us to completely and successfully change our species identification training model during the height of the COVID-19 pandemic and continue monitoring for state and local agencies as well as driving new research regarding the life stages of dinoflagellates. As a result of the pandemic experience, we have identified ways in which imaging technologies can be used to accelerate and increase efficiency in phytoplankton species identification training, as well as enhancing sample throughput and HAB monitoring, and facilitating regional and global inter-comparison studies, and tracking regional and global intraspecific variability.

Although we have outlined many advantages of using imaging technologies for phytoplankton species identification training and research, we do not advocate that imaging tools should completely replace microscopy. There are particular applications that imaging is not well suited for, in particular for achieving high taxonomic resolution in identifications, which light microscopy allows thanks to the higher optical resolution and ability to scrutinize the cell at different focal levels. Imaging is also poorly suited to identifying and enumerating picophytoplankton and nanophytoplankton that are smaller than $\sim 5\,\mu m$ and not always detected by imaging instruments. While imaging coupled with automated classification tools can process a larger number of samples than microscopy, microscopy has been shown to achieve higher taxonomic resolution (Álvarez et al., 2014). We also recognize that not all laboratories have the resources required to implement the use of imaging in their training or analysis work, due to the expense of imaging instruments such as the IFCB or FlowCam. These instruments, although useful for generating very large quantities of images, are not required to apply some of the workflow described here. The important step is to generate images that can then be shared with trainees in some way. This could be achieved with light microscopes equipped with cameras (which are commonly found in laboratories), or by using cheap open-source instruments such as the PlanktoScope imaging microscope that can be built by non-experts for < $1,000 each. Finally, we reiterate that we do not advocate that imaging technologies should replace microscopy, but rather that both can be used in concert to expand capacity in taxonomic analysis and through that expanded capacity, drive new advances in the fields of phytoplankton taxonomy and ecology.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

SC, MM, and LG-S contributed to the conception of the manuscript. SC, LG-S, KMo, and MM drafted the manuscript. All

authors contributed to manuscript revision, read, and approved the submitted version.

REFERENCES

Álvarez, E., Moyano, M., López-Urrutia, Á, Nogueira, E., and Scharek, R. (2014). Routine determination of plankton community composition and size structure: a comparison between FlowCAM and light microscopy. *J. Plankt. Res.* 36, 170–184. doi: 10.1093/plankt/fbt069

Drew, L. (2011). Are We Losing the Science of Taxonomy? *BioScience* 61, 942–946. doi: 10.1525/bio.2011.61.12.4

Gobler, C. J., Doherty, O. M., and Theresa, K. (2017). Hattenrath-Lehmann, Andrew W. Griffith, Yoonja Kang, and R. Wayne Litaker. "Ocean warming since 1982 has expanded the niche of toxic algal blooms in the North Atlantic and North Pacific oceans." *Proc. Natl. Acad. Sci.* 114, 4975–4980. doi: 10.1073/pnas.1619575114

Griffith, A. W., and Gobler, C. J. (2020). Harmful algal blooms: a climate change co-stressor in marine and freshwater ecosystems. *Harmful Algae* 91:101590. doi: 10.1016/j.hal.2019.03.008

HAB RDDTT (2008). *Harmful Algal Bloom Research, Development, Demonstration, and Technology Transfer National Workshop Report* (Woods Hole, Massachusetts: National HAB Committee; Woods Hole Oceanographic Institute).

Hardison, D. R., Holland, W. C., Currier, R. D., Kirkpatrick, B., Stumpf, R., Fanara, T., et al. (2019). HABscope: A tool for use by citizen scientists to facilitate early warning of respiratory irritation caused by toxic blooms of Karenia brevis. *PLoS One* 14:e0218489. doi: 10.1371/journal.pone.0218489

Harred, L. B., and Campbell, L. (2014). Predicting harmful algal blooms: a case study with Dinophysis ovum in the Gulf of Mexico. *J. Plankt. Res.* 36, 1434–1445. doi: 10.1093/plankt/fbu070

HARRNESS (2005). *Harmful Algal Research and Response: A National Environmental Science Strategy 2005–2015*. (Washington DC: Ecological Society of America), 96.

Jewett, E. B., Lopez, C. B., Dortch, Q., and Etheridge, S. M. (2007). *National Assessment of Efforts to Predict and Respond to Harmful Algal Blooms in U.S. Waters. Interim Report*. Washington, DC: Interagency Working Group on Harmful Algal Blooms, Hypoxia, and Human Health of the Joint Subcommittee on Ocean Science and Technology.

Lindstrom, E., Gunn, J., Fischer, A., McCurdy, A., and Glover, L. K. (2012). *A Framework for Ocean Observing*. (Paris: UNESCO, IOC Information Document), 1284. doi: 10.5270/OceanObs09- FOO

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). Globally consistent quantitative observations of planktonic ecosystems. *Front. Marin. Sci.* 6:196. doi: 10.3389/fmars.2019.00196

McQuatters-Gollop, A., Johns, D. G., Bresnan, E., Skinner, J., Rombouts, I., Stern, R., et al. (2017). From microscope to management: the critical value of plankton taxonomy to marine policy and biodiversity conservation. *Marin. Policy* 83, 1–10. doi: 10.1016/j.marpol.2017.05.022

Mulholland, M. R., Morse, R., Egerton, T., Bernhardt, P. W., and Filippino, K. C. (2018). Blooms of dinoflagellate mixotrophs in a lower Chesapeake Bay tributary: carbon and nitrogen uptake over diurnal, seasonal, and interannual timescales. *Estuar. coasts* 41, 1744–1765. doi: 10.1007/s12237-018-0 388-5

Najjar, R., Pyke, C., Adams, M. B., Breitburg, D., Hershner, C., Kemp, W. M., et al. (2010). Potential climate-change impacts on the Chesapeake Bay. *Estuar. Coast. Shelf Sci.* 80, 1–20. doi: 10.1016/j.ecss.2009.09.026

Olson, R. J., and Sosik, H. M. (2007). A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot. *Limnol. Oceanograph.* 5, 195–203. doi: 10.4319/lom.2007.5.195

Orr, M. C., Ferrari, R. R., Hughes, A. C., Chen, J., Ascher, J. S., Yan, Y. H., et al. (2021). Taxonomy must engage with new technologies and evolve to face future challenges. *Nat. Ecol. Evol.* 5, 3–4. doi: 10.1038/s41559-020-01 360-5

Pearson, D. L., Hamilton, A. L., and Erwin, T. L. (2011). Recovery plan for the endangered taxonomy profession. *BioScience* 61, 58–63. doi: 10.1525/bio.2011.61.1.11

Picheral, M., Colin, S., and Irisson, J.-O. (2017). EcoTaxa, a Tool for the Taxonomic Classification of Images. Available online at: http://ecotaxa.obs-vlfr.fr

Pollina, T., Larson, A. G., Lombard, F., Li, H., Colin, S., de Vargas, C., et al. (2020). PlanktonScope: Affordable modular imaging platform for citizen oceanography. *bioRxiv* [Preprint], doi: 10.1101/2020.04.23.05 6978

Sieracki, C. K., Sieracki, M. E., and Yentsch, C. S. (1998). An imaging-in-flow system for automated analysis of marine microplankton. *Marin. Ecol. Prog. Series* 168, 285–296. doi: 10.3354/meps168285

Sosik, H. M., Peacock, E. E., and Brownlee, E. F. (2015). WHOI-Plankton, annotated plankton images - data set for developing and evaluating classification methods. *arXiv* [Preprint], doi: 10.1575/1912/7341

Stauffer, B. A., Bowers, H. A., Buckley, E., Davis, T. W., Johengen, T. H., Kudela, R., et al. (2019). Considerations in harmful algal bloom research and monitoring: perspectives from a consensus-building workshop and technology testing. *Front. Marin. Sci.* 6:399. doi: 10.3389/fmars.2019.00399

Throndsen, J. (1978). "Preservation and storage," in *Phytoplankton Manual*, ed. A. Sournia (Paris: UNESCO), 69–74.

Cas1 and Fen1 Display Equivalent Functions During Archaeal DNA Repair

Julia Wörtz[1], Victoria Smith[2], Jörg Fallmann[3], Sabine König[4,5], Tharani Thuraisingam[6], Paul Walther[6], Henning Urlaub[4,5], Peter F. Stadler[3,7,8,9,10,11,12,13], Thorsten Allers[2], Frank Hille[1] and Anita Marchfelder[1]*

[1] Biology II, Ulm University, Ulm, Germany, [2] School of Life Sciences, University of Nottingham, Nottingham, United Kingdom, [3] Department of Computer Science, Bioinformatics Group, Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, Germany, [4] Bioanalytical Mass Spectrometry Group, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany, [5] Institute of Clinical Chemistry, University Medical Center Göttingen, Göttingen, Germany, [6] Central Facility for Electron Microscopy, Ulm University, Ulm, Germany, [7] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany, [8] Competence Center for Scalable Data Services and Solutions, Leipzig Research Center for Civilization Diseases, University Leipzig, Leipzig, Germany, [9] Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia, [10] Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria, [11] Center for RNA in Technology and Health, University of Copenhagen, Copenhagen, Denmark, [12] Santa Fe Institute, Santa Fe, NM, United States, [13] Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

CRISPR-Cas constitutes an adaptive prokaryotic defence system against invasive nucleic acids like viruses and plasmids. Beyond their role in immunity, CRISPR-Cas systems have been shown to closely interact with components of cellular DNA repair pathways, either by regulating their expression or via direct protein-protein contact and enzymatic activity. The integrase Cas1 is usually involved in the adaptation phase of CRISPR-Cas immunity but an additional role in cellular DNA repair pathways has been proposed previously. Here, we analysed the capacity of an archaeal Cas1 from *Haloferax volcanii* to act upon DNA damage induced by oxidative stress and found that a deletion of the *cas1* gene led to reduced survival rates following stress induction. In addition, our results indicate that Cas1 is directly involved in DNA repair as the enzymatically active site of the protein is crucial for growth under oxidative conditions. Based on biochemical assays, we propose a mechanism by which Cas1 plays a similar function to DNA repair protein Fen1 by cleaving branched intermediate structures. The present study broadens our understanding of the functional link between CRISPR-Cas immunity and DNA repair by demonstrating that Cas1 and Fen1 display equivalent roles during archaeal DNA damage repair.

Keywords: CRISPR-Cas, Cas1, DNA repair, Fen1, archaea, *Haloferax volcanii*

INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR) and their associated genes (Cas) comprise a prokaryotic immune system, which defends bacteria and archaea from predatory mobile genetic elements, including viruses and plasmids (Hille et al., 2018; Nussenzweig and Marraffini, 2020). Its unique adaptive nature is enabled by proteins Cas1 and Cas2, which form

a complex that integrates short pieces of invading DNA, called spacers, into the chromosomal CRISPR array (Barrangou et al., 2007; Nuñez et al., 2014). Transcription of the CRISPR array generates a long precursor CRISPR-RNA (crRNA), which is processed by Cas proteins or cellular RNases into short mature crRNAs, each harbouring the sequence of a previously acquired spacer (Brouns et al., 2008). Cas effector nucleases bind and utilise crRNAs to sequence-specifically target and cleave complementary sequences, causing the degradation of invading DNA, thus protecting the cell (Barrangou et al., 2007).

Besides their role in prokaryotic immunity, the involvement of CRISPR-Cas systems in cellular functions beyond anti-viral defence has been subject of multiple studies (Faure et al., 2019a,b). One such function is the contribution to the repair of chromosomal DNA damage. In fact, CRISPR-Cas systems were first hypothesised to represent a novel DNA repair system due to similarity of enzymatic domains in Cas proteins with those found in DNA repair proteins like DNA helicases and nucleases such as RecB (Makarova et al., 2002, 2006). While this assumption has now been revised, a functional link between components of DNA repair pathways and CRISPR-Cas systems has been reported previously; during spacer acquisition, the DNA repair complexes RecBCD and AddAB (in Gram-negative and Gram-positive bacteria, respectively) are able to degrade invading DNA, thereby providing Cas1-Cas2 with DNA fragments for integration (Marraffini and Sontheimer, 2008). This mechanism biases the uptake of new spacers toward foreign genomes, thus avoiding auto-immunity (Levy et al., 2015; Modell et al., 2017). Moreover, the integration of a new spacer generates temporarily single-stranded regions within the CRISPR array. Sealing of those gaps by repair proteins like DNA polymerases and ligases are crucial to retain chromosome integrity (Ivančić-Baće et al., 2015).

There are also indications that Cas proteins may actively participate in DNA damage repair. Deletion of *cas1* in *Escherichia coli* yielded a phenotype that was sensitive to DNA damage and showed impaired chromosome segregation. The mechanism behind these observations likely relies on the ability of Cas1 to cleave branched DNA substrates that usually occur during DNA repair and recombination (Babu et al., 2011; Rollie et al., 2015). Moreover, Cas1 has been shown to associate with various repair proteins in *E. coli*, further strengthening the assumption that components of the *E. coli* CRISPR-Cas system are involved in DNA repair (Babu et al., 2011).

Another example illustrating the interplay of the CRISPR immune system and DNA repair was revealed in *Sulfolobus solfataricus*. Here, the CRISPR-associated protein Csa3a controls the expression of major CRISPR adaptation genes (Liu et al., 2015), is a key player in the activation of DNA repair genes (Liu et al., 2017), and regulates the DNA damage response (DDR) (Liu et al., 2020). The reason for the synergistic activation of both CRISPR-Cas and DNA repair pathways by Csa3a is probably due to the frequent acquisition of spacers from its own genome (roughly 7%) (Liu et al., 2017). The simultaneous activation of DDR genes thus reduces auto-immunity effects caused by self-targeting spacers (Liu et al., 2020).

The recognition of similar DNA substrates by Cas and repair nucleases might explain the mutually exclusive prevalence of type II-A CRISPR-Cas systems and NHEJ (non-homologous end joining) genes. Csn2, an essential adaptation protein of type II-A systems (Heler et al., 2015; Wei et al., 2015), likely inhibits NHEJ repair by binding DNA at double-stranded breaks (DSBs) (Arslan et al., 2013), outcompeting the repair machinery and forcing bacteria to select against either of those systems (Bernheim et al., 2017).

In this study, we describe the interplay between CRISPR-Cas and DNA repair in the archaeon *Haloferax volcanii* by showing that Cas1 and repair protein Fen1 display equivalent roles during DNA repair. Fen1 is a highly conserved component of the cellular DNA metabolism found in Archaea and Eukarya that targets and removes 5' flap structures commonly associated with Okazaki fragments or base excision repair (Pan et al., 2011; Balakrishnan and Bambara, 2013; Grasso and Tell, 2014). *Haloferax volcanii* has a single *fen1* gene located on the main chromosome (Lestini et al., 2010) while *cas1* is encoded on a megaplasmid within a type I-B CRISPR-Cas system (Maier et al., 2019). The *cas* gene cassette is flanked by two CRISPR loci, with a third CRISPR locus present on the main chromosome (Maier et al., 2015).

We found that deletions of either *cas1* or *fen1* caused similar phenotypes that exhibited sensitivity to oxidative stress and DNA damage. Double deletions caused cell elongation as well as increased cell size and when exposed to ultraviolet light (UV) stress survival rates were greatly reduced. We hypothesise that, given the similar substrate specificity and enzymatic activities of Cas1 and Fen1, both proteins are able to process 5' flap secondary DNA structures that occur during DNA repair and DNA replication.

RESULTS

Cas1 and Fen1 Are Required for DNA Repair

In order to investigate whether there is a functional link between components of the CRISPR-Cas system and the DNA repair machinery, we induced DNA damage in the *H. volcanii* wild-type strains and mutant daughter strains, harbouring a deletion of either *cas1, fen1* or both genes, respectively, and evaluated their fitness. To evaluate DNA repair, all strains were exposed to oxidative stress using H_2O_2 and survival rates were measured. Both single deletion strains displayed a significant reduction in survival compared to the wild-type, with $\Delta fen1$ showing a higher sensitivity than $\Delta cas1$ (**Figure 1A**). The $\Delta cas1 \Delta fen1$ double mutant displayed a growth defect comparable to the $\Delta fen1$ single mutant. These results indicate that both Cas1 and Fen1 are involved in the cellular oxidative stress response.

To exclude any polar effects of the gene deletions, we complemented the $\Delta cas1$ strain with a plasmid-born copy of the gene and were able to detect growth rescue in the complemented strain (**Figure 1B**). Moreover, when mutating the active site of Cas1 in the complementing gene (E158A, **Supplementary Figure 1**) (Babu et al., 2011; Kim et al., 2013; Rollie et al., 2015), growth rescue was impaired, further strengthening our
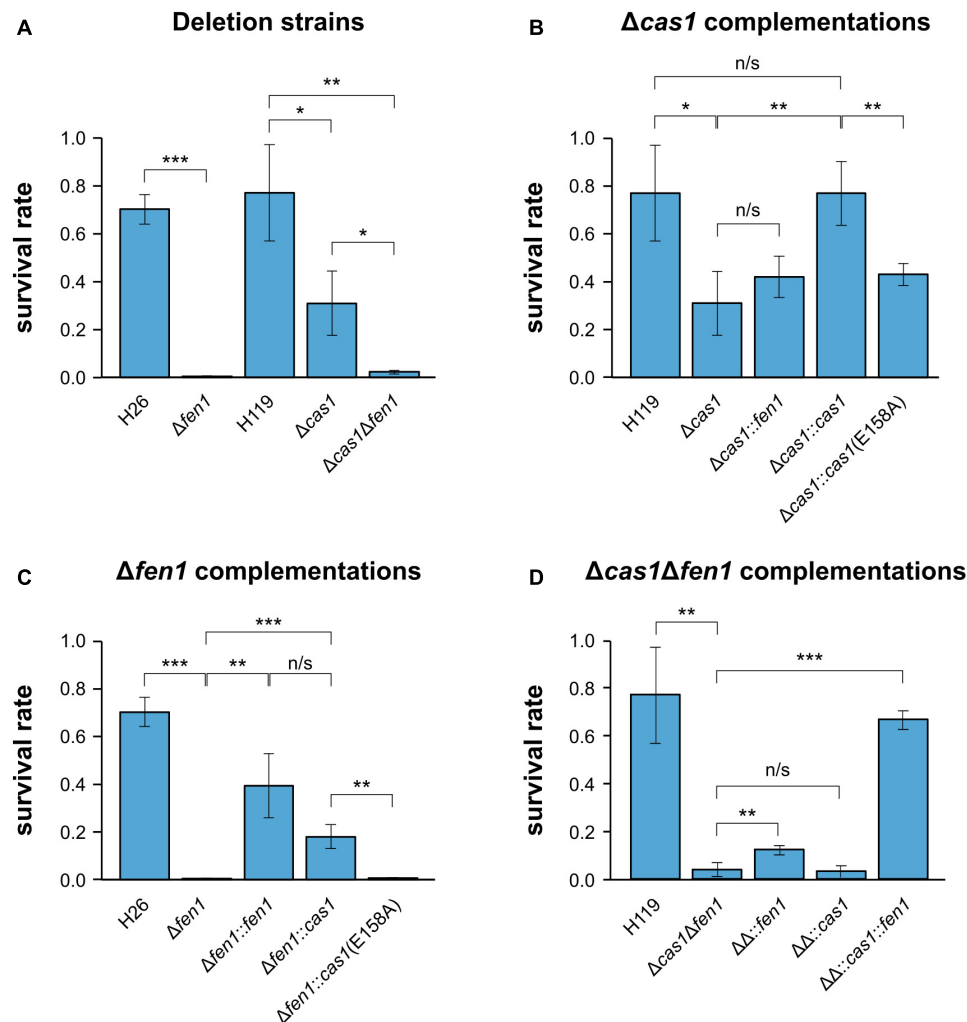
FIGURE 1 | Survival rates of *Haloferax volcanii* exposed to 4 mM H$_2$O$_2$. *Haloferax volcanii* strains were grown to the mid-log growth phase, exposed for 1 h to 4 mM H$_2$O$_2$ and survival rates were calculated by dividing the number of colony forming units (CFUs) determined for the exposed strains by the number of CFUs of an unexposed control. **(A)** Wild-type and single deletion strains as well as the double deletion strain were analysed. H26 is the wild-type strain for Δ*fen1*, H119 is the wild-type strain for Δ*cas1* and Δ*cas1*Δ*fen1* (**Supplementary Table 4**). **(B)** Δ*cas1* was complemented with *fen1*, *cas1* or a catalytically dead *cas1* mutant [*cas1(E158A)*]. **(C)** Δ*fen1* was complemented with *fen1*, *cas1* or the *cas1* mutant. **(D)** Δ*cas1*Δ*fen1* (Δ Δ) was complemented with *fen1*, *cas1* or the *cas1* mutant. Asterisks indicate significant differences (*t*-test) between the survival rates based on three independent experiments; ***: highly significant (p-value < 0.001), ** very significant (p-value < 0.01), * significant (p-value < 0.05), n/s = not significant.

assumption that the enzymatic activity of Cas1 is directly involved in DNA repair (**Figure 1B**).

We also supplemented the Δ*cas1* strain with a plasmid-borne copy of *fen1* but were not able to observe increased rescue, indicating that the DNA damage induced by H$_2$O$_2$ cannot be compensated by higher levels of Fen1 (**Figure 1B**). Previous reports showed that Cas1 has not only a 5′ flap processing activity but additional activities (Han et al., 2009; Wiedenheft et al., 2009; Beloglazova et al., 2015), which might explain why Fen1 cannot completely rescue a missing Cas1. Interestingly, when we overexpressed Cas1 in the Δ*fen1* background, survival rates increased compared to both the uncomplemented Δ*fen1* strain and the Δ*fen1* strain harbouring a plasmid expressed active-site mutant of *cas1* (**Figure 1C**), clearly demonstrating the capacity

of Cas1 to mitigate DNA damage. It is noteworthy that the Cas1 protein is only present in low levels in wild-type cells under standard conditions (Jevtić et al., 2019), which could explain why the overexpression of Cas1 showed a significant increase in cell survival, as opposed to Fen1 in the Δ*cas1* strain.

In the Δ*cas1*Δ*fen1* background, complementation with only *cas1* did not significantly increase the survival rate, while complementation with only *fen1* showed a slight but statistically significant growth rescue (p-value: 1,46E-02) (**Figure 1D**). These data are consistent with the results of the single deletion strains, where Δ*fen1* showed an elevated sensitivity toward oxidative stress (p-value: 3.47E-06) compared to Δ*cas1* (p-value: 1.68E-02).

To further evaluate the ability of Cas1 and Fen1 to repair DNA damage, we exposed each single deletion strain and the double
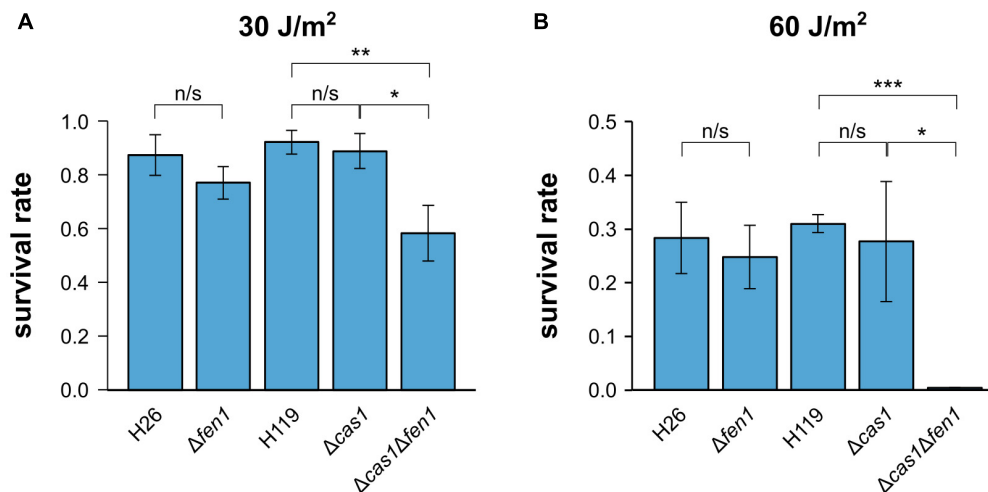
FIGURE 2 | Survival rates of *Haloferax volcanii* exposed to UV radiation. *H. volcanii* strains were grown to the mid-log growth phase, spotted on Hv-YPC plates and exposed to UV radiation [**(A)** 30 J/m², **(B)** 60 J/m²]. After growth, survival rates were calculated by dividing the number of CFUs determined for the exposed strains by the number of CFUs of an unexposed control. Asterisks indicate significant differences (*t*-test) between the survival rates based on three independent experiments; ***: highly significant (p-value < 0.001), **: very significant (p-value < 0.01), *: significant (p-value < 0.05), n/s = not significant.

mutant to UV radiation (**Figure 2**). We observed a significant decrease in cell survival at both 30 and 60 J m⁻², whereby only 1% of Δ*cas1*Δ*fen1* cells survived at 60 J m⁻² (**Figure 2B**). Interestingly, the single mutants did not display an increased sensitivity to UV radiation, indicating that UV-induced damage can be efficiently repaired in the presence of either enzyme alone.

Absence of Cas1 and Fen1 Cause Increased Cell Size and Cell Elongation

Our results indicate that both Cas1 and Fen1 play vital roles during the DNA damage response. To gain further insights into their roles in cellular growth, we compared morphological features of the wild-type and the Δ*cas1*Δ*fen1* deletion strain.

Analysis of the strains via light microscopy revealed that the double mutant displayed an inflated phenotype under normal growth conditions, while the single deletion strains (Δ*cas1* and Δ*fen1*, respectively) lacked any obvious change in phenotype (**Figure 3A**), indicating that the deletion of both genes has effects on cellular pathways other than DNA damage response. The altered morphology of the double mutant was confirmed by scanning electron microscopy (**Figure 3B**). Here, the cells showed a variety of elongated and inflated shapes, indicating defects in the growth cycle and division of the double mutant strain. This assumption was further strengthened by analysing cell size and DNA content using flow cytometry (**Figure 3C** and **Supplementary Figure 2**). Here, we observed a shift of the peak absorbance of the DNA content in the Δ*cas1*Δ*fen1* strain compared to the wild-type, indicating fewer genome copies per cell in the mutant. Considering the enlarged phenotype of the mutant, these results suggest that the mutant fails to properly replicate its genome, leading to impaired cell division, resulting in the cell inflating. It can therefore be hypothesised that Cas1 and Fen1 may play a role in DNA-related pathways, such as DNA

replication, to maintain genome integrity not only during stress responses but under normal growth conditions as well.

To assess whether the inflated phenotype of the Δ*cas1*Δ*fen1* mutant affects its ability to deal with osmotic stress, we monitored the growth of *H. volcanii* in media supplemented with different salt concentrations[1] via OD₆₅₀ₙₘ measurements (**Figure 4**). Under standard conditions (18% salt), we observed a similar growth rate of the wild-type and double deletion strain (**Figure 4A**).

Interestingly, in both low and high salt conditions (15% and 23% salt, respectively), the mutant displayed a shorter lag phase compared to the wild-type, indicating an increased adaptation rate to osmotic stress in the absence of Cas1 and Fen1 (**Figures 4B,C**). However, when grown at 23% salt, the optical density of the double mutant culture declined shortly after reaching the stationary phase at roughly 80 h post-inoculation, while the wild-type was still in the logarithmic growth phase at a similar timepoint.

Cas1 Interacts With DNA Repair Proteins

To exclude the possibility that the deletion of *cas1* and *fen1* has an effect on the expression level of other DNA repair proteins, thus indirectly causing impaired fitness during the cellular DNA damage response, we investigated gene expression patterns in both the wild-type and Δ*cas1*Δ*fen1* deletion strains using RNA sequencing. When comparing up- and down-regulated genes (**Tables 1**, **2** and **Supplementary Tables 1, 2**), only minor changes were observed in the gene expression levels of various functional groups that are not involved in DNA damage repair or DNA metabolism. However, the absence of Cas1 and Fen1 caused the

---

[1]Standard growth conditions (18% salt water) correspond to 15% NaCl, low salt (15% salt water) corresponds to 10.8% NaCl and high salt (23% salt water) corresponds to 19.2% NaCl.
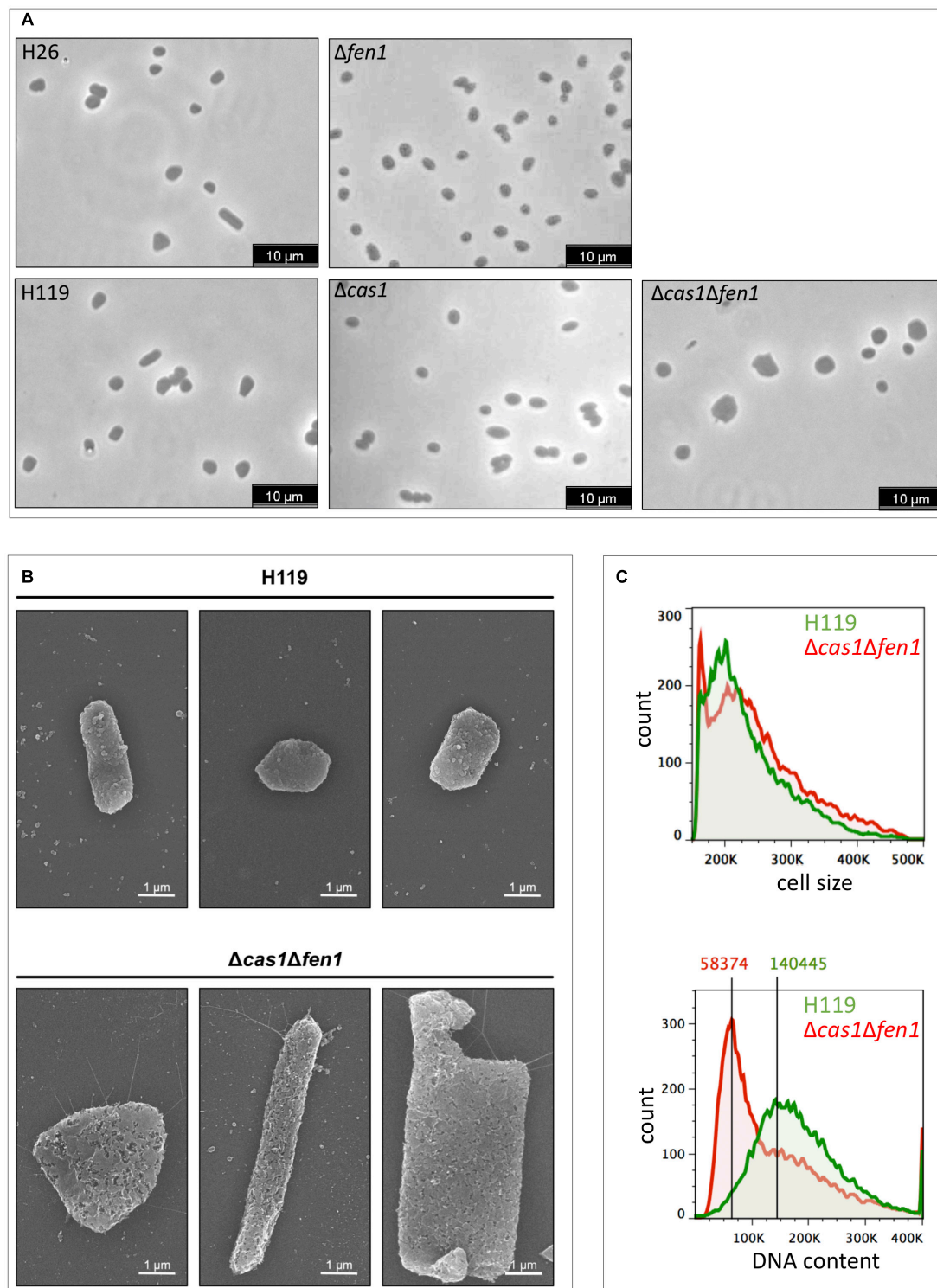
FIGURE 3 | Cell morphology and DNA content of wild-type and the Δ*cas1*Δ*fen1* deletion strain. **(A)** *Haloferax volcanii* wild-type and deletion strains were grown to mid-log growth phase and cell morphology was analysed by light microscopy. **(B)** Scanning electron microscopic pictures of wild-type strain H119 (upper panel) and Δ*cas1*Δ*fen1* (lower panel) grown to stationary phase. **(C)** Flow cytometry of wild-type H119 (green) and Δ*cas1*Δ*fen1* deletion strains (red) grown to mid-log growth phase. The FS (forward light scatter) channel was used to evaluate cell size (upper panel) and the FL1 (fluorescence 1) channel was used to evaluate the DNA content of acridine orange-stained cells (lower panel). Modal population values for DNA content highlights a decreased DNA content in the Δ*cas1*Δ*fen1* mutant (41.5% of wild-type value).

FIGURE 4 | Growth of deletion mutants in standard, low and high-salt conditions. Growth of *Haloferax volcanii* wild-type (orange) and deletion strains (green) under standard (18%) **(A)**, low (15%) **(B)** and high (23%) **(C)** salt conditions in Hv-YPC medium. Vertical lines represent standard deviation at different measuring points of three independent experiments. The x-axis shows the time of growth and the y-axis the OD$_{650nm}$.

up-regulation of altogether 15 putative proviral genes (Dyall-Smith et al., 2021) in the genome of *H. volcanii* (**Table 1** and **Supplementary Table 1**), which might be relevant for the immunity-related functions of Cas1.

Considering the minor changes in observed expression patterns, we conclude that the deletion of *cas1* and *fen1* does not strongly impact cellular pathways on a transcriptional level.

We further analysed direct protein-protein interactions of Cas1 by conducting pull-down analyses. Cas1 co-purified with

a variety of proteins involved in replication and DNA repair (**Table 3**), including several helicases and the DNA mismatch

TABLE 1 | RNA expression profile of the Δ*cas1*Δ*fen1* deletion mutant compared to the wild-type.

Gene_ID	Annotation/Gene name	logFC
Provirus regions		
HVO_0369	hypothetical protein; ProVir2 prediction	2.18
HVO_1434	hypothetical protein; ProVir5 prediction	2.16
HVO_A0218	oxidoreductase; ProVir4 prediction	2.01
HVO_0276A	homolog to HGPV1-ORF9; ProVir6 prediction	2.00
Iron metabolism		
HVO_1721	2Fe-2S iron-sulfur cluster binding domain-containing protein, *ferA3*	2.04
Transposases		
HVO_2817	transposase (ISH51)	2.06
HVO_A0258	ISH4-type transposase homolog	2.06
Other		
HVO_B0028	xylose dehydrogenase (NAD/NADP dependent), *xacA*	2.32
HVO_B0342	oxidoreductase (Luciferase family protein), *mer2*	2.19
HVO_0641	cob(II)yrinic acid a,c-diamide reductase, *bluB*	2.10
HVO_B0343	hydrolytic enzyme LplD, *lplD*	2.07
HVO_0694	purine phosphoribosyltransferase, *gptA*	2.05
HVO_1205	taxis cluster protein CheD, *cheD*	2.00

*Shown are up-regulated genes (logFC ≥ 2.00) (logFC: log₂ fold change.). Six genes for hypothetical proteins are also up-regulated with logFC ≥ 2.00 (**Supplementary Table 1**), for a complete list of up-regulated genes see **Supplementary Table 1**. Provirus prediction according to Dyall-Smith et al. (2021).*

TABLE 2 | RNA expression profile of the Δ*cas1*Δ*fen1* deletion mutant compared to the wild-type.

Gene_ID	Annotation/Gene name	logFC
Deleted genes		
HVO_2873	flap endonuclease, *fen1*	−8.13
HVO_A0211	Cas1 protein, *cas1*	−7.63
Transposases		
HVO_A0279	Transposase (ISH18)	−4.41
HVO_2051	Transposase (ISH51)	−2.23
Transcription regulators		
HVO_2507	Asn family transcription regulator, *trh7*	−3.17
HVO_2522	Asn family transcriptional regulator, *trh8*	−2.97
tRNA metabolism		
HVO_1092	ribonuclease P protein component 2, *rnp2*	−2.68
HVO_3052	tRNA Gly	−2.35
Iron metabolism		
HVO_A0541	ABC-type transport system periplasmic substrate-binding protein (probable substrate iron-III)	−2.45
HVO_B0044	iucA iron transport protein A, *iucA*	−2.43
HVO_2588	isocitrate dehydrogenase, *iucD*	−2.28
Other		
HVO_1228	halocyanin domain protein (membrane), *hcpE*	−2.92
HVO_2508	carbamoyl-phosphate synthase small subunit, *carA*	−2.56
HVO_2361	carbamoyl-phosphate synthase large subunit, *carB*	−2.04
HVO_B0045	daminobutyrate decarboxylase, *bdb*	−2.47
HVO_B0046	diaminobutyrate pyruvate aminotransferase, *dat*	−2.26
HVO_2606	PQQ repeat-containing protein	−2.72
HVO_2607	PQQ repeat-containing protein	−2.29

*Shown are down-regulated genes (logFC ≤ 2.00) (logFC:log2 fold change.). Nine genes for hypothetical proteins are also down-regulated with logFC ≤ 2.00 (**Supplementary Table 2**), for a complete table of down-regulated genes see **Supplementary Table 2**.*

Gene ID	Annotation/Gene name	Peptides/Unique spectra counts
Replication and repair		
HVO_0393	UvrABC system protein A, *uvrA*	73
HVO_0029	UvrABC system protein B, *uvrB*	32
HVO_0415	repair helicase UvrD, *uvrD*	38
HVO_0349	DNA-directed RNA polymerase subunit A, *rpoA1*	56
HVO_0347	DNA-directed RNA polymerase subunit B, *rpoB2*	44
HVO_0858	DNA-directed DNA polymerase B (intein-containing), *polB1*	36
HVO_2380	AAA-type ATPase (CDC48 subfamily), *cdc48a*	66
HVO_0854	DNA double-strand break repair ATPase Rad50, *rad50*	56
HVO_B0118	Smc-like protein Sph2; homolog of Rad50, *sph2*	30
HVO_0552	DNA mismatch repair protein MutS, *muts1b*	49
HVO_0014	ATP-dependent DNA helicase Hel308a, *hel308a*	44
HVO_1018	Hef-associated 3 exonuclease, *recJ3*	35
HVO_2889	DHH/RecJ family phosphoesterase RecJ4, *recJ4*	33
HVO_0220	ATP-dependent DNA helicase MCM, *mcm*	34
Cas protein		
HVO_A0206	Cas8b, *cas8*	58
Ribonucleases		
HVO_0874	zinc-dependent nuclease CPSF1, *cpsf1*	44
HVO_2724	ribonuclease J, *rnJ*	35
Sensing kinases		
HVO_1811	sensor box histidine kinase	32
HVO_B0154	receiver/sensor box histidine kinase	41

*Two independent pull-down analyses using a FLAG-Cas1 fusion protein resulted in the co-purification of interacting proteins. Corresponding peptides of the co-purifying proteins were sequenced by mass spectrometry (MS). Only proteins that are functionally interesting, e.g., involved in DNA repair, CRISPR-Cas functions, and having unique spectra/peptide counts ≥ 30 are listed. The complete table of co-purified proteins with a spectra/peptide count ≥ 30 is found in **Supplementary Table 3**. The complete set of MS data including all co-purified proteins (including those with a spectra/peptide count < 30) has been deposited in PRIDE. Column peptides/unique spectra counts: number of peptides and unique spectra counts identified.*

repair protein MutS, and was associated with three members of the UvrABC system that acts in nucleotide excision repair. For a final confirmation of these interactions reverse pull-down experiments have to be carried out where the proteins identified by the Cas1 pull-down are FLAG-tagged and the presence of Cas1 in the elution fraction is determined after FLAG-purification.

Other proteins that co-purified with Cas1 included RNases, sensing kinases, and proteins associated with amino acid synthesis, vitamin and glucose metabolism (**Table 3** and **Supplementary Table 3**). Interestingly, only five peptides of Fen1 were detected during the MS analysis (data not shown, the complete set of co-purified proteins has been deposited in

PRIDE), suggesting that the interaction between Cas1 and Fen1 is only weak or does not occur at all. This indicates that their functional link might not be based on physical interactions within the cell.
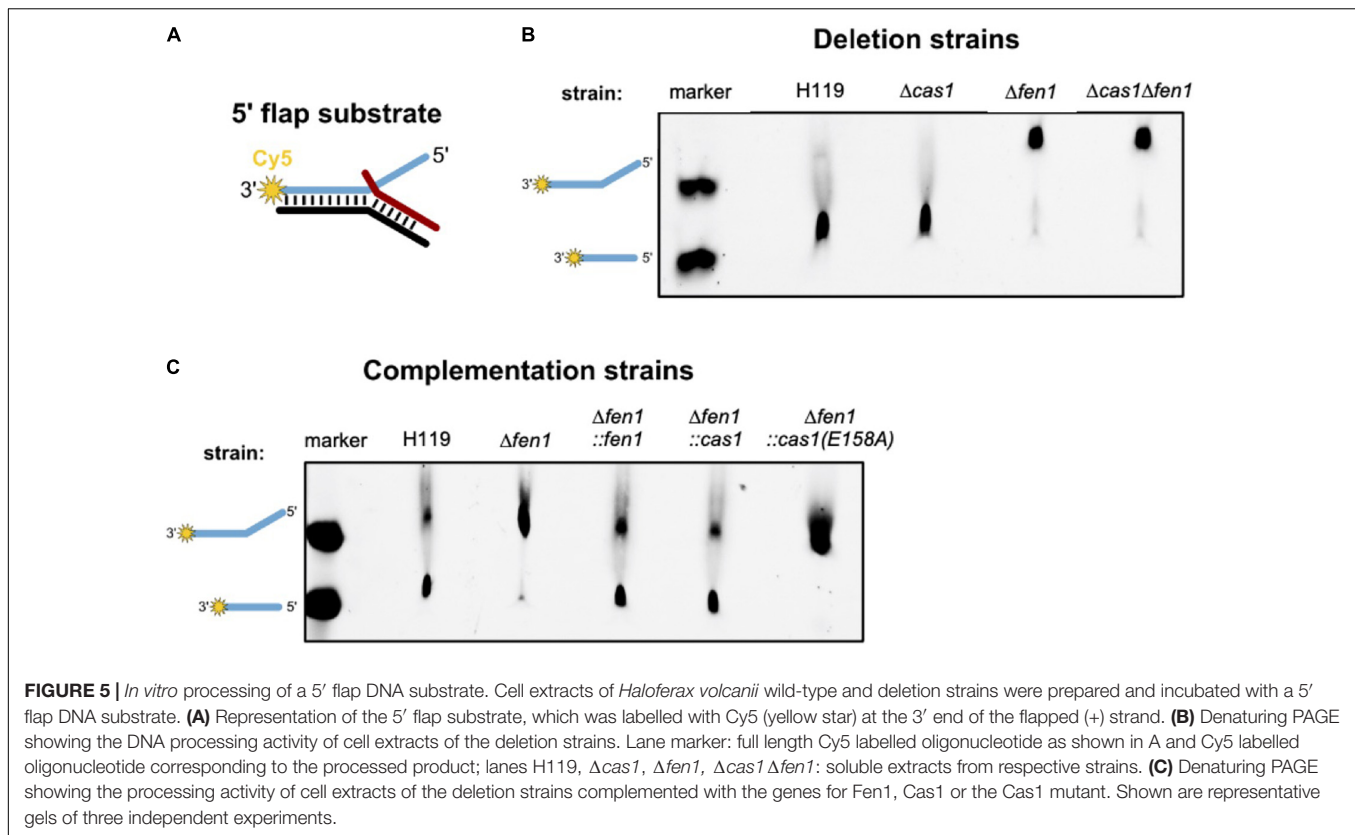
Cas1 Has 5′ Flap Endonuclease Activity

Previous studies have revealed that Cas1 has DNA processing activities and is able to cleave branched DNA substrates such as Holliday junctions and 5′ flap structures (Babu et al., 2011; Rollie et al., 2015). We hypothesised that the type I-B Cas1 homolog of *H. volcanii* might possess similar activities on branched DNA substrates. In the context of DNA repair, processing of such structures by Cas1 might compensate for a deletion of the flap endonuclease Fen1. To test this, we attempted to express and purify both nucleases recombinantly in *E. coli* but failed to achieve adequate yield and purity. Therefore, we prepared cell extracts of *H. volcanii* wild-type as well as *cas1* and *fen1* deletion mutants and compared their ability to process a double flap DNA substrate (**Figure 5**).

We observed processing of the branched substrate by cell extracts of the wild-type and Δ*cas1* strain, whereas cell extracts of the Δ*fen1* strain yielded no DNA cleavage product, suggesting that only Fen1 is able to process the DNA double flap substrate (**Figure 5B**). It is noteworthy that Cas1 levels are considerably low under normal growth conditions (Jevtić et al., 2019), potentially too low to observe substrate cleavage in our assay. Therefore, we repeated the cleavage assay using cell extracts of a Δ*fen1* strain that constitutively expressed an *in trans* copy of *cas1* and were able to detect processing of the double flap substrate (**Figure 5C**). Moreover, this effect was not observable in the Δ*fen1* strain expressing a nuclease-defective variant of Cas1 (E158A), which strongly indicates that both Fen1 and Cas1 are able to recognise and process a DNA flap.

DISCUSSION

Thus far it has only been demonstrated in bacteria that components of the prokaryotic immune system CRISPR-Cas may be involved in other cellular pathways, such as DNA repair. In this study, we demonstrate a functional link between CRISPR-Cas systems and DNA metabolism in Archaea, by showing that Cas1 and the replication enzyme Fen1 play equivalent roles during the DNA damage response in *H. volcanii*.

Fen1 is found in Archaea and Eukarya and cleaves flapped DNA structures that occur during DNA replication and repair (Pan et al., 2011; Balakrishnan and Bambara, 2013). Accordingly, deletion of the homologous gene in *H. volcanii* causes reduced survival rates following DNA-damaging oxidative stress. Deletion of *cas1* had similarly adverse effects on cell survival under these conditions, indicating that the CRISPR-Cas immune protein is involved in DNA repair as well. Additionally, we found that the enzymatic activity of the Cas1 nuclease is directly responsible for DNA repair, as a mutation in the active site of the protein also resulted in reduced survival after oxidative stress. The mechanistic basis for repairing DNA damage is likely linked to

FIGURE 5 | *In vitro* processing of a 5′ flap DNA substrate. Cell extracts of *Haloferax volcanii* wild-type and deletion strains were prepared and incubated with a 5′ flap DNA substrate. **(A)** Representation of the 5′ flap substrate, which was labelled with Cy5 (yellow star) at the 3′ end of the flapped (+) strand. **(B)** Denaturing PAGE showing the DNA processing activity of cell extracts of the deletion strains. Lane marker: full length Cy5 labelled oligonucleotide as shown in A and Cy5 labelled oligonucleotide corresponding to the processed product; lanes H119, Δ*cas1*, Δ*fen1*, Δ*cas1*Δ*fen1*: soluble extracts from respective strains. **(C)** Denaturing PAGE showing the processing activity of cell extracts of the deletion strains complemented with the genes for Fen1, Cas1 or the Cas1 mutant. Shown are representative gels of three independent experiments.

the ability of the nuclease Cas1 to cleave flapped DNA substrates in a similar manner to Fen1.

Pull-down experiments showed that Cas1 seems to physically interact with a variety of DNA binding and processing enzymes, some of which are involved in DNA repair. Whether these interactions affect the role of Cas1 in DNA repair remains to be revealed, however, it is tempting to speculate that Cas1 is part of a multi-protein repair complex in *H. volcanii*. The formation of such a complex might determine the role of Cas1 in either DNA repair or CRISPR-Cas immunity, which would be advantageous for the cell. During viral defence, Cas1 forms a complex with Cas2 and integrates short pieces of DNA into the CRISPR array (Nuñez et al., 2014). Recruiting the Cas1-Cas2 complex to DNA lesions during repair could trigger spacer acquisition from the organism's own genome and consequently lead to auto-immunity. The interaction of Cas1 with DNA repair enzymes would prevent this problem, since complex formation with Cas2 is constrained and thus prevents spacer acquisition. Further experiments are required to reveal the factors that influence the role of Cas1 in either damage repair or immunity. In our pull-down experiments Cas2 could not be identified as interaction partner of Cas1. Cas2 is not detectable in proteome analyses of cells grown under standard conditions (Jevtić et al., 2019; Schulze et al., 2020) which might explain this observation.

In addition to its role in DNA repair, Cas1 might also play a role in DNA replication. In *E. coli*, it has been proposed that Cas1 is involved in genome segregation, based on the detection of abnormally elongated cells of a *cas1* knockout

strain after mitomycin C treatment (Babu et al., 2011). In *H. volcanii*, a single deletion of either *fen1* or *cas1* did not cause any notable morphological changes in cell shape or size, compared to the wild-type under unstressed conditions. However, a simultaneous deletion of both genes led to an inflated phenotype and reduced DNA content per cell, with the latter being indicative of fewer genome copies per cell. Similar morphological alterations have been observed before, where UV-induced DNA damage in *H. volcanii* results in increased cell size and cell elongation (Delmas et al., 2013). These results suggest that both Cas1 and Fen1 might process flapped DNA substrates during replication and thus maintain genome integrity under conditions of genomic stress.

CONCLUSION

This study expands our knowledge of the role of CRISPR-Cas systems beyond immunity, by demonstrating that, similar to its bacterial homologs, the archaeal Cas1 in *H. volcanii* may participate in cellular DNA repair.

MATERIALS AND METHODS

Strains, plasmids and oligonucleotides used are listed in **Supplementary Tables 4–6**. *E. coli* strain DH5a (Invitrogen, Thermo Fischer Scientific, Waltham, MA, United States) was used for plasmid cloning and grown aerobically at 37°C in 2YT medium (Miller, 1972).

Growth Curves

Haloferax volcanii strains were grown in Hv-YPC medium (Dyall-Smith, 2008) at 45°C with an optimal concentration of 18% salt water (2.46 M NaCl, 88 mM MgCl$_2$, 85 mM MgSO$_4$, 56 mM KCl, 12 mM Tris-HCl, pH 7.5). For low and high salt stress, salt concentration was adjusted to 15% and 23%, respectively, without modifying the ion ratios. While standard conditions (18% salt water) correspond to 15% NaCl, low salt (15% salt water) corresponds to 10.8% NaCl and high salt (23% salt water) corresponds to 19.2% NaCl. Growth of cultures was monitored in 96 well microtiter plates using an Epoch2 NS Microplate Spectrophotometer (BioTek Instruments, Bad Friedrichshall, Germany). Strains were precultured in Hv-YPC medium to OD$_{650nm}$ of 0.4–0.7, after dilution to an OD$_{650nm}$ of 0.05 they were transferred to microtiter plates. Cultures were incubated aerobically with orbital shaking at 45°C and OD$_{650nm}$ was measured every 30 min. Outer wells were filled with salt water as evaporation barriers (Liao et al., 2016). The growth curves represent the average of at least three biological replicates.

Plasmid Cloning

pBlue-*Nde*I-Cas1-*Eco*RV was generated by ligating a linearised pBluescriptII KS (cleaved with *Eco*RV) with the insert that contained *cas1* and the restriction sites *Nde*I and *Eco*RV. The insert was amplified from *H. volcanii* genomic DNA using primers Cas1up-FLAGC and Cas1do-FLAGC, primers contained the restriction sites *Nde*I and *Eco*RV.

pTA927-Cas1E158A-CFLAG was generated using a site-directed mutagenesis kit (Agilent, Santa Clara, CA, United States) according to the manufacturer's instructions with primers Cas1E158A fw and Cas1E158A rev and pTA927-Cas1-CFLAG as a template.

Plasmid pTA927-FLAGcontrol was cloned as follows. Using primers FLAG *Nde*I fw and FLAG *Eco*RI rev with pTA927-NZ-NFLAG as template, the FLAG sequence (ATG-3xFLAG-TGA) was amplified. The primers contained restriction sites for *Apa*I (5′) and *Eco*RI (3′), after digestion of the resulting polymerase chain reaction (PCR) product with *Apa*I and *Eco*RI ligation with pTA927 (digested with *Apa*I and *Eco*RI) resulted in pTA927-FLAGcontrol.

Plasmid pTA231.p.fdx was obtained by ligation of pTA231 (digested with *Not*I and *Eco*RI) with a p.fdx-t.syn fragment excised with *Not*I and *Eco*RI from pMA-T-FdxSyn (synthesised by GeneArt, Thermofisher Scientific), resulting in pTA231-p.fdx.

In order to generate pTA231-p.fdx-NFLAG, a PCR with primers 231NFLAG fw and 231NFLAG rev and pTA927-Cas1-NFLAG as a template was performed resulting in PCR fragment ATG-3xFLAG, primers contained a *Nde*I restriction site (5′ terminus) and a *Sna*BI as well as a *Xba*I restriction site (3′ terminus). The purified PCR product was digested with *Nde*I and *Sna*BI and ligated to pTA231-p.fdx (digested with *Nde*I and *Xba*I), to generate pTA231-p.fdx-NFLAG.

Plasmid pTA230-p.fdx-NFLAG was generated by digesting pTA231-p.fdx-NFLAG with *Not*I and *Eco*RI to obtain the p.fdx-NFLAG-t.syn fragment (p.fdx-ATG-3x FLAG-t.syn), which was subsequently ligated with the pTA230 vector (linearised with *Not*I and *Eco*RI).

Plasmids pTA230-p.fdx-Cas1-NFLAG and pTA230-p.fdx-Fen 1-NFLAG were obtained as follows. *fen1* and *cas1* genes were amplified from *H. volcanii* genomic DNA using primers Fen1 *Sna*BI fw/Fen1 *Xba*I rev and Cas1 *Eco*RV fw/Cas1 *Xba*I rev, respectively. The resulting PCR products were ligated with pBluescriptII KS (*Eco*RV digested), resulting in pBlue-*Sna*BI-Fen1-*Xba*I and pBlue-*Eco*RV-Cas1-*Xba*I. These plasmids were then digested with *Sna*BI/*Xba*I or *Eco*RV/*Xba*I to excise the *fen1* and *cas1* gene fragments, respectively, which were subsequently ligated with pTA230-p.fdx-NFLAG (linearised with *Sna*BI and *Xba*I) yielding pTA230-p.fdx-Fen1-NFLAG and pTA230-p.fdx-Cas1-NFLAG.

For generation of pTA230-p.fdx-Cas1E158A-NFLAG, a site-directed mutagenesis kit (Agilent, Santa Clara, CA, United States) was used according to the manufacturer's instructions with primers Cas1E158A fw and Cas1E158A rev and pTA230-p.fdx-Cas1-NFLAG as a template.

To obtain the pTA962-p.fdx-Fen1-NFLAG-Cas1 overexpression plasmid, the *cas1* insert was generated with PCR using primers 5′XmaJI Cas1 + Start fw/3′BglII Cas1 + Stopp rev and plasmid pTA230-p.fdx-Cas1-NFLAG as template. The resulting PCR product was cleaved with *Xma*JI/*Bgl*II and then ligated with pTA962 (cleaved with *Xma*JI/*Bgl*II). To generate p.fdx-Fen1-NFLAG, the *fen1* insert was obtain with PCR using primers 5′ApaI p.fdx fw/3′XmaJI Fen1 + Stopp rev and pTA230-p.fdx-Fen1-NFLAG as template. The resulting PCR product was cloned into pTA962 via *Apa*I/*Xma*JI restriction sites.

The *cas1* deletion plasmid pTA131-UPDO(*cas1*) was obtained by amplifying the *cas1* gene with 480 bp up- and 587 bp down-stream sequences using primers Cas1KOUP/Cas1KODO. The resulting fragment was ligated with pTA131 (linearised with *Eco*RV). The resulting plasmid pTA131-UP-Cas1-DO was used as template for an inverted PCR with 5′ phosphorylated primers iCas1 fw/iCas1 rev, the resulting PCR product was ligated to obtain pTA131-UPDO(*cas1*). This plasmid contains only the up- and downstream regions of *cas1*.

Generation of Knockout Mutants

Deletion strains were generated using the pop-in pop-out method as described previously (Bitan-Banin et al., 2003). Transformations were performed according to the PEG600 protocol as described in the HaloHandbook (Dyall-Smith, 2008). Strain Δ*cas1* was generated as described (Klein, 2011), in short: For the generation of the Δ*cas1* knockout strain, the H119 wild-type strain was transformed with the *cas1* deletion plasmid pTA131-UPDO(*cas1*) and plated onto Hv-Ca medium (Dyall-Smith et al., 2021) supplemented with L-tryptophan (40 µg/ml) to select for pop-in candidates which were subsequently identified by PCR using the primers Cas1 KO UP/Cas1 KO DO. Positive clones were then grown in Hv-Ca medium supplemented with L-tryptophan (40 µg/ml) and subsequently plated onto Hv-Ca agar plates with L-tryptophan (40 µg/ml), 5-FOA (5-fluoro-orotic acid) (10 µg/ml) and 10 µg/ml uracil to select for pop-out clones. Knockout candidates were identified by PCR using the same primers as mentioned above and the additional primer pair Cas1 *Eco*RV fw/Cas1 *Apa*I rev. Homozygous deletion strains were finally confirmed via Southern Blot analysis (Sambrook and Russell, 2001; Klein, 2011).

In order to establish a $\Delta cas1 \Delta fen1$ knockout strain, the $\Delta cas1$ strain was transformed with *fen1* deletion plasmid pCN6 (Meslet-Cladiére et al., 2007). Transformants were selected by plating on Hv-Ca medium supplemented with tryptophan (40 µg/ml). Pop-in candidates were subsequently plated onto Hv-Ca agar plates supplemented with L-tryptophan (40 µg/ml), 5-FOA (10 µg/ml) and 10 µg/ml uracil. Pop-in and pop-out clones were identified by PCR using the primers Fen1 HVO UP/Fen1 HVO DO and Fen1 *Sna*BI fw/Fen1 *Xba*I rev. Finally, homozygous pop-out candidates were verified by Southern Blot analysis (Sambrook and Russell, 2001) (**Supplementary Figure 3**). Genomic DNAs of candidate clones K1, K4, K17, K20, K22 and of wild-type H119 were extracted via spooling method (Dyall-Smith, 2008). Concentrations of genomic DNAs were determined with a NanoPhotometer N60, the DNA was stored at 4°C. Genomic DNA was digested with *Sal*I and a Southern blot analysis was made as described in Sambrook and Russell (2001). Radioactively labeled PCR products were used as hybridisation probes to detect DNA fragments on the southern blot membrane. PCR products amplified with Fen1 Sonde fw/Fen1 Sonde rev and Fen1 (HVO) UP/Fen1 (HVO) DO, respectively, using H119 gDNA as a template were used as probes. For radiolabeling, 50 µCi [α-^{32}P]-dCTP and the DECAprime^{TM} II DNA labeling kit (Life Technologies/Merck, Germany) were used. Two separate probes were generated, one that binds upstream of the *fen1* coding sequence, and another one that binds directly in the *fen1* gene. After hybrididsation the membrane was exposed to an x-ray film (**Supplementary Figure 3**).

H$_2$O$_2$ and UV Radiation Survival Assays

Wild-type and deletion strains were inoculated in 4 ml Hv-YPC medium, strains that carried a plasmid were inoculated in 4 ml selective medium with the appropriate supplements and incubated over night at 45°C until culture density reached an OD$_{650nm}$ of 0.3–0.5. To determine survival rates after H$_2$O$_2$ exposure, cultures that reached the adequate OD$_{650nm}$ were aliquoted into two 2 ml reaction tubes by pipetting 490 µl culture into each tube. One tube was supplemented with 10 µl of a 200 mM H$_2$O$_2$ stock solution (final concentration 4 mM), and 10 µl of 18% saltwater was added to the other tube (control). H$_2$O$_2$ stock solutions were freshly prepared for each experiment by diluting 30% H$_2$O$_2$ with 18% saltwater. Samples were incubated at 45°C and 450 rpm for 1h. Subsequently, samples were serially diluted (1:10) in 18% saltwater and 20 µl of the 10^{-3} to 10^{-6} dilutions were spotted onto prewarmed Hv-YPC agar plates in duplicate. After the liquid was absorbed by the agar, plates were incubated at 45°C for 3 days until colonies could be counted. Survival rates were determined by dividing the number of colonies of the 4 mM H$_2$O$_2$ approach by the number of colonies of the control. P-values were determined by *t*-tests (two sample assuming equal variances).

To compare the sensitivity of the wild-type and deletions strains toward UV radiation, strains were incubated as described above until an OD$_{650nm}$ of 0.3–0.5. Next, cultures were serially diluted (1:10) in 18% saltwater and 20 µl of the dilutions 10^{-2} to 10^{-6} were spotted onto Hv-YPC agar plates in duplicate. After the liquid was absorbed by the agar, plates were exposed to UV radiation (either 30 or 60 J m^{-2}) in a Stratalinker® UV Crosslinker (Stratagene, Agilent, Santa Clara, CA, United States). Control plates were not exposed to UV radiation. To prevent photoreactivation, all plates were incubated in the dark at 45°C for at least three days. Colonies were counted and survival rates were calculated by dividing the number of colonies on irradiated plates by the number of colonies on the control plates. All experiments were performed in triplicate.

Light Microscopy

Liquid cultures were inoculated in 4 ml Hv-YPC medium and incubated at 45°C overnight. After cultures reached an OD$_{650nm}$ of 0.3–0.5 (exponential phase) and 0.9–1.2 (stationary phase), respectively, 1 ml of each culture was transferred into a reaction tube and centrifuged at 1,500 x *g* for 10 min. The supernatant was discarded, and the obtained pellets were resuspended in 50 µl (derived from exponential culture) and 150 µl (derived from stationary culture) Hv-YPC medium, respectively, and applied onto agarose-covered microscope slides. Agarose slides were prepared as follows: 0.1 g agarose was solved in 4 ml bidestilled water by heating the suspension in a microwave until the agarose was completely dissolved. Meanwhile, 6 ml 18% saltwater was prewarmed in a water bath and added to the hot agarose solution. With the help of a brush, the agarose solution was applied onto the slides. 2 µl of the resuspended pellets were pipetted onto the hardened agarose slides and covered with a coverslip. The edges of the cover glasses were sealed with clear nail polish and microscopic analysis was done at 100 x magnification with oil immersion on a Leica DM5500 B light microscope.

Scanning Electron Microscopy

For scanning electron microscopy H119 and $\Delta cas1 \Delta fen1$ cells were grown to stationary phase (OD$_{650nm}$: 1.2) in Hv-YPC, adsorbed on silicon platelets and fixed for 1 h with a final concentration of 2.5% glutaraldehyde in 0.1 M phosphate buffer with 1% saccharose. Afterward, samples were fixed with 2% of osmium tetroxide in phosphate buffer for 20 min, dehydrated in a graded series of propanol and then critical point dried using carbon dioxide in a CPD BalTec030 critical point dryer. Finally, samples were mounted on specimen stubs and rotary coated by electron beam evaporation with about 2 nm of platinum. Samples were imaged in a Hitachi S-5200 scanning electron microscope at an accelerating voltage of 10 kV using the secondary electron signal.

Co-immunoprecipitation and Mass Spectrometry

The FLAG-Cas1 fusion protein and as a control the FLAG protein were expressed in *Haloferax* cells (strain $\Delta cas1$). 500 ml Hv-Ca medium supplemented with 40 µg/ml L-tryptophan were inoculated from a plate and grown until OD$_{650nm}$ of 0.7. After incubation for 1 h cultures were harvested by centrifugation at 10,000 x *g* for 30 min at 4°C. The resulting pellets were washed with enriched PBS [2.5 M NaCl, 150 mM MgCl$_2$,1xPBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$PO$_4$, 2 mM KHPO$_4$, pH 7.4)] and stored at −80°C. Pellets were thawed on ice

and resuspended in 10 ml lysis buffer (50 mM Tris-HCl, pH 7.4, 1 mM EDTA, 10 mM $MgCl_2$, 1 mM $CaCl_2$) containing 150 µl of proteinase inhibitor (Sigma-Aldrich/Merck, Darmstadt, Germany) and 100 µl RQ1 RNase-free DNase (Promega, Walldorf, Germany). After 30 min incubation on ice, cells were disrupted by ultrasonic treatment and subsequently centrifuged at 100,000 x g at 4°C for 30 min. The supernatants (S100) were transferred into fresh falcons and RNase A was added at a final concentration of 400 µg/ml. After incubation for 30 min at 37°C, NaCl was added to a final concentration of 150 mM. A total of one FLAG control purification and two FLAG-Cas1 purifications were carried out.

For affinity purification, 400 µl of anti-FLAG M2 affinity gel (Sigma-Aldrich/Merck, Darmstadt, Germany) was washed 10 times with 10 ml of ice-cold washing buffer (50 mM Tris/HCl, pH 7.4, 150 mM NaCl) before the lysate was added. After incubation overnight at 4°C, anti-FLAG M2 affinity gel was washed 10 times with 10 ml of ice-cold washing buffer. Proteins were eluted from the affinity gel using washing buffer supplemented with 150 ng/l 3xFLAG peptide and fractions were collected in 2 ml reaction tubes and finally precipitated using acetone (99%). Protein concentrations were determined using ROTI®Nanoquant according to the manufacturer's protocol, extracts were snap frozen in liquid nitrogen and stored at −80°C. The FLAG elution fraction was concentrated using an Amicon Ultra – 4 10,000 NMWL device (centrifugation for 15 min at 7,500 xg and 4°C). 50 mM Na_3PO_4/150 mM NaCl buffer was added to the retentate to a final volume of 200 µl. The sample was then loaded onto a Superdex™ 75 Increase 10/300 gel filtration column. 1 ml fractions containing Cas1 and eluting with the complex size of approximately 670 kDa were collected and used for mass spectrometry analyses.

SDS-PAGE-separated protein samples were processed as described by Shevchenko et al. (1996). The resulting peptides were loaded to nano HPLC coupled with Exploris Orbitrap Mass spectrometer (Thermo Fisher Scientific, Waltham United States). The peptides were separated with a linear gradient of 5–40% buffer B (80% acetonitrile and 0.1% formic acid) at flow rate of 300 nL/min over 48 min total gradient time. The MS data was aquired by scanning the precursors in mass range from 350 to 1400 m/z at a resolution of 60,000 at m/z 200. Top30 precursor ion were chosen for MS2 by using data-dependent acquisition (DDA) mode at a resolution of 15,000 at m/z 200 with auto IT. Data analysis and search was performed against Uniprot_Haloferax_volcanii database (July 2021, 3911 entries) using Maxquant Software (1.6.17.0) and results were annotated with Scaffold5 software. Search parameters were 0.1 FDR on peptide and protein level; Carbamidomethyl was set as fixed modification and oxidation of methionine and acetylation at the N-terminus of the protein as variable modifications.

To evaluate the data, first all proteins for which more than five peptides were found in the FLAG control were removed from the final tables (**Table 3** and **Supplementary Table 3**). In addition, for proteins that were found in the FLAG only control (identified with four or less peptides) the number of peptides found in the FLAG only control was subtracted from those found in the FLAG-Cas1 samples. In addition, only proteins found in both FLAG-Cas1 purifications were included in **Table 3** and **Supplementary Table 3**.

Flow Cytometry

DNA content and cell size of *H. volcanii* cells were determined via flow cytometry. Cultures were prepared in 5 ml Hv-YPC broth and grown at 45°C with 8 rpm rotation in two successive overnight dilutions until an OD_{650nm} of ∼0.4 was reached. Acridine orange solution was added to the cells at a final concentration of 1 µg/ml. Samples were analysed using an FC500 flow cytometer (Beckman Coulter; University of Nottingham Flow Cytometry facility) equipped with a 488 nm laser and 528/28 emission filter to measure acridine orange fluorescence. Samples were run on the lowest speed setting and at least 20,000 cells were acquired for each sample. Data was analysed using Flow Jo v7.6 (Tree Star Inc.). Cells were gated based on forward and side scatter and doublets were excluded by height/area analysis.

Preparation of Cell Extracts and Nuclease Assays

For the preparation of total protein extracts, 150 ml cultures were grown in Hv-YPC medium (strains without plasmids) or Hv-Min medium (strains carrying a plasmid) until they reached an OD_{650nm} of 0.6–0.8. Cells were harvested by centrifugation at 10.000 x g and 4°C for 30 min, cell pellets were washed once with enriched PBS [2.5 M NaCl, 150 mM $MgCl_2$, 1xPBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2PO_4, 2 mM $KHPO_4$, pH 7.4)] and stored at −80°C until further usage. Pellets were thawed on ice and resuspended in 2 ml Lysis buffer (100 mM Tris-HCl pH 7.5, 10 mM EDTA) and subsequently disrupted by ultrasonic treatment. To separate soluble and non-soluble lysate fractions, suspensions were centrifuged at 100,000 x g and 4°C for 30 min. The supernatants (S100) were transferred into 2 ml reaction tubes and total protein concentration was measured using ROTI® Nanoquant solution according to the manufacturer's instructions. Extracts were snap frozen in liquid nitrogen and stored at −80°C until further use.

For the *in vitro* 5′ flap processing experiments, 20 µg of the S100 protein extracts were mixed with 150 ng double flap DNA substrate and 10x IVP buffer (500 Tris-HCl pH 7.4, 500 mM $MgCl_2$, 50 mM KCl) and incubated for 2 h at 45°C. The longer flap strand was ordered as oligonucleotide labelled with Cy5 at the 3′ end (Sigma-Aldrich, Taufkirchen, Germany). (**Figure 5**, as described in Craggs et al., 2014). Proteins were inactivated and DNA was denatured by heat treatment (90°C, 10 min) and reactions were loaded onto a denaturing 10% PAA gel. Cy5 signals were detected on a ChemiDoc using standard parameters for Cy5.

RNA Sequencing

Three replicates of wild-type H119 and deletion strain Δ*cas1*Δ*fen1* were cultured in Hv-YPC medium and grown to OD_{650nm} of 0.6–0.7. Total RNA was isolated using NucleoZOL™ (Machery and Nagel, Düren, Germany) and RNA samples were sent to vertis Biotechnologie AG (Martinsried, Germany) for sequencing. Total RNA was treated with T4 Polynucleotide Kinase (NEB, Frankfurt,

Germany) and rRNA was depleted using an in-house protocol. Samples were fragmented using an ultrasonic treatment before cDNA library was prepared. For this, first-strand cDNA synthesis was performed using a 3′-adapter primer and the M-MLV reverse transcriptase. After cDNA purification, the 5′ Illumina TruSeq sequencing adapter was ligated to the 3′ end of the antisense cDNA and the sample amplified using TruSeq_Sense_primer and TruSeq_Antisense_primer to 10–20 ng/µl using a high-fidelity DNA polymerase. Finally, cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics), samples were pooled (equimolar) and the pool size was fractionated (200–550 bp) by preparative agarose gel electrophoresis before the samples were sequenced on an Illumina NextSeq 500 system using 1×75 bp read length.

Bioinformatics Analyses

Quality control was conducted on raw reads and after each processing step using fastqc (v0.11.9) ([2]accessed 09/09/20) and MultiQC (v1.10.1) (Ewels et al., 2016). Adapters were trimmed using trimgalore (v0.6.3) ([3]accessed 09/09/20) and reads were mapped against the *H. volcanii* genome with STAR (v2.7.3a) (Dobin et al., 2013). Unique reads were extracted and counted using featureCounts (v2.0.1) (Liao et al., 2014) and DESeq2 (v1.30.0) (Love et al., 2014) was used for differential expression analysis. Command lines for the processing steps are available in the **Supplementary Material**. A modular workflow assembler which was used to generate workflows for this analysis is available at https://github.com/jfallmann/MONSDA (Fallmann et al., 2022).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA776127; https://www.ebi.ac.uk/pride/archive/, PXD029952 (Vizcaíno et al., 2016).

---

[2] https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[3] https://github.com/FelixKrueger/TrimGalore

AUTHOR CONTRIBUTIONS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.822304/full#supplementary-material

REFERENCES

Arslan, Z., Wurm, R., Brener, O., Ellinger, P., Nagel-Steger, L., Oesterhelt, F., et al. (2013). Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2. *Nucleic Acids Res.* 41, 6347–6359. doi: 10.1093/nar/gkt315

Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., et al. (2011). 'A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair'. *Mol. Microbiol.* 79, 484–502.

Balakrishnan, L., and Bambara, R. A. (2013). 'Flap Endonuclease 1'. *Annu. Rev. Biochem.* 82, 119–138.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). 'CRISPR provides acquired resistance against viruses in prokaryotes'. *Science* 315, 1709–1712.

Beloglazova, N., Lemak, S., Flick, R., and Yakunin, A. F. (2015). "Analysis of nuclease activity of Cas1 proteins against complex DNA substrates," in *CRISPR: Methods and Protocols*, eds M. Lundgren, E. Charpentier, and P. C. Fineran (New York, NY: Humana Press-Springer). doi: 10.1007/978-1-4939-2687-9_16

Bernheim, A., Calvo-Villamañán, A., Basier, C., Cui, L., Rocha, E. P. C., Touchon, M., et al. (2017). 'Inhibition of NHEJ repair by type II-A CRISPR-Cas systems in bacteria'. *Nat. Comm.* 8:2094. doi: 10.1038/s41467-017-02350-1

Bitan-Banin, G., Ortenberg, R., and Mevarech, M. (2003). 'Development of a Gene Knockout System for the Halophilic Archaeon Haloferax volcanii by Use of the pyrE Gene'. *J. Bacteriol.* 185, 772–778. doi: 10.1128/JB.185.3.772-778.2003

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., et al. (2008). 'Small CRISPR RNAs guide antiviral defense in prokaryotes'. *Science* 321, 960–964. doi: 10.1126/science.1159689

Craggs, T. D., Hutton, R. D., Brenlla, A., White, M. F., and Penedo, J. C. (2014). 'Single-molecule characterization of Fen1 and Fen1/PCNA complexes acting on flap substrates'. *Nucleic Acids Res.* 42, 1857–1872. doi: 10.1093/nar/gkt1116

Delmas, S. I, Duggin, G., and Allers, T. (2013). DNA damage induces nucleoid compaction via the Mre11-Rad50 complex in the archaeon Haloferax volcanii. *Mole. Microbiol.* 87, 168–179. doi: 10.1111/mmi.12091

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). 'STAR: ultrafast universal RNA-seq aligner'. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Dyall-Smith, M. (2008). *The Halohandbook: protocols for haloarchaeal genetics.* Melbourne: Haloarchaeal Genetics Laboratory, 14.

Dyall-Smith, M., Pfeiffer, F., Chiang, P. W., and Tang, S. L. (2021). The Novel Halovirus Hardycor1, and the Presence of Active (Induced) Proviruses in Four Haloarchaea. *Genes* 2021, 12. doi: 10.3390/genes12020149

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). 'MultiQC: summarize analysis results for multiple tools and samples in a single report'. *Bioinform.* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354

Fallmann, J., Goldmann, R., and Stadler, P. F. (2022). *MONSDA, Modular Organizer of Nextflow and Snakemake Driven hts Data Analysis.* (in preparation)

Faure, G., Makarova, K. S., and Koonin, E. V. (2019a). 'CRISPR–Cas: Complex Functional Networks and Multiple Roles beyond Adaptive Immunity'. *J. Mole. Biol.* 431, 3–20. doi: 10.1016/j.jmb.2018.08.030

Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., et al. (2019b). 'CRISPR–Cas in mobile genetic elements: counter-defence and beyond'. *Nat. Rev. Microb.* 17, 513–525. doi: 10.1038/s41579-019-0204-7

Grasso, S., and Tell, G. (2014). 'Base excision repair in Archaea: Back to the future in DNA repair'. *DNA Repair* 21, 148–157. doi: 10.1016/j.dnarep.2014.05.006

Han, D., Lehmann, K., and Krauss, G. (2009). 'SSO1450–a CAS1 protein from Sulfolobus solfataricus P2 with high affinity for RNA and DNA'. *FEBS Lett.* 583, 1928–1932. doi: 10.1016/j.febslet.2009.04.047

Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., et al. (2015). 'Cas9 specifies functional viral targets during CRISPR-Cas adaptation'. *Nature* 519, 199–202. doi: 10.1038/nature14245

Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). 'The Biology of CRISPR-Cas: Backward and Forward'. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032

Ivančić-Baće, I., Cass, S. D., Wearne, S. J., and Bolt, E. L. (2015). 'Different genome stability proteins underpin primed and naïve adaptation in E. coli CRISPR-Cas immunity'. *Nucleic Acids Res.* 43, 10821–10830. doi: 10.1093/nar/gkv1213

Jevtić, Z., Stoll, B., Pfeiffer, F., Sharma, K., Urlaub, H., Marchfelder, A., et al. (2019). The Response of Haloferax volcanii to Salt and Temperature Stress: A Proteome Study by Label-Free Mass Spectrometry. *Proteomics* 19:1800491. doi: 10.1002/pmic.201800491

Kim, T. Y., Shin, M., Huynh, L., Yen, T., and Kim, J. S. (2013). Crystal structure of Cas1 from Archaeoglobus fulgidus and characterization of its nucleolytic activity. *Biochem. Biophys. Res. Commun.* 441, 720–725. doi: 10.1016/j.bbrc.2013.10.122

Klein, C. (2011). *Die Untersuchung des CRISPR/Cas Systems in Haloferax volcanii.* Germany: Ulm University.

Lestini, R., Duan, Z., and Allers, T. (2010). 'The archaeal Xpf/Mus81/FANCM homolog Hef and the Holliday junction resolvase Hjc define alternative pathways that are essential for cell viability in Haloferax volcanii'. *DNA Repair* 9, 994–1002. doi: 10.16/j.dnarep.2010.06.012

Levy, A., Goren, M. G., Yosef, I., Auster, O., Manor, M., Amitai, G., et al. (2015). 'CRISPR adaptation biases explain preference for acquisition of foreign DNA'. *Nature* 520, 505–510. doi: 10.1038/nature14302

Liao, Y., Smyth, G. K., and Shi, W. (2014). 'featureCounts: an efficient general purpose program for assigning sequence reads to genomic features'. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Liao, Y., Williams, T. J., Walsh, J. C., Ji, M., Poljak, A., Curmi, P. M. I, et al. (2016). 'Developing a genetic manipulation system for the Antarctic archaeon, Halorubrum lacusprofundi: investigating acetamidase gene function. *Sci. Rep.* 6:34639. doi: 10.1038/srep34639

Liu, T., Li, Y., Wang, X., Ye, Q., Li, H., Liang, Y., et al. (2015). 'Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR de novo

spacer acquisition'. *Nucleic Acids Res.* 43, 1044–1055. doi: 10.1093/nar/gku1383

Liu, T., Liu, Z., Ye, Q., Pan, S., Wang, X., Li, Y., et al. (2017). 'Coupling transcriptional activation of CRISPR-Cas system and DNA repair genes by Csa3a in Sulfolobus islandicus'. *Nucleic Acids Res.* 45, 8978–8992. doi: 10.1093/nar/gkx612

Liu, Z., Sun, M., Liu, J., Liu, T., Ye, Q., Li, Y., et al. (2020). 'A CRISPR-associated factor Csa3a regulates DNA damage repair in Crenarchaeon Sulfolobus islandicus'. *Nucleic Acids Res.* 48, 9681–9693. doi: 10.1093/nar/gkaa694

Love, M. I., Huber, W., and Anders, S. (2014). 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2'. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Maier, L.-K., Dyall-Smith, M., and Marchfelder, A. (2015). 'The Adaptive Immune System of Haloferax volcanii'. *Life* 5, 521–537. doi: 10.3390/life5010521

Maier, L.-K., Stachler, A.-E., Brendel, J., Stoll, B., Fischer, S., Haas, K. A., et al. (2019). 'The nuts and bolts of the Haloferax CRISPR-Cas system I-B'. *RNA Biol.* 16, 469–480. doi: 10.1080/15476286.2018.1460994

Makarova, K. S., Aravind, L., Grishin, N. V. I, Rogozin, B., and Koonin, E. V. (2002). 'A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis'. *Nucleic Acids Res.* 30, 482–496. doi: 10.1093/nar/30.2.482

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). 'A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action'. *Biol. Direct.* 1:7. doi: 10.1186/1745-6150-1-7

Marraffini, L. A., and Sontheimer, E. J. (2008). 'CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA'. *Science* 322, 1843–1845. doi: 10.1126/science.1165771

Meslet-Cladiére, L., Norais, C., Kuhn, J., Briffotaux, J., Sloostra, J. W., Ferrari, E., et al. (2007). 'A novel proteomic approach identifies new interaction partners for proliferating cell nuclear antigen'. *J. Mol. Biol.* 372, 1137–1148. doi: 10.1016/j.jmb.2007.06.056

Miller, J. H. (1972). *Experiments in Molecular Genetics.* Cold Spring Harbour, N. Y.: Cold Spring Harbour Laboratory Press.

Modell, J. W., Jiang, W., and Marraffini, L. A. (2017). 'CRISPR–Cas systems exploit viral DNA injection to establish and maintain adaptive immunity'. *Nature* 544, 101–104. doi: 10.1038/nature21719

Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014). 'Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity'. *Nat. Struct. Mol. Biol.* 21, 528–534. doi: 10.1038/nsmb.2820

Nussenzweig, P. M., and Marraffini, L. A. (2020). 'Molecular Mechanisms of CRISPR-Cas Immunity in Bacteria'. *Annu. Rev. Genet.* 54, 93–120. doi: 10.1146/annurev-genet-022120-112523

Pan, M., Kelman, L. M., and Kelman, Z. (2011). 'The archaeal PCNA proteins'. *Biochem. Soc. Trans.* 39, 20–24. doi: 10.1042/BST0390020

Rollie, C., Schneider, S., Brinkmann, A. S., Bolt, E. L., and White, M. F. (2015). 'Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *eLife* 4:e08716. doi: 10.7554/eLife.08716

Sambrook, J., and Russell, D. (2001). *Molecular Cloning: A Laboratory Manual.* New York, NY: Cold Spring Harbour Press.

Schulze, S., Adams, Z., Cerletti, M., De Castro, R., Ferreira-Cerca, S., Fufezan, C., et al. (2020). 'The Archaeal Proteome Project advances knowledge about archaeal cell biology through comprehensive proteomics'. *Nat. Comm.* 11:3145. doi: 10.1038/s41467-020-16784-7

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). 'Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels'. *Anal. Chem.* 68, 850–858. doi: 10.1021/ac950914h

Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., et al. (2016). '2016 update of the PRIDE database and its related tools'. *Nucleic Acids Res.* 44, D447–D456.

Wei, Y., Terns, R. M., and Terns, M. P. (2015). 'Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation'. *Genes Dev.* 29, 356–361. doi: 10.1101/gad.257550.114

Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W., and Doudna, J. A. (2009). 'Structural basis for DNase activity of a conserved protein implicated

in CRISPR-mediated genome defense'. *Structure* 17, 904–912. doi: 10.1016/j.str.
2009.03.019

Conflict of Interest: The authors declare that the research was conducted in the
absence of any commercial or financial relationships that could be construed as a
potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors
and do not necessarily represent those of their affiliated organizations, or those of
the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or
endorsed by the publisher.

Check for
updates

Plants Dictate Root Microbial Composition in Hydroponics and Aquaponics

Victor Lobanov[1], Karel J. Keesman[2] and Alyssa Joyce[1]*

[1] Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden, [2] Mathematical and Statistical Methods Group – Biometris, Wageningen University & Research, Wageningen, Netherlands

The role of the microbial community in mediating fish and plant co-culture is often considered the black box of aquaponics. Despite widespread recognition regarding the dependency of plants on their rhizosphere, the extent to which upstream aquaculture influences downstream hydroponic root communities has been poorly described in the literature. In this study we performed a taxonomic survey (16S rRNA metabarcoding) of microbial communities originating in the facility water source, hydroponic nutrient solution (HNS) sump, nutrient supplemented biofilter effluent (BF) sump, and recirculating aquaculture system tanks stocked with Nile tilapia (*Oreochromis niloticus*). Lettuce (*Lactuca sativa*) was then grown using the HNS and BF effluent under sterilized or mature (prior aquaponics/hydroponics lettuce culture water) conditions, likewise, the influence of probiotic addition or inoculation with soil-grown lettuce rhizosphere was assessed. Compositional similarities across treatments suggest that under soil-less conditions, plants are able to exert a stronger discriminatory influence on their rhizosphere composition than is done by colonization from upstream sources. Furthermore, cluster dendrograms grouped the sterilized and unsterilized treatments more consistently together than hydroponics and aquaponics treatments. These findings contradict conventional beliefs that microbial communities in the water column colonize roots based on their presence alone, ignoring the role that plants play in rhizosphere community selection.

Keywords: rhizosphere, community analysis, rhizobiome, aquaponics, hydroponics

INTRODUCTION

The region in and around plant roots, the rhizosphere, is an interspecies nutrient and electron trade zone with stakeholders representing all kingdoms (Hacquard, 2017; Wang et al., 2017; Garcia and Kao-Kniffin, 2018; Geisen et al., 2018; Guyonnet et al., 2018; Hu et al., 2018). Recent studies have shown that soil-based plants exert significant pressure in terms of nutrient composition on their rooting communities (Sasse et al., 2018; Zhalnina et al., 2018; Compant et al., 2019). The extent to which these findings may be transposed onto plants grown in soil-less cultivation conditions is less clear for two reasons. Firstly, it is unclear whether the release of soluble plant exudates into an aqueous milieu diminishes their effect on the microbial community. Secondly, the greater

ease by which the microbial community may be transferred within the aqueous environment could contribute to a greater capacity for root colonization.

The rhizosphere community (rhizobiome) manages nutrient uptake needs (Yang and Crowley, 2000; Scagliola et al., 2016; Garcia and Kao-Kniffin, 2018; Vadstein et al., 2018), abiotic stress resistance (Hussain et al., 2018; Sasse et al., 2018; Topalovic et al., 2020), and host defense (Gourion et al., 2015; Yasin and Ahmed, 2016; Elhady et al., 2018). It is composed of a core component fulfilling essential functions required by the plant at each stage of its growth, and a satellite component consisting of strains present at low abundances (Compant et al., 2019). The core community consists of taxa that are necessarily drawn to the root environment in contrast to bulk soil (Yeoh et al., 2017). As only 7% of bulk soil microorganisms are found in the rhizosphere (DeAngelis et al., 2009), the carbon-rich environment of the rhizosphere has been described as a precursory selection pressure. The relatively stable flow of 10–250 mg/g organic acids from the plant into the rhizosphere enriches microbial taxa two orders of magnitude greater than surrounding soil (Lynch and de Leij, 2001), with root exudates including amino acids, organic anions, sugars (Phillips et al., 2004; Badri and Vivanco, 2009; Kawasaki et al., 2016; Jacoby et al., 2020). The complex dynamics of rhizobiome development has given rise to many metagenomic studies on the rhizosphere (Kawasaki et al., 2018; Sasse et al., 2018; Vadstein et al., 2018; Ayipio et al., 2019; Compant et al., 2019). Research on soil-based studies indicates that investments into the root community is a high priority for terrestrial plants, but it is not evident how well this relationship is preserved in a nutrient solution environment such as soil-less hydroponic or integrated agriculture systems (e.g., aquaponics). Furthermore, the capacity of probiotics to mediate host plant/rhizosphere interactions was explored through the application of the commercially relevant bacterium *Bacillus amyloliquefaciens*, which has been developed as a probiotic in hydroponics but not aquaculture (Kidoglu et al., 2008; Nautiyal et al., 2013; Chowdhury et al., 2015).

In this study, a decoupled aquaponics design was used to study downstream colonization of the rhizosphere by upstream microbial communities (Goddek et al., 2019a). From a nutrient perspective, there are two inputs: fish feed for the aquaculture unit and any fertilizer addition in the hydroponics unit. Sources for microbial inoculation may arise from the local aqueous or airborne environment, as well as through the import of foreign material into the system (i.e., via feed). Recent publications focusing on the diversity of microorganisms in aquaponic systems have given rise to many hypotheses as to how the microbial community may lead to increased performance based on the increased abundance of chelating agents, cofactors, enzymes, or hormones facilitating nutrient bioavailability, either directly or indirectly (Munguia-Fragozo et al., 2015; Rurangwa and Verdegem, 2015; Schmautz et al., 2017; Sheridan et al., 2017; Kasozi et al., 2021). While the microbial community is widely recognized as important to the success of aquaponic systems (Delaide et al., 2016, 2017; Goddek et al., 2016b; Wielgosz et al., 2017;

Bartelme et al., 2018; Goddek and Vermeulen, 2018; Eck et al., 2019), it has likewise been suspected as a vector for pathogen proliferation (Mori and Smith, 2019; Kasozi et al., 2021).

With the objective of determining the source of the microbial community colonizing the rhizosphere, lettuce (*Lactuca sativa*) was grown under a variety of hydroponic conditions including nutrient supplementation with a commercial hydroponic solution alone, nutrient supplemented aquaculture-derived water stocked with Nile tilapia (*Oreochromis niloticus*) and after inoculation with a probiotic or soil culture. Through multiple discriminating analyses (cluster dendrogram, principal component analysis), this study highlights the important role of plants in determining their own rhizosphere composition in soil-less cultivation systems.

MATERIALS AND METHODS

A decoupled (unidirectional flow) aquaponics system was stocked with Nile tilapia (*Oreochromis niloticus*) and Batavian lettuce (*Lactuca sativa*) Exaudio RZ 79-43 (Rijk Zwaan, Netherlands) grown at the Wageningen UR Greenhouse Horticulture Unit (Bleiswijk, Netherlands). The lettuce was grown in hydroponic boxes (3 plants/ea.) with three replicates per treatment. Boxes, insomuch as they were self-contained provided better control over microbial exposure to the plants than normal media-based, raft or nutrient film systems, but did not completely prevent bacterial transfer as growth conditions were not sterile, nor were seeds sterilized prior to planting. Each box contained a Styrofoam sheet floating on nutrient solution, mimicking a deep-water culture environment. Four microcentrifuge tubes with sheared tips were filled with 2% w/v agar-agar (Sigma, Netherlands) and inserted into the sheet with seeds immersed in the agar. Roots growing into the aqueous milieu were considered to be representative of the plants' rhizosphere, as this most closely resembles root structure in hydroponic cultivation conditions.

For all treatments, seeds were incubated in darkness overnight (8 h) at 25°C. Filter sterilized (0.22 μm) hydroponic nutrient solution (HNS) was added to each box at the beginning of cultivation and exchanged for the treatment-specific nutrient solution after 2 days. Nutrient solutions were prepared weekly, at which time half of the volume was exchanged. Supplementation of the sump solution was done as necessary to maintain the following approximate macronutrient composition (mmol/L): 15.0 NO_3, 1.5 NH_4, 5.0 K, 1.5 Na, 3.0 Ca, 1.5 Mg, 0.1 Si, 0.1 Cl, 1.5 SO_4, 0.5 HCO_3, 0.5–1.0 P. The following trace elements set points were also maintained (μmol/l): 20.0 Fe, 7.0 Mn, 5.0 Zn, 20.0 B, 0.5 Cu, 0.1 Mo, while pH (set to 6–7) and EC (set to 2–2.5 mS/cm) were adjusted as needed to maintain desired ranges. Studies directly comparing yields between aquaponics and hydroponics have proven difficult to reproduce (Ayipio et al., 2019; Yep and Zheng, 2019). As most aquaponic and hydroponic systems strive to maximize crop productivity through the same conventional means (greenhouse design, cultivar selection, etc.),

nutrient concentrations were kept constant in this study to avoid confounding the relationship between nutrient loading and plant health.

Treatments were watered from either the aquaponics system (BF) or a commercial hydroponic nutrient solution (HNS) (**Figure 1**). Aquaponics crops received effluent from the biofilter, with nutrient supplementation carried out in a decantation tank prior to the hydroponics unit. Here, we refer to HNS from two full crop cycle as mature HNS (HNS.m). To make sterilized HNS (HNS.s) or BF (BF.s), freshly made nutrient stock solutions were filter sterilized (0.22 μm). The probiotic effect of *B. amyloliquefaciens* was added to sterilized HNS and to unsterilized BF (corresponding to treatments Probio.s and Probio.m, respectively). A DSMZ (Germany) culture stock of *B. amyloliquefaciens* (ex Fukumoto 1943) grown in pure culture to 5×10^{11} CFU/g stock was applied to achieve a final concentration of 2 mg/L. Soil inoculum (ca. 50 mg) was sourced from Batavian lettuce grown in potting soil for 4 weeks. The soil sample was sequenced as a control (referred to as "Soil"), inoculated treatments are referred to as Soil.inoc. The water column from Batavian lettuce grown aquaponic (BF.aqueous) and hydroponic (HNS.aqueous) basins were furthermore sampled as a control for the pelagic microbial community, as was the facility water source (WS) and the aquaculture tanks (RAS).

Water samples during all three trials were analyzed weekly for nutrient concentrations, pH, and EC (Groen Agro Control, Netherlands). Dissolved oxygen (DO) was kept saturated for both experiments. Temperature was controlled at 16°C. Broad spectrum lighting was maintained at 200 μmol/s/m^2 for 16 h/day for all trials, although supplemental lighting was not used for trial 2 (due to summer conditions providing adequate irradiation). Crops were harvested after 6 weeks.

For microbial community profiling, DNA was isolated from the roots of each technical replicate using the DNeasy PowerSoil Kit (Qiagen, Germany). All plant roots in an individual box (technical replicate) were combined for DNA extraction. Roots were lightly shaken but not directly dried so as not to influence the rhizosphere community prior placement inside the microcentrifuge tube used in the protocol. A noticeable film of water enveloped the roots after shaking; 0.25 g of wet roots were used for the DNA extraction. The PowerSoil kit was chosen as it is well adapted to extract DNA from complex matrices such as the extracellular polymeric substances consistent with biofilm structure. For soil samples (soil inoculum referred to above as "Soil"), 0.25 g of soil from around the root was used. Purified DNA was PCR amplified using universal 16S rDNA bacterial primers (**Table 1**) targeting the V3-V4 region of the 16S rRNA gene. Primers were provided by BaseClear B.V. (Netherlands) and sequenced using their MiSeq system. Sequenced operational taxonomic units (OTUs) were processed as per BaseClear protocols whereby sequenced amplicons are merged into overlapping pseudo-reads and subsequently aligned against the NCBI 16S rRNA database for putative taxonomic identification.

In R, the OTU data set was subdivided into six data frames related to the taxonomic rank using the Tidyverse packages tidyr

and dplyr. Subsequent analyses were restricted to the genus, family, and order ranks as a compromise between the large amount of OTUs generated in the data set (obscuring clear visualization of the data) and to avoid a lack of resolution as occurring at higher ranks. Firstly, vegan was used for the diversity analysis, ade4 and labdsv were used for multivariate data analyses, pvclust for hierarchical cluster analysis, vegclust and vegsoup for data clustering, picante for community analysis, and finally corrplot for the correlation plot. Packages for visualization of the data included gclus to generate the clustering graphics, dendextend for dendrograms, and ggplot2 for the correlation plot.

Due to the effect of outliers, several normalization strategies were explored: presence/absence, maximum abundance per treatment, relative abundance per species, relative frequency per site, normalization to the Euclidian norm (Chord transformation), normalization to the relative frequency per site (Hellinger transformation), double profile normalization (Chi-squared transformation), and normalization first by species maxima then by site totals (Wisconsin standardization). Normalization by Hellinger transformation were chosen for this study based on the tightness of the variance range in the processed data sets.

Three types of neighbor clustering were used to organize the data: nearest, furthest, and Ward. Nearest neighbor clustering agglomerates groups based on the shortest pairwise dissimilarities between members, while the furthest neighbor method defines the group membership based on the maximum distance between any two clusters. Ward's minimum variance clustering minimizes the total within-cluster variance and appeared the most logical to follow based on the robustness of the groups. The optimal number of clusters were calculated using Ward correlation, Pearson correlation, IndVal method, simple structure index (ssi) criteria, and Calinski criteria (Dolnicar et al., 2000; Borcard et al., 2018). While the range of optima was fairly consistent across taxonomic ranks, the optimization algorithms never converged on a single figure. The clustering result was then independently confirmed by a principal component analysis and correlation plot of the treatments. Finally, network analyses of both the treatments and microbial taxa allowed us to visualize which treatments most closely resemble each other at different taxonomic ranks.

RESULTS

In terms of plant health, treatments were not nutrient limited nor displayed signs of disease. Lacking obvious indications of stress, it was assumed that plants interacted with the surrounding microbial environment under homogenous circumstances across treatments with differences in community composition originating from the source water and not physicochemical or stress factors. To elucidate the relationship between the host plants and the composition of the rhizosphere microbial community, this study investigated patterns in taxonomic prevalence across treatments through hierarchal classification and clustering analyses.

FIGURE 1 | Summary of treatments in the current study.
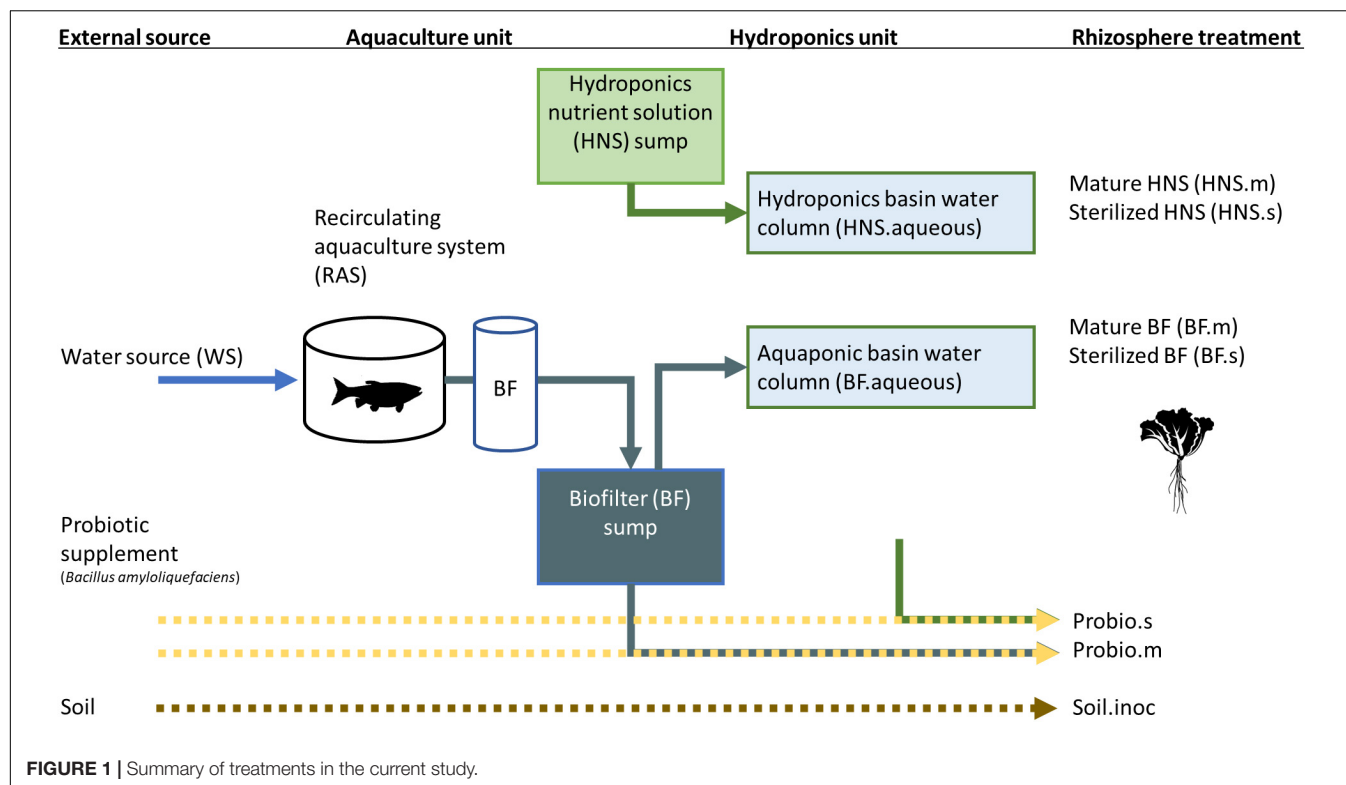
TABLE 1 | Primer sequences used for the taxonomic community analysis in this study.

Domain target	Direction	Sequence	Length (bp)	Melting temperature	GC%
Bacteria	Fwd primer	AGAGTTTGATCCTGGCTCAG	20	56.92	50.00
	Rv primer	ATTACCGCGGCTGCTGG	17	60.18	64.71

Plotting the distribution of OTUs across the treatments provided evidence for the existence of a core microbiome present in many (9–10 out of 28 treatments), although most OTUs are unique to 1–2 treatments (**Figure 2**); plots of the family and order rank (**Supplementary Figures 1**, **2**) were similar but not identical. Approximate unbiased (red, significant ≥ 0.95) and bootstrap probability (green, value indicates the amount of bootstrapping until robust) p-values for the edge dendrograms (edge number in gray) are indicated. Blue squares indicate significance at $p \geq 0.90$ with red bars indicating high robustness at $p \geq 0.95$. At the genus rank it is visible that the source of colonization does not strongly predict clustering. Aquaculture derived water appears to influence community dynamics whether sterilization is imposed (BF.s) or not (BF.m) (group 2), however, mature BF or HNS box communities (group 3) were not closely related to the aqueous community used to inoculate the boxes. The control treatments (BF, WS, HNS, RAS, Soil) cluster similarly (group 4), with the aqueous communities (aquaculture linked or independent) clustering closely together. Probiotic supplementation mostly clustered in group 5, however, some branches were mixed with other treatments. At the family rank (**Supplementary Figure 1**) no clear pattern was visible, although it is visible that the probiotic treatments populated one branch at the first fork, while the controls and most mature and sterilized HNS treatments populated the second fork.

Partitioning based on ssi criteria resulted in multiple equally optimal partitions for a range of cluster objects. Ultimately, this indicates a high degree of interchangeability between most treatments, suggesting that the microbial communities present are more similar than different. Looking at a dissimilatory matrix of the treatments (**Figure 3**), we see that the aquaculture impacted (BF series) and probiotic supplemented (Probio) treatments tend to be more similar within themselves that to each other, with the soil and standard hydroponics (HNS series) being less cohesive groups. The principle component analysis (**Figure 4**) places the mature HNS treatments (HNS.m) at the center of the distribution, with the two most discriminating factors at 23.8% (dimension 1) and 15.5% (dimension 2) causing a split between the controls (WS, RAS, BF, HNS, Soil, HNS.aqueous, and BF.aqueous) and experimental treatments (HNS.m, HNS.s, BF.m, and BF.s). The probiotic (probio) and soil inoculated (soil) treatments were less dependent on the two principal dimensions.

Finally, co-occurrence networks were generated for microbial communities. At higher ranks the superstructure for community similarity across treatments is more clearly defined. At the order rank this appears as three clusters, two of which are more closely

FIGURE 2 | Cluster dendrogram of the distribution of microbial communities at the genus rank across treatments with the five most robust clades highlighted. Similar patterns were observed at higher ranks. Treatments include hydroponic nutrient solution sump (HNS) and biofilter effluent sump (BF) under mature (.m), sterilized (.s), and basin water column (.aqueous) conditions. Additionally, soil inoculum (Soil) and HNS inoculated culture (soil) and probiotic (probio) inoculated sterilized (.s) and unsterilized biofilter effluent (BF) samples, as well as the facility water source (WS) and recirculating aquaculture system water column (RAS) are also included.

related (**Figure 5A**). At lower ranks (**Figure 5B**) these clusters begin to splinter as the quantity of unique labels corresponding to microbial taxa increases exponentially.

DISCUSSION

This study is the first to investigate how the rhizosphere microbial community is shaped by upstream influences under soil-less cultivation conditions. Lettuce (*Lactuca sativa*) was grown hydroponically or in aquaponics co-culture with Nile tilapia (*Oreochromis niloticus*). As shown in **Figure 1**,

treatments included nutrient supplementation with a commercial hydroponic solution alone, nutrient-supplemented aquaculture-derived water, or the commercial nutrient solution inoculated with a probiotic or soil culture. Filter sterilization vs. inoculation with mature media (nutrient solution derived from a previously harvested lettuce culture) were tested for both hydroponic and aquaponic treatments as well as the probiotic addition.

As indicated by the cluster dendrogram (**Figure 2**), no divisive split grouping all aquaponic (BF.m and BF.s) apart from all commercial hydroponic (HNS.m and HNS.s) treatments exists at the genus rank despite a highly robust clustering model with a cophenetic correlation of 0.93, a pattern consistent across
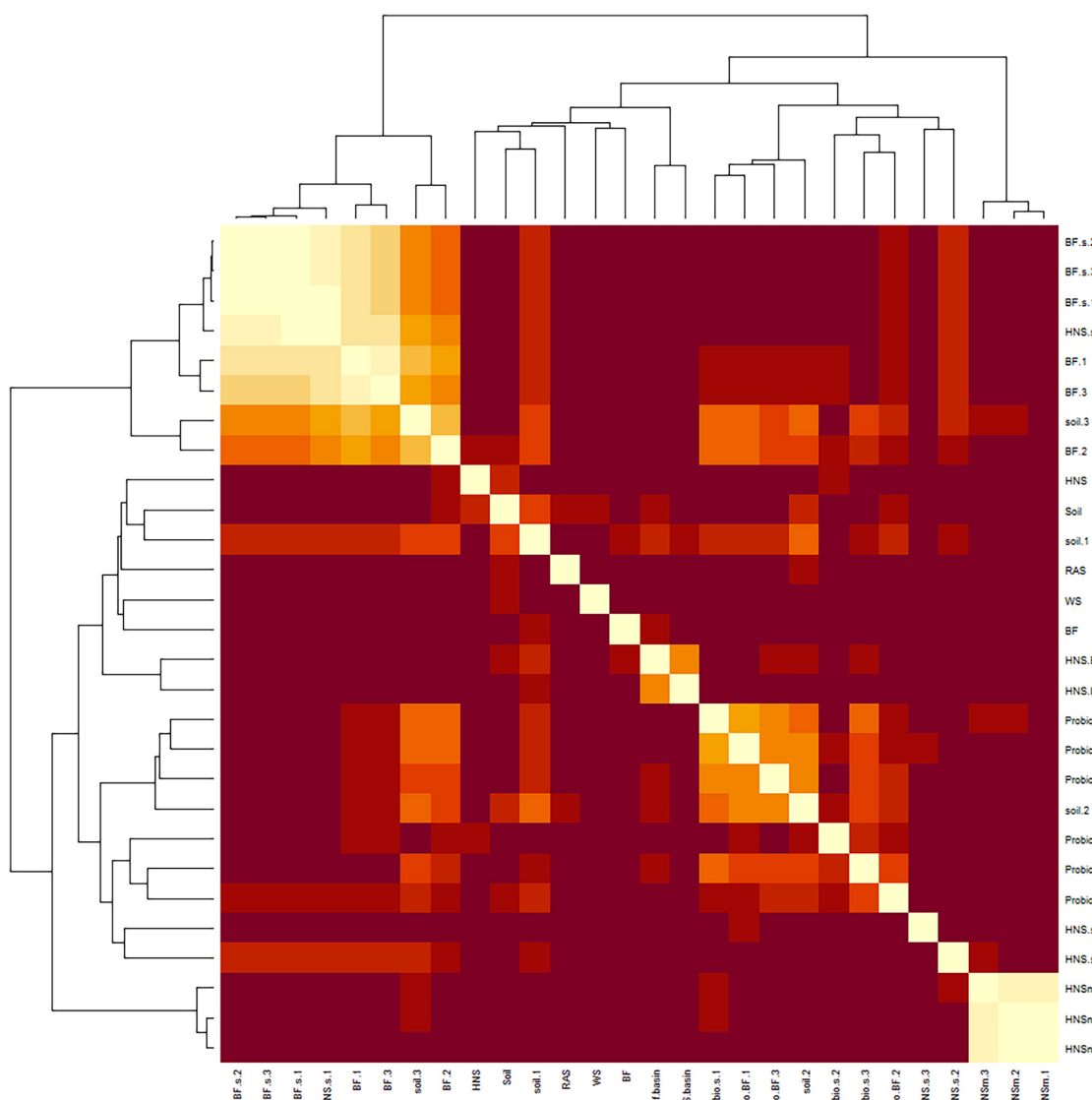
FIGURE 3 | Dissimilatory matrix between microbial communities of the treatments in the study at the genus rank. Similar patterns were observed at higher ranks. Treatments include hydroponic nutrient solution sump (HNS) and biofilter effluent sump (BF) under mature (.m), sterilized (.s), and basin water column (.aqueous) conditions. Additionally, soil inoculum (Soil) and HNS inoculated culture (soil) and probiotic (probio) inoculated sterilized (.s) and unsterilized biofilter effluent 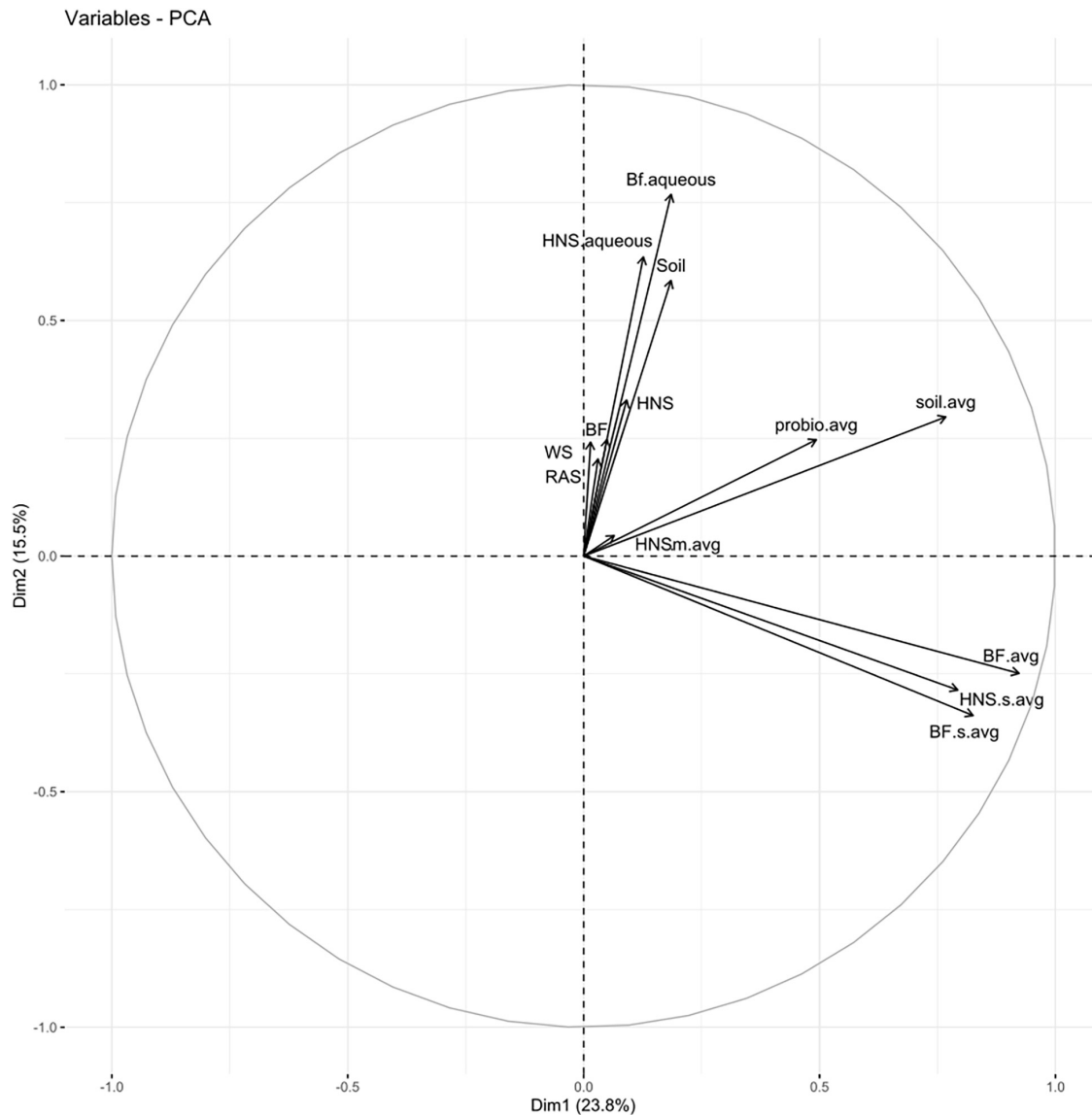(BF) samples, as well as the facility water source (WS) and recirculating aquaculture system water column (RAS) are also included.

different clustering methods and at higher taxonomic ranks. The number of optimal clusters, however, varied from 2 to 9 clusters between the five methods tested, mirroring the overall dendrogram shape when viewed based on cluster height. After an initial branching into 2–4 groups, the height difference between clusters drops sharply – reflecting a higher degree of replaceability. Either as part of a main or sub-branch, the controls (BF, WS, RAS, HNS, and Soil) tended to cluster closely. These controls mainly serve to identify environmentally prevalent microorganisms from the water supply (WS), aquaculture unit (RAS), nutrient-supplemented biofilter effluent sump (BF), hydroponic nutrient solution sump (HNS), and local soil-based lettuce rhizosphere (Soil). Their high degree of similarity at

low taxonomic ranks suggests that most microorganisms are ubiquitously present, in agreement with the rare biosphere ecological model (Lynch and Neufeld, 2015; Jousset et al., 2017).

Another perspective of community similarity is portrayed in the dissimilatory matrix (**Figure 3**), comparing treatments by virtue of their degree of similarity instead of being clustered based on a threshold consensus as well as the co-occurrence networks (**Figures 5A,B**), where clustering is allowed to overlap if treatments are sufficiently similar. At the class rank (**Figure 5A**), an agglomeration of the aquaponics treatments (3 BF.s, 2 BF.m) is visible in cluster III. The BF control and BF water column samples were distinct from this group (cluster I), sharing a greater degree of similarity
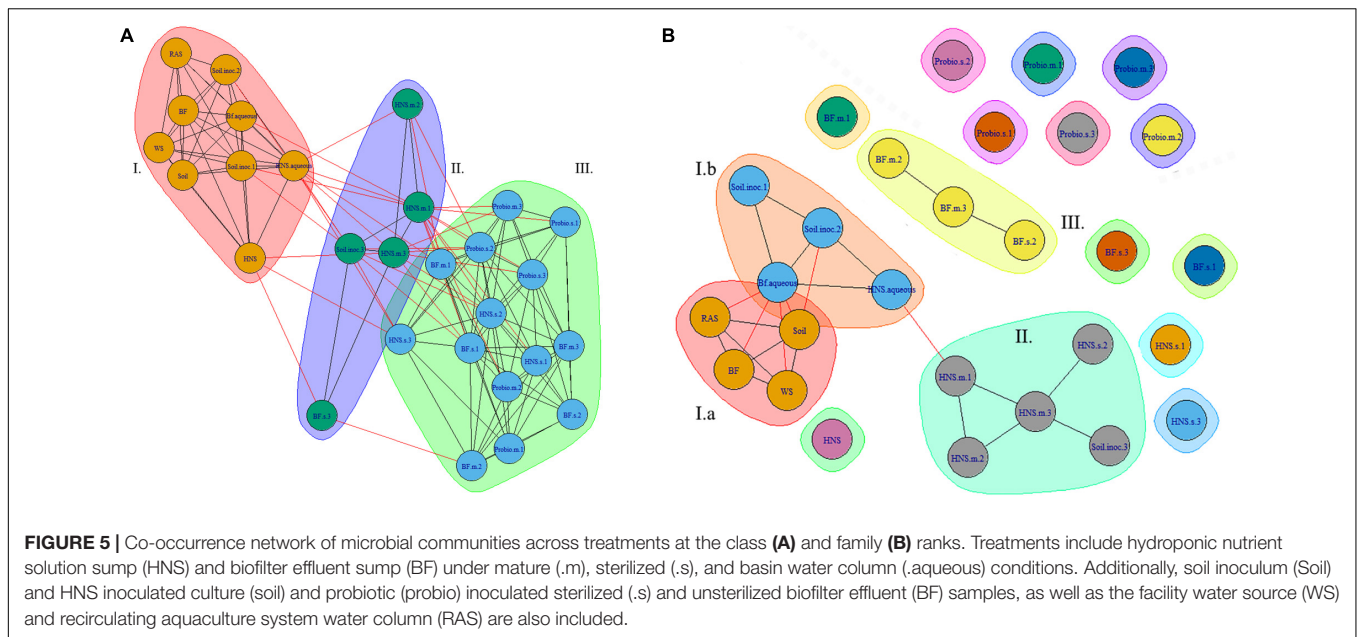
FIGURE 4 | Principle component analysis for all treatments; abundance data across technical replicates were averaged for each set. Treatments include hydroponic nutrient solution sump (HNS) and biofilter effluent sump (BF) averaged for each technical replicate and basin water column (.aqueous) conditions. Additionally, soil inoculum (Soil) and HNS inoculated culture (soil) and probiotic (probio) samples averaged for all technical replicates, as well as the facility water source (WS) and recirculating aquaculture system water column (RAS) are also included.

with the other controls instead. While the mature HNS treatments (HNS.m, cluster II) clustered together, their cluster partially overlapped with cluster III containing the three HNS.s treatments. The probiotic treatments clustered together (cluster III), however, the soil treatments were distributed between clusters I and II. Much of the cluster similarity disappeared at the genus rank, albeit the nodes of clusters I and II are still visible (**Figure 5B**).

From the PCA (**Figure 4**) and the co-occurrence network analysis at the genus rank (**Figure 5B**), it is visible that the probiotic treatments poorly grouped together on average, meaning that their taxonomic composition was the least

consistent within technical replicates. One may speculate that the colonizing influence of the probiotic shifted the community as a whole, rearranging the rhizobiome into a different configuration than in other treatments. Insofar as this may be attributed to the probiotic itself is outside the scope of this study.

Soil treatments did not consistently cluster together (**Figures 2**, **4**, **5**), ostensibly reflecting the shift in community composition from bulk soil to the rhizosphere environment as described elsewhere (DeAngelis et al., 2009). Community diversity was poorly retained when soil-based lettuce roots were used to inoculate sterile HNS. These treatments gravitated toward the same global consensus as the other hydroponic treatments

FIGURE 5 | Co-occurrence network of microbial communities across treatments at the class **(A)** and family **(B)** ranks. Treatments include hydroponic nutrient solution sump (HNS) and biofilter effluent sump (BF) under mature (.m), sterilized (.s), and basin water column (.aqueous) conditions. Additionally, soil inoculum (Soil) and HNS inoculated culture (soil) and probiotic (probio) inoculated sterilized (.s) and unsterilized biofilter effluent (BF) samples, as well as the facility water source (WS) and recirculating aquaculture system water column (RAS) are also included.

rather than forming a robust branch independently, despite filter sterilization of the HNS and no direct contact between media. Although a portion of the HNS sump microbiome is shared with HNS.m and HNS.s treatments, several taxa undergo major shifts in abundance during this transition. As a model this suggests that under similar nutrient concentrations, rhizosphere involvement plays a greater role in driving microbial community composition than water source.

An array of factors influencing rhizobiome composition have been identified, originating both from the plant (genotype, life stage) and the environment (water source, nutrient profile) (Compant et al., 2019; Chen et al., 2020). As shown by Bartelme et al. (2019), facility conditions strongly dictate the microbial populations present in RAS and aquaponic systems. Our results suggest that a similar facility-specific microbiome forms within the rhizosphere in hydroponic systems. Studies on the rhizobiome in other type of cultivation systems such as soil or air have indicated a similar pattern of consolidation. For instance, Schreiter et al. (2014) observed that the lettuce rhizobiome was consistent across varying soil types, while Edmonds et al. (2020) observed a rhizobiome unique from the circulating nutrient solution that formed after 12 days of plant growth in aeroponic conditions (Edmonds et al., 2020). This trend appears to be a hallmark of terrestrial plants (Berg et al., 2014; Vandenkoornhuyse et al., 2015; Sasse et al., 2018; Zhalnina et al., 2018). In combination with the results from this study, it appears that selection pressures exerted by the plant to consolidate the rhizobiome around a particular profile are a fundamental aspect of plant physiology despite the influence of the exogenous microbial environment. That profile, although observed as a collection of taxa, mirrors the functional needs required by the plant at a particular life stage and under particular environmental conditions.

At a more global level, microbial communities will occupy all available niches as they become available. For instance, among its many discoveries, the Tara Oceans project revealed that physiochemical parameters such as pH and temperature play a more decisive role in the relative taxonomic abundance than does taxonomic presence (Gorsky et al., 2019; Sunagawa et al., 2020). Co-occurrence networks (**Supplementary Figure 3**) at the family and order rank indicate consistent grouping of certain microbial clades. However, further research should combine our top-down approach with a bottom-up strategies to study community organization [e.g., identification of keystone species (Herren and McMahon, 2018)], as well as omics based techniques for community functional analysis, to elucidate how select microorganisms or clades may impact facility productivity through their disproportionate influence on community structure.

Understanding the potential impact of upstream microbial communities on downstream hydroponic units has direct implications for preventative disease management. Demonstrating that the rhizosphere community composition is associated with the plant more strongly than the presence of exogenous colonizing bacteria implies that focusing efforts on supporting plant health rather than on water sterilization will better protect crops. Sterilization of incoming water and media is widely used in hydroponics to discourage the proliferation of pathogens (Ehret et al., 2001; Shimizu et al., 2007; Raudales et al., 2014; Liu and Huang, 2019; Zheng et al., 2019) albeit at the cost of reducing overall microbial diversity – both beneficial and harmful microorganisms – potentially opening niches for rapid colonization by r-strategists (Vadstein et al., 2018).

Some aquaponic studies advocate for continuous cycling of water between RAS and HP components (coupled aquaponics) (Nichols and Savidov, 2011; Palm et al., 2019), while others have advocated for a discrete separation (decoupled aquaponics)

with no return of water and hence microorganisms from the HP to the RAS (Delaide et al., 2016; Goddek et al., 2016a, 2019b; Goddek and Keesman, 2018; Goddek and Korner, 2019; Monsees et al., 2019). In this context, we sought to determine whether sterilization (reducing microbial proliferation across units) succeeds in significantly shaping the microbial community structure. Clustering did not indicate a mature/sterilization split at the genus (**Figure 2**), family, or order ranks (**Supplementary Figures 1, 2**), nor was a strong split visible via the dissimilatory matrix (**Figure 3**). Most treatments furthermore clustered together at the class rank (**Figure 5A**), with the notable exception being the mature HNS treatments in cluster II. In a prior investigation into the effect of sterilization in the context of RAS coupling, Wielgosz et al. (2017) concluded that the beneficial effects on plant growth from RAS effluent were most likely conferred through microbial exudates, and thus unaffected by the sterilization process itself. While the identity of those exudates remains unknown, our results further support their hypothesis by showing that the community composition is not principally determined by the source water (HNS/BF) or source community (mature/sterilized).

In terms of microbial compositional diversity, the most profound shift occurred between controls and treatments in a stepwise manner (**Supplementary Figure 3**). The soil control indicated a high level of diversity with a couple phyla disappearing in soil-inoculated treatments (*Fibrobacteres, Nitrospinae*), however, the majority of phyla were present at reduced concentrations. The facility water supply control (WS) was relatively enriched with some phyla compared to the recirculating aquaculture system (RAS) and biofilter (BF): *Bacillariophyta, Chlamydiae, Aquificae, Candidatus Saccharibacteria*. The RAS and BF conditions enriched the phyla *Fusobacteria, Nitrospirae,* and *Lentisphaerae*. Few members of these phyla could be detected in subsequent aqueous (BF.aqueous) or rhizosphere (BF.1-3) environments suggesting a lack of viability in the oligotrophic, ammonia-poor, hydroponic environment. Probiotic treatments (Probio.m.1-3, Probio.s.1-3) most significantly perturbed the total microbial composition. While no mechanism could be determined within the context of this study, probiotics have been shown to restrict microbial diversity in the gut microbiome (Suez et al., 2018). Regardless, further studies must corroborate our results to elucidate the relationship between probiotic administration and community diversity. In terms of co-occurrence, no clear patterns of taxonomic clustering could be discerned. At high taxonomic ranks, the amount of overlap consolidates most taxa together while at lower ranks the diversity creates an unmanageable number of sub-groups. At the order rank, some discernable clustering is visible (**Supplementary Figure 4**), however, the significance of these co-occurrences could not be determined within the scope of this study.

Our study focused on the microbial community dynamics at the main interface between the aqueous milieu and the plant in soil-less cultivation systems – the rhizosphere. The above trends indicate community consolidation in our system, suggesting that prioritizing plant health metrics will likewise reduce the potential for disease. We have recently demonstrated that trace nutrients are not taken up by plants proportionally to their external aqueous concentrations (Lobanov et al., 2021), which suggests that fundamental issues such as plant nutritional needs should be prioritized. Given the slow growth requirements of k-strategists [e.g., anammox (Laureni et al., 2015; Zhang et al., 2017; Zielińska et al., 2018), archaea (Seyler et al., 2014; Bartelme et al., 2017)], system-wide maturation of the microbial population may take months or years (Savidov and Brooks, 2004). It would not be unreasonable to expect successive waves of colonization to mark this period, as is similarly observed within the rhizobiome during plant growth (Micallef et al., 2009; Kristin and Miranda, 2013; Chaparro et al., 2014). Archaea and eukaryotic phyla (algae) were observed in the study at the phylum rank (**Supplementary Figure 4**), however, their contribution to rhizosphere structure, organization, and nutrient flow in aqueous environments remains an open question. While not investigated here, community succession in the rhizobiome during facility maturation may indicate the duration within which a facility microbiome stabilizes and thus is able to maximally resist pathogen colonization.

CONCLUSION

In this study, we have provided evidence that plant crop health is poorly predicted by exposure to upstream microbial communities in soil-less aquaponic cultivation systems. This study is the first to address the question of rhizosphere-colonizing microbial transfer in aquaponics by selectively exposing hydroponically grown plants to a range of treatments intended to shape the root microbiome. Prior literature has suggested that upstream aquaculture directly contributes to crop productivity through microbial colonization (Bartelme et al., 2018; Eck et al., 2019), or in other cases, may represent an entry point for pathogens into the system (Mori and Smith, 2019; Kasozi et al., 2021). While our data do not exclude these possibilities, they instead suggest that the introduction of upstream bacteria is less impactful than previously assumed. More likely, plant health weaknesses are exploited by pathogenic microorganisms ubiquitous in the local environment, thus not uniquely introduced through the water column. We expect the findings of this study to be transferable to cultivation conditions where healthy plants are not subject to excessive stress (i.e., due to nutrient deficiency or other water quality perturbations), however, future research must investigate how these systems respond to acute abiotic or biotic stressors.

This work paves the way for two important future directions. Firstly, our study suggests that aqueous nutrient concentration play a more predicative role in determining community composition than sterilization. While sterilization is a routine technique in aquaculture as well as hydroponics, it is nonetheless a tradeoff between pathogen suppression and total microbial diversity reduction. Future studies must likewise determine whether aquaponic facilities benefit from sterilization, or whether the co-cultivation of plants and fish in an environment promoting diversity leads to a more resilient facility-wide microbiome. Secondly, in line with previous work on the relationship between aqueous nutrient concentrations and plant health (Lobanov et al., 2021), more research is needed to determine whether a greater focus on maintaining plant health as opposed to only maximizing

yield will lead to more disease-tolerant crops, and ultimately more productive crops.

DATA AVAILABILITY STATEMENT

AUTHOR CONTRIBUTIONS

AJ and VL conceived the presented idea. VL carried out the experiments and wrote the manuscript with input from all authors.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.848057/full#supplementary-material

Supplementary Figure 1 | Cluster dendrogram of the distribution of microbial communities at the order rank across treatments with the five most robust clades highlighted. Similar patterns were observed at higher ranks. Treatments include hydroponic nutrient solution sump (HNS) and biofilter effluent sump (BF) under mature (.m), sterilized (.s), and basin water column (.aqueous) conditions. Additionally, soil inoculum (Soil) and HNS inoculated culture (soil) and probiotic (probio) inoculated sterilized (.s) and unsterilized biofilter effluent (BF) samples, as well as the facility water source (WS) and recirculating aquaculture system water column (RAS) are also included.

Supplementary Figure 2 | Cluster dendrogram of the distribution of microbial communities at the family rank across treatments with the five most robust clades highlighted. Similar patterns were observed at higher ranks. Treatments include hydroponic nutrient solution sump (HNS) and biofilter effluent sump (BF) under mature (.m), sterilized (.s), and basin water column (.aqueous) conditions. Additionally, soil inoculum (Soil) and HNS inoculated culture (soil) and probiotic (probio) inoculated sterilized (.s) and unsterilized biofilter effluent (BF) samples, as well as the facility water source (WS) and recirculating aquaculture system water column (RAS) are also included.

Supplementary Figure 3 | Distribution of phyla across treatments and controls.

Supplementary Figure 4 | Co-occurrence network of microbial taxa at the order rank across treatments.

REFERENCES

Ayipio, E., Wells, D. E., McQuilling, A., and Wilson, A. E. (2019). Comparisons between aquaponic and conventional hydroponic crop yields: a meta-analysis. *Sustainability* 11:6511.

Badri, D. V., and Vivanco, J. M. (2009). Regulation and function of root exudates. *Plant Cell Environ.* 32, 666–681. doi: 10.1111/j.1365-3040.2008.01926.x

Bartelme, R. P., McLellan, S. L., and Newton, R. J. (2017). Freshwater recirculating aquaculture system operations drive biofilter bacterial community shifts around a stable nitrifying consortium of ammonia-oxidizing archaea and comammox nitrospira. *Front. Microbiol.* 8:101. doi: 10.3389/fmicb.2017.00101

Bartelme, R. P., Oyserman, B. O., Blom, J. E., Sepulveda-Villet, O. J., and Newton, R. J. (2018). Stripping away the soil: plant growth promoting microbiology opportunities in aquaponics. *Front. Microbiol.* 9:8. doi: 10.3389/fmicb.2018.00008

Bartelme, R. P., Smith, M. C., Sepulveda-Villet, O. J., and Newton, R. J. (2019). Component microenvironments and system biogeography structure microorganism distributions in recirculating aquaculture and aquaponic systems. *Msphere* 4:e00143-19.

Berg, G., Grube, M., Schloter, M., and Smalla, K. (2014). Unraveling the plant microbiome: looking back and future perspectives. *Front. Microbiol.* 5:148. doi: 10.3389/fmicb.2014.00148

Borcard, D., Gillet, F., and Legendre, P. (2018). *Numerical Ecology with R*. Berlin: Springer.

Chaparro, J. M., Badri, D. V., and Vivanco, J. M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* 8, 790–803. doi: 10.1038/ismej.2013.196

Chen, T., Nomura, K., Wang, X., Sohrabi, R., Xu, J., Yao, L., et al. (2020). A plant genetic network for preventing dysbiosis in the phyllosphere. *Nature* 580, 653–657. doi: 10.1038/s41586-020-2185-0

Chowdhury, S. P., Hartmann, A., Gao, X., and Borriss, R. (2015). Biocontrol mechanism by root-associated Bacillus amyloliquefaciens FZB42–a review. *Front. Microbiol.* 6:780. doi: 10.3389/fmicb.2015.00780

Compant, S., Samad, A., Faist, H., and Sessitsch, A. (2019). A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *J. Adv. Res.* 19, 29–37. doi: 10.1016/j.jare.2019.03.004

DeAngelis, K. M., Brodie, E. L., DeSantis, T. Z., Andersen, G. L., Lindow, S. E., and Firestone, M. K. (2009). Selective progressive response of soil microbial community to wild oat roots. *ISME J.* 3, 168–178. doi: 10.1038/ismej.2008.103

Delaide, B., Delhaye, G., Dermience, M., Gott, J., Soyeurt, H., and Jijakli, M. H. (2017). Plant and fish production performance, nutrient mass balances, energy and water use of the PAFF Box, a small-scale aquaponic system. *Aquac. Eng.* 78, 130–139.

Delaide, B., Goddek, S., Gott, J., Soyeurt, H., and Jijakli, M. H. (2016). Lettuce (*Lactuca sativa* L. var. Sucrine) growth performance in complemented aquaponic solution outperforms hydroponics. *Water* 8:467.

Dolnicar, S., Grabler, K., and Mazanec, J. A. (2000). Analyzing destination images: a perceptual charting approach. *J. Travel Tour. Mark.* 8, 43–57.

Eck, M., Sare, A. R., Massart, S., Schmautz, Z., Junge, R., Smits, T. H. M., et al. (2019). Exploring bacterial communities in aquaponic systems. *Water* 11:260. doi: 10.1186/s12866-021-02273-4

Edmonds, J. W., Sackett, J. D., Lomprey, H., Hudson, H. L., and Moser, D. P. (2020). The aeroponic rhizosphere microbiome: community dynamics in early succession suggest strong selectional forces. *Antonie Van Leeuwenhoek* 113, 83–99. doi: 10.1007/s10482-019-01319-y

Ehret, D. L., Alsanius, B., Wohanka, W., Menzies, J. G., and Utkhede, R. (2001). Disinfestation of recirculating nutrient solutions in greenhouse horticulture. *Agronomie* 21, 323–339.

Elhady, A., Adss, S., Hallmann, J., and Heuer, H. (2018). Rhizosphere microbiomes modulated by pre-crops assisted plants in defense against plant-parasitic nematodes. *Front. Microbiol.* 9:1133. doi: 10.3389/fmicb.2018.01133

Garcia, J., and Kao-Kniffin, J. (2018). Microbial group dynamics in plant rhizospheres and their implications on nutrient cycling. *Front. Microbiol.* 9:1516. doi: 10.3389/fmicb.2018.01516

Geisen, S., Mitchell, E. A. D., Adl, S., Bonkowski, M., Dunthorn, M., Ekelund, F., et al. (2018). Soil protists: a fertile frontier in soil biology research. *FEMS Microbiol. Rev.* 42, 293–323. doi: 10.1093/femsre/fuy006

Goddek, S., Schmautz, Z., Scott, B., Delaide, B., Keesman, K. J., Wuertz, S., et al. (2016b). The effect of anaerobic and aerobic fish sludge supernatant on hydroponic lettuce. *Agronomy* 6:37.

Goddek, S., Espinal, C. A., Delaide, B., Jijakli, M. H., Schmautz, Z., Wuertz, S., et al. (2016a). Navigating towards decoupled aquaponic systems: a system dynamics design approach. *Water* 8:303.

Goddek, S., Joyce, A., Kotzen, B., and Burnell, G. M. (2019a). *Aquaponics Food Production Systems*. Berlin: Springer.

Goddek, S., Joyce, A., Wuertz, S., Körner, O., Bläser, I., Reuter, M., et al. (2019b). *Decoupled Aquaponics Systems. Aquaponics Food Production Systems*. Berlin: Springer, 201–229.

Goddek, S., and Keesman, K. J. (2018). The necessity of desalination technology for designing and sizing multi-loop aquaponics systems. *Desalination* 428, 76–85.

Goddek, S., and Korner, O. (2019). A fully integrated simulation model of multi-loop aquaponics: a case study for system sizing in different environments. *Agr. Syst.* 171, 143–154.

Goddek, S., and Vermeulen, T. (2018). Comparison of *Lactuca sativa* growth performance in conventional and RAS-based hydroponic systems. *Aquac. Int.* 26, 1377–1386. doi: 10.1007/s10499-018-0293-8

Gorsky, G., Bourdin, G., Lombard, F., Pedrotti, M. L., Audrain, S., Bin, N., et al. (2019). Expanding Tara oceans protocols for underway, ecosystemic sampling of the ocean-atmosphere interface during Tara Pacific expedition (2016–2018). *Front. Mar. Sci.* 6:750. doi: 10.3389/fmars.2019.00750

Gourion, B., Berrabah, F., Ratet, P., and Stacey, G. (2015). Rhizobium-legume symbioses: the crucial role of plant immunity. *Trends Plant Sci.* 20, 186–194. doi: 10.1016/j.tplants.2014.11.008

Guyonnet, J. P., Guillemet, M., Dubost, A., Simon, L., Ortet, P., Barakat, M., et al. (2018). Plant nutrient resource use strategies shape active rhizosphere microbiota through root exudation. *Front. Plant Sci.* 9:1662. doi: 10.3389/fpls.2018.01662

Hacquard, S. (2017). Commentary: microbial small talk: volatiles in fungal-bacterial interactions. *Front. Microbiol.* 8:1. doi: 10.3389/fmicb.2017.00001

Herren, C. M., and McMahon, K. D. (2018). Keystone taxa predict compositional change in microbial communities. *Environ. Microbiol.* 20, 2207–2217. doi: 10.1111/1462-2920.14257

Hu, L., Robert, C. A. M., Cadot, S., Zhang, X., Ye, M., Li, B., et al. (2018). Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nat. Commun.* 9:2738. doi: 10.1038/s41467-018-05122-7

Hussain, S. S., Mehnaz, S., and Siddique, K. H. (2018). "Harnessing the plant microbiome for improved abiotic stress tolerance," *Plant Microbiome: Stress Response*, eds D. Egamberdieva and P. Ahmad (Berlin: Springer) 21–43.

Jacoby, R. P., Chen, L., Schwier, M., Koprivova, A., and Kopriva, S. (2020). Recent advances in the role of plant metabolites in shaping the root microbiome. *F1000Res.* 9:F1000FacultyRev-151. doi: 10.12688/f1000research.21796.1

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., et al. (2017). Where less may be more: How the rare biosphere pulls ecosystems strings. *ISME J.* 11, 853–862. doi: 10.1038/ismej.2016.174

Kasozi, N., Abraham, B., Kaiser, H., and Wilhelmi, B. (2021). The complex microbiome in aquaponics: significance of the bacterial ecosystem. *Ann. Microbiol.* 71, 1–13.

Kawasaki, A., Donn, S., Ryan, P. R., Mathesius, U., Devilla, R., Jones, A., et al. (2016). Microbiome and exudates of the root and rhizosphere of brachypodium distachyon, a model for wheat. *PLoS One* 11:e0164533. doi: 10.1371/journal.pone.0164533

Kawasaki, A., Okada, S., Zhang, C., Delhaize, E., Mathesius, U., Richardson, A. E., et al. (2018). A sterile hydroponic system for characterising root exudates from specific root types and whole-root systems of large crop plants. *Plant Methods* 14:114. doi: 10.1186/s13007-018-0380-x

Kidoglu, F., Gül, A., Tüzel, Y., and Özaktan, H. (2008). "Yield enhancement of hydroponically grown tomatoes by rhizobacteria," in *Proceedings of the International Symposium on Strategies Towards Sustainability of Protected Cultivation in Mild Winter Climate*, Antalya.

Kristin, A., and Miranda, H. (2013). The root microbiota—a fingerprint in the soil? *Plant Soil* 370, 671–686.

Laureni, M., Weissbrodt, D. G., Szivak, I., Robin, O., Nielsen, J. L., Morgenroth, E., et al. (2015). Activity and growth of anammox biomass on aerobically pre-treated municipal wastewater. *Water Res.* 80, 325–336. doi: 10.1016/j.watres.2015.04.026

Liu, Y. W., and Huang, C. K. (2019). Effects of the circulation pump type and ultraviolet sterilization on nutrient solutions and plant growth in plant factories. *Horttechnology* 29, 189–198.

Lobanov, V. P., Combot, D., Pelissier, P., Labbé, L., and Joyce, A. (2021). Improving plant health through nutrient remineralization in aquaponic systems. *Front. Plant Sci.* 12:683690. doi: 10.3389/fpls.2021.683690

Lynch, J. M., and de Leij, F. (2001). *Rhizosphere, in ELS*.Atlanta, GA: American Cancer Society.

Lynch, M. D., and Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* 13, 217–229. doi: 10.1038/nrmicro3400

Micallef, S. A., Channer, S., Shiaris, M. P., and Colon-Carmona, A. (2009). Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. *Plant Signal Behav.* 4, 777–780. doi: 10.1093/jxb/erp053

Monsees, H., Suhl, J., Paul, M., Kloas, W., Dannehl, D., and Wurtz, S. (2019). Lettuce (*Lactuca sativa*, variety Salanova) production in decoupled aquaponic systems: same yield and similar quality as in conventional hydroponic systems but drastically reduced greenhouse gas emissions by saving inorganic fertilizer. *PLoS One* 14:e0218368. doi: 10.1371/journal.pone.0218368

Mori, J., and Smith, R. (2019). Transmission of waterborne fish and plant pathogens in aquaponics and their control with physical disinfection and filtration: a systematized review. *Aquaculture* 504, 380–395.

Munguia-Fragozo, P., Alatorre-Jacome, O., Rico-Garcia, E., Torres-Pacheco, I., Cruz-Hernandez, A., Ocampo-Velazquez, R. V., et al. (2015). Perspective for aquaponic systems:"omic" technologies for microbial community analysis. *BioMed Res. Int.* 2015:480386. doi: 10.1155/2015/480386

Nautiyal, C. S., Srivastava, S., Chauhan, P. S., Seem, K., Mishra, A., and Sopory, S. K. (2013). Plant growth-promoting bacteria Bacillus amyloliquefaciens NBRISN13 modulates gene expression profile of leaf and rhizosphere community in rice during salt stress. *Plant Physiol. Biochem.* 66, 1–9. doi: 10.1016/j.plaphy.2013.01.020

Nichols, M., and Savidov, N. (eds) (2011). "Aquaponics: a nutrient and water efficient production system," in *Proceedings of the II International Symposium on Soilless Culture and Hydroponics*, Mexico 947.

Palm, H. W., Knaus, U., Appelbaum, S., Strauch, S. M., and Kotzen, B. (2019). *Coupled aquaponics systems. Aquaponics Food Production Systems*. Berlin: Springer, 163–199.

Phillips, D. A., Fox, T. C., King, M. D., Bhuvaneswari, T. V., and Teuber, L. R. (2004). Microbial products trigger amino acid exudation from plant roots. *Plant Physiol.* 136, 2887–2894. doi: 10.1104/pp.104.044222

Raudales, R. E., Parke, J. L., Guy, C. L., and Fisher, P. R. (2014). Control of waterborne microbes in irrigation: a review. *Agr. Water Manage.* 143, 9–28.

Rurangwa, E., and Verdegem, M. C. J. (2015). Microorganisms in recirculating aquaculture systems and their management. *Rev. Aquacult.* 7, 117–130.

Sasse, J., Martinoia, E., and Northen, T. (2018). Feed your friends: Do plant exudates shape the root microbiome? *Trends Plant Sci.* 23, 25–41. doi: 10.1016/j.tplants.2017.09.003

Savidov, N., and Brooks, A. (2004). *Evaluation and Development of Aquaponics Production and Product Market Capabilities in Alberta*. Toronto: Crop Diversification Centre South.

Scagliola, M., Pii, Y., Mimmo, T., Cesco, S., Ricciuti, P., and Crecchio, C. (2016). Characterization of plant growth promoting traits of bacterial isolates from the rhizosphere of barley (*Hordeum vulgare* L.) and tomato (*Solanum lycopersicon*

L.) grown under Fe sufficiency and deficiency. *Plant Physiol. Biochem.* 107, 187–196. doi: 10.1016/j.plaphy.2016.06.002

Schmautz, Z., Graber, A., Jaenicke, S., Goesmann, A., Junge, R., and Smits, T. H. (2017). Microbial diversity in different compartments of an aquaponics system. *Arch. Microbiol.* 199, 613–620. doi: 10.1007/s00203-016-1334-1

Schreiter, S., Ding, G. C., Heuer, H., Neumann, G., Sandmann, M., Grosch, R., et al. (2014). Effect of the soil type on the microbiome in the rhizosphere of field-grown lettuce. *Front. Microbiol.* 5:144. doi: 10.3389/fmicb.2014.00144

Seyler, L. M., McGuinness, L. M., and Kerkhof, L. J. (2014). Crenarchaeal heterotrophy in salt marsh sediments. *ISME J.* 8, 1534–1543. doi: 10.1038/ismej.2014.15

Sheridan, C., Depuydt, P., De Ro, M., Petit, C., Van Gysegem, E., Delaere, P., et al. (2017). Microbial community dynamics and response to plant growth-promoting microorganisms in the rhizosphere of four common food crops cultivated in hydroponics. *Microb. Ecol.* 73, 378–393. doi: 10.1007/s00248-016-0855-0

Shimizu, K., Matsuda, Y., Nonomura, T., Ikeda, H., Tamura, N., Kusakari, S., et al. (2007). Dual protection of hydroponic tomatoes from rhizosphere pathogens Ralstonia solanacearum and Fusarium oxysporum f.sp radicis-lycopersici and airborne conidia of Oidium neolycopersici with an ozone-generative electrostatic spore precipitator. *Plant Pathol.* 56, 987–997.

Suez, J., Zmora, N., Zilberman-Schapira, G., Mor, U., Dori-Bachash, M., Bashiardes, S., et al. (2018). Post-antibiotic gut mucosal microbiome reconstitution is impaired by probiotics and improved by autologous FMT. *Cell* 174, 1406.–1423. doi: 10.1016/j.cell.2018.08.047

Sunagawa, S., Acinas, S. G., Bork, P., Bowler, C., Eveillard, D., Gorsky, G., et al. (2020). Tara Oceans: towards global ocean ecosystems biology. *Nat. Rev. Microbiol.* 18, 428–445. doi: 10.1038/s41579-020-0364-5

Topalovic, O., Hussain, M., and Heuer, H. (2020). Plants and Associated Soil Microbiota Cooperatively Suppress Plant-Parasitic Nematodes. *Front. Microbiol.* 11:313. doi: 10.3389/fmicb.2020.00313

Vadstein, O., Attramadal, K. J. K., Bakke, I., and Olsen, Y. (2018). K - selection as microbial community management strategy: a method for improved viability of larvae in aquaculture. *Front. Microbiol.* 9:2730. doi: 10.3389/fmicb.2018.02730

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., and Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *New Phytol.* 206, 1196–1206. doi: 10.1111/nph.13312

Wang, Z., Li, T., Wen, X., Liu, Y., Han, J., Liao, Y., et al. (2017). Fungal communities in rhizosphere soil under conservation tillage shift in response to plant growth. *Front. Microbiol.* 8:1301. doi: 10.3389/fmicb.2017.01301

Wielgosz, Z. J., Anderson, T. S., and Timmons, M. B. (2017). Microbial effects on the production of aquaponically grown lettuce. *Horticulturae* 3:46. doi: 10.3390/foods2040488

Yang, C. H., and Crowley, D. E. (2000). Rhizosphere microbial community structure in relation to root location and plant iron nutritional status. *Appl. Environ. Microbiol.* 66, 345–351. doi: 10.1128/AEM.66.1.345-351.2000

Yasin, N. A., and Ahmed, S. (2016). Induction of defence-related biochemicals by rhizosphere bacteria against black spot disease of rose. *Biol. Agric. Hortic.* 32, 34–46.

Yeoh, Y. K., Dennis, P. G., Paungfoo-Lonhienne, C., Weber, L., Brackin, R., Ragan, M. A., et al. (2017). Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. *Nat. Commun.* 8:215. doi: 10.1038/s41467-017-00262-8

Yep, B., and Zheng, Y. B. (2019). Aquaponic trends and challenges - a review. *J. Clean. Prod.* 228, 1586–1599. doi: 10.1016/j.tifs.2020.06.013

Zhalnina, K., Louie, K. B., Hao, Z., Mansoori, N., da Rocha, U. N., Shi, S., et al. (2018). Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat. Microbiol.* 3, 470–480. doi: 10.1038/s41564-018-0129-3

Zhang, L., Narita, Y., Gao, L., Ali, M., Oshiki, M., and Okabe, S. (2017). Maximum specific growth rate of anammox bacteria revisited. *Water Res.* 116, 296–303. doi: 10.1016/j.watres.2017.03.027

Zheng, L., Yang, Q., and Song, W. (2019). Ozonated nutrient solution treatment as an alternative method for the control of root-knot nematodes in soilless cultivation. *Ozone Sci. Eng.* 42, 371–376.

Zielińska, M., Cydzik-Kwiatkowska, A., Bernat, K., Zieliński, M., Kulikowska, D., and Wojnowska-Baryła, I. (2018). Start-up of a one-stage biofilm reactor for the removal of nitrogen from digester supernatant in the partial nitrification-anammox process. *Rocznik Ochrona środowiska* 20, 241–257.

Check for updates

Community Interaction Co-limitation: Nutrient Limitation in a Marine Microbial Community Context

Catherine Bannon[1†], Insa Rapp[1,2†] and Erin M. Bertrand[1*]

[1] Department of Biology and Institute for Comparative Genomics, Dalhousie University, Halifax, NS, Canada, [2] Marine Biogeochemistry Division, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

The simultaneous limitation of productivity by two or more nutrients, commonly referred to as nutrient co-limitation, affects microbial communities throughout the marine environment and is of profound importance because of its impacts on various biogeochemical cycles. Multiple types of co-limitation have been described, enabling distinctions based on the hypothesized mechanisms of co-limitation at a biochemical level. These definitions usually pertain to individuals and do not explicitly, or even implicitly, consider complex ecological dynamics found within a microbial community. However, limiting and co-limiting nutrients can be produced in situ by a subset of microbial community members, suggesting that interactions within communities can underpin co-limitation. To address this, we propose a new category of nutrient co-limitation, community interaction co-limitation (CIC). During CIC, one part of the community is limited by one nutrient, which results in the insufficient production or transformation of a biologically produced nutrient that is required by another part of the community, often primary producers. Using cobalamin (vitamin B_{12}) and nitrogen fixation as our models, we outline three different ways CIC can arise based on current literature and discuss CIC's role in biogeochemical cycles. Accounting for the inherent and complex roles microbial community interactions play in generating this type of co-limitation requires an expanded toolset – beyond the traditional approaches used to identify and study other types of co-limitation. We propose incorporating processes and theories well-known in microbial ecology and evolution to provide meaningful insight into the controls of community-based feedback loops and mechanisms that give rise to CIC in the environment. Finally, we highlight the data gaps that limit our understanding of CIC mechanisms and suggest methods to overcome these and further identify causes and consequences of CIC. By providing this framework for understanding and identifying CIC, we enable systematic examination of the impacts this co-limitation can have on current and future marine biogeochemical processes.

Keywords: nutrient co-limitation, microbial community interactions, cobalamin, nitrogen fixation, marine microbes, primary producers

INTRODUCTION

Major nutrients (N and P) and micronutrients (Fe, Co, Zn, Mn, B_1, B_{12}) have been found to directly limit phytoplankton growth alone or in combination (Bertrand et al., 2007; Moore et al., 2013; Paerl et al., 2017; Wu et al., 2019; Browning et al., 2021), thereby impacting marine food webs and ocean productivity. Nutrient stoichiometries and molecular measurements in the global ocean suggest that primary production in large parts of the surface ocean may be subject to co-limitation and stress (defined as a physiological response to nutrient scarcity, Moore et al., 2013) due to simultaneous deficiencies in multiple nutrients (Moore et al., 2013; Saito et al., 2015; Browning et al., 2017). Co-limitation has been suggested to be a common condition in heterogenous microbial communities because of different nutrient demands between microbial groups and taxa-specific abilities to adapt to low nutrient conditions (Danger et al., 2008). Cellular stoichiometry for micronutrients, such as trace metals, can vary by more than two orders of magnitude, with variability between taxa as well as between growth conditions and nutrient gradients in the ocean (Moore et al., 2013; Twining and Baines, 2013; Twining et al., 2021). Cellular stoichiometries for macronutrients show less variation but deviate significantly from the Redfield ratio for individual species (Ho et al., 2003; Garcia et al., 2018). Additionally, certain members of a phytoplankton community may have different abilities to access or preferences for nutrient pools, such as ligand-bound Fe (Hutchins et al., 1999) and forms of nitrogen [e.g., N_2, NO_x, NH_4 or urea (dissolved organic nitrogen: DON)] (Moore et al., 2002; Collier et al., 2009, 2012). Comparatively little is known about the requirements for B-vitamins, though it is clear that eukaryotic phytoplankton and the majority of cyanobacteria require different forms of cobalamin (Helliwell et al., 2016; Heal et al., 2017), and that different species can satisfy their thiamine requirements using different thiamine-related compounds (Paerl et al., 2017, 2018a,b).

Co-limitation is commonly divided into 3 types: (I) independent co-limitation, (II) biochemical substitution co-limitation and (III) biochemically dependent co-limitation (Saito et al., 2008). Independent co-limitation occurs when two entirely independent nutrients are simultaneously drawn down to limiting conditions either for a single organism or, in a community context, different parts of the community are limited by different nutrients. This type can either occur as true co-limitation or as serial limitation, where the secondary limiting nutrient only becomes limiting when the limitation by the primary limiting nutrient has been relieved. Biochemical substitution co-limitation occurs when one limiting nutrient can be substituted for the other limiting nutrient or the addition of one nutrient can partly relieve the nutrient stress imposed by the lack of another limiting nutrient and vice versa (Saito et al., 2008). Biochemically dependent co-limitation occurs when one nutrient is required for acquisition of another nutrient at low concentrations (Saito et al., 2008; Moore et al., 2013). These definitions describe scenarios where two nutrients are simultaneously limiting phytoplankton growth but can be further extended to three or more nutrients. These different categories

of nutrient limitation have been subject to many studies and reviews to date (e.g., see Saito et al., 2008; Harpole et al., 2011; Moore et al., 2013).

Currently, the mechanisms describing these types of co-limitation do not consider complex interactions that arise when a limiting nutrient is produced *in situ* by a subset of the microbial community. Such conditions have been identified in the ocean (Bertrand et al., 2015) and have the potential to respond to change much differently than other types of co-limitation. We suggest this necessitates an additional definition that acknowledges the interdependencies and exchange of essential materials in microbial communities. To this end, we introduce the concept of community interaction co-limitation (CIC). CIC describes a community where multiple limiting nutrient cycles are affected by one another through the interactions between groups within the community. In other words, one part of the community is limited by one nutrient, which results in the insufficient production or transformation of a biologically produced nutrient (BPN) that is required by another part of the community, often primary producers. This results in a community that is limited by multiple nutrients because of the interactions among different microbial groups present. Unlike other categories of nutrient co-limitation, community interaction co-limitation cannot be observed in a monoculture and is not relevant in the context of an individual. Some examples of CIC are given in **Table 1** and include limitation of bacterial vitamin production by a specific nutrient, which results in the limitation of the phytoplankton community that requires this vitamin but lacks the ability to synthesize it. Another example is the limitation of diazotrophs by the availability of iron and/or phosphorus in a nitrogen depleted region, resulting in nitrogen limitation of the non-diazotrophic community.

Here we aim to describe the importance of considering community dynamics in the interpretation of nutrient limitation in the ocean. Some of the examples provided above (**Table 1**) have previously been classified as one of the three traditional types of co-limitation, e.g., iron and B_{12} co-limitation as type I, and iron limitation of nitrogen fixation as type III (Saito et al., 2008). These traditional classifications are of most utility when focusing on one species or dominant plankton group at a time. We are suggesting that these classifications are not invalid but, instead, are incomplete. Marine microbial communities are webs of interactions which control the communities' structure, function, and resource availability – factors that influence the community's productivity, composition, and contribution to global-scale processes.

Community interactions are commonly overlooked during experiments investigating resource limitation, perhaps because of the complexity required to sample and interpret them. Yet, community interactions are commonly overlooked during experiments investigating resource limitation, perhaps because of the complexity required to sample and interpret them. Yet, community interactions control nutrient availability and metabolic status of individual community members through the exchange of nutrients and metabolites via direct, associated relationships (in the phycosphere) or through cross feeding of public goods in bulk medium (Seymour et al., 2017;

Examples for CIC	Biologically produced nutrient	Producers	Nutrient limiting production of biologically produced nutrient	Consumers of biologically produced nutrient	Select references
Limited availability of one or more nutrients limits B_{12} production by bacteria and archaea, resulting in phytoplankton growth being limited by B_{12}	B_{12}	Select bacteria and archaea	Iron, nitrogen, labile carbon, cobalt	Eukaryotic phytoplankton	Bertrand et al., 2007, 2015; Gobler et al., 2007; Koch et al., 2011; Barber-Lluch et al., 2019
Limitation of nitrogen fixers (e.g., by P or Fe) resulting in limitation of non-diazotrophic phytoplankton by N	Fixed nitrogen (e.g., NO_x, NH_4)	Diazotrophs	Iron, phosphate, labile carbon	Non-diazotrophic primary producers	Mills et al., 2004
Limitation of B_1 production resulting in co-limitation of B_1 auxotrophs by B_1 and N/C	B_1	Select bacteria, archaea, and eukaryotic phytoplankton	Nitrogen, carbon	B_1-auxotrophic phytoplankton (e.g., chlorophyta) and select bacteria	Gobler et al., 2007; Koch et al., 2012; Paerl et al., 2018b

Pacheco et al., 2019). A central focus of microbial ecology is to identify mechanisms through which communities assemble in order to predict potential for change and how that change could influence community function (e.g., Zomorrodi and Segrè, 2017; Fu et al., 2020). Incorporating ideas about microbial evolution, such as the Black Queen Hypothesis (discussed in Section "Incorporating Ecological and Evolutionary Theory Into Community Interaction Co-limitation"), aids in identifying the principles that govern the evolution of microbial dependencies and "public good" nutrient production, which could underpin such instances of this co-limitation (Mas et al., 2016). Incorporating ecological and evolution theories into resource limitation studies will enable stronger interpretations of community-based feedback loops that control resource availability, which could provide a framework to predict the response of microbial communities (e.g., Mas et al., 2016; Zomorrodi and Segrè, 2017; Fu et al., 2020). Considering CIC as a fourth and separate type of co-limitation will enable researchers to systematically leverage findings from microbial ecology and evolutionary ecology, highlighting the impact microbial interactions have on larger biogeochemical cycles and generating new insights on patterns and trends of nutrient limitation in the ocean.

COMMUNITY INTERACTION CO-LIMITATION

Cobalamin

While community interaction co-limitation can extend to other BPNs (**Table 1**), we use the micronutrient cobalamin (vitamin B_{12}) to provide evidence for CIC because of its importance in microbial community interactions and potential to co-limit primary production in open ocean and coastal waters (Croft et al., 2005; Bertrand et al., 2007; Ramanan et al., 2016; Sokolovskaya et al., 2020). Cobalamin is a cobalt-containing organometallic micronutrient that's structure was first elucidated in 1956 by

Nobel laureate Dr. Dorothy Crowfoot Hodgkin (Hodgkin et al., 1956). It is a powerful cofactor that performs rearrangements and methylation during biochemical reactions (Banerjee and Ragsdale, 2003; Dowling et al., 2012). Cobalamin can only be synthesized *de novo* by select bacteria and archaea, and it is required by auxotrophic bacteria and an estimated 50% of eukaryotic phytoplankton (Martens et al., 2002; Croft et al., 2005). Auxotrophy in eukaryotic phytoplankton is believed to arise from cobalamin's role as a cofactor for the enzyme methionine synthase (MetH), which is the final step in methionine synthesis (Croft et al., 2005; Helliwell et al., 2011). All eukaryotic phytoplankton appear to have a cobalamin dependent MetH, while a subset have an additional cobalamin independent one (MetE) that is less efficient and more costly to use (Bertrand et al., 2013). Cyanobacteria produce and use their own cobalamin-like molecule, pseudo-cobalamin, that is largely unavailable to eukaryotic algae but can be re-modeled into cobalamin by a subset of heterotrophic bacteria and a small group of algae, if the alpha ligand (DMB) is available (Helliwell et al., 2016; Ma et al., 2020). This discovery improved, and complicated, our understanding of the marine cobalamin cycle, which now must consider pseudo-cobalamin and remodelers as a potential source of cobalamin. Cobalamin has a significant history at the intersection between biological oceanography and microbial ecology, which has been reviewed (Sañudo-Wilhelmy et al., 2014; Helliwell, 2017) but we will briefly outline its role in algal–bacterial interactions and as a limiting micronutrient here because of its relevance to community interaction co-limitation.

Cobalamin has an important role in microbial interactions and is regularly considered as a model metabolite exchanged in beneficial microbial interactions (Ramanan et al., 2016). Auxotrophic phytoplankton can acquire cobalamin through sustaining obligate or facultative interactions with B_{12}-producing bacteria and/or archaea or by uptake of the vitamin from its surrounding environment, referred to as scavenging (Amin et al., 2012; Kazamia et al., 2012; Bertrand et al., 2015). However, the strategy an organism uses to obtain sufficient cobalamin is likely

affected by the availability of the nutrient or associated producer. The factors that influence methods of cobalamin acquisition and the regulation of cobalamin transfer or exchange between organisms are still poorly understood although relevant for understanding CIC. Interactions between cobalamin producers and consumers can be selected for and strengthened, in part, by the reciprocal transfer of nutrients and/or the active selection of beneficial bacteria through secondary metabolites (Kazamia et al., 2012; Grant et al., 2014; Shibl et al., 2020). In particular, the exchange of organic, labile carbon for cobalamin is relatively well documented, and it is expected that other nutrients, such as nitrogen species, amino acids, sulfonates, sugar derivatives or specific growth factors like indol-3-acetic acid could be exchanged as well (Amin et al., 2012; Ramanan et al., 2016; Durham et al., 2019).

Cobalamin is found at picomolar to sub-picomolar concentrations in the open ocean (Sañudo-Wilhelmy et al., 2014) and has already been shown to limit or co-limit phytoplankton communities in multiple regions of the ocean (e.g., coastal McMurdo Sound and the Ross Sea, Gulf of Alaska, coastal NE Atlantic) (Bertrand et al., 2007, 2015; Koch et al., 2011; Barber-Lluch et al., 2019). These studies investigated cobalamin limitation using bottle incubation nutrient addition experiments, occasionally layering omics approaches to investigate mechanisms of limitation (Bertrand et al., 2015). A few studies have directly measured the concentration of cobalamin(s) in the water column (Gobler et al., 2007; Heal et al., 2014). Determining potential for cobalamin limitation through examining stoichiometry is not yet feasible since cobalamin production rates, requirements, and quotas are still poorly defined. Cobalamin availability is thought to depend on the rate of photodegradation, remodeling, and balance of supply and demand (driven by B_{12} producers and B_{12} consumers). But, cobalamin production can be limited by other nutrients required for bacterial growth (iron or nitrogen) or cobalamin production (e.g., cobalt).

Nitrogen Fixation

It is also useful to consider the more extensively studied example of nitrogen fixation in the context of CIC. Since a number of reviews summarizing the current advances in understanding marine nitrogen fixation are available (e.g., Sohm et al., 2011; Zehr and Capone, 2020), we will only briefly highlight some key findings on community interactions as it pertains to CIC. Nitrogen fixation is performed by diazotrophs and comprises an important source of fixed nitrogen for non-diazotrophs in the ocean, especially during periods of low fixed nitrogen availability (Karl et al., 2002). However, diazotrophic activity, and therefore nitrogen supply, can be limited by iron and phosphorous availability, generating a community that is co-limited by iron and/or phosphorous and nitrogen (Held et al., 2020). In many cases, diazotrophs, such as *Trichodesmium* spp., independently fix atmospheric carbon but it remains unclear if the diazotroph benefits in any way from the release/exchange of fixed nitrogen. Some diazotrophic cyanobacteria observed to be associated with haptophytes appear to lack the ability to fix atmospheric carbon, thereby relying on the supply of organic carbon from
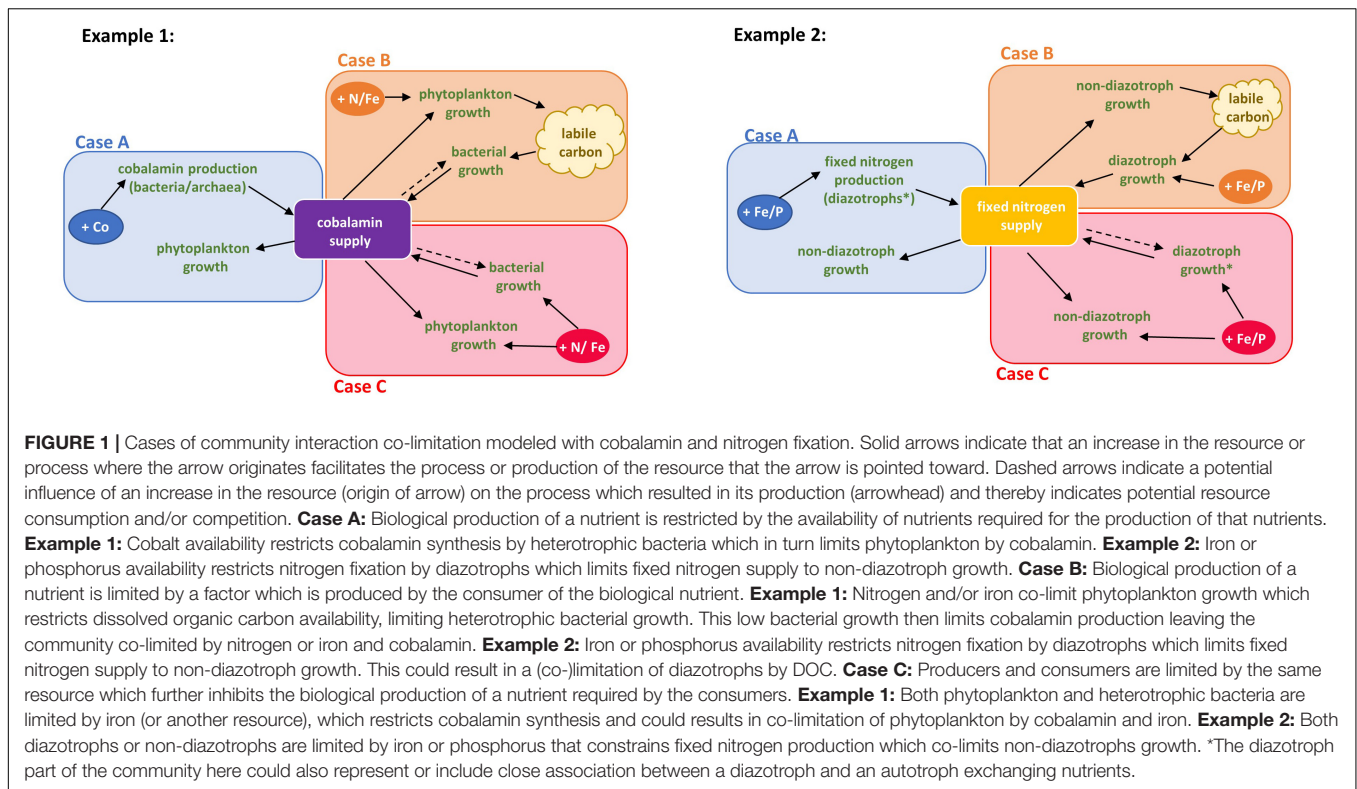
the haptophytes (Thompson et al., 2012). Carbon and nitrogen exchange between the two symbionts therefore relies on adequate growth conditions for both organisms. Several recent studies also suggest a direct influence of DOM availability on nitrogen fixation rates when (non-cyanobacterial) heterotrophs dominate the diazotrophic community (Bonnet et al., 2013; Rahav et al., 2015, 2016). The nutrients produced within these symbiosis could be available to the surrounding community through active or passive transfer and, as such, suggest that nitrogen fixation is well suited to be investigated through the lens of CIC.

CASES OF COMMUNITY INTERACTION CO-LIMITATION

The production of a biological nutrient *in situ* can be restricted in a variety of ways, each of which have the potential to limit the growth of other community members (like phytoplankton) that require the BPN (**Figure 1**). Here we outline different cases of CIC based on the mechanisms that limit the availability of the BPN. The underpinning mechanism of CIC may determine how efficiently a microbial community can respond to changes in nutrient supply or recover from nutrient stress once the availability of the limiting nutrient increases. Categorizing CIC into specific cases clarifies mechanisms of limitation, potential metabolic interactions, and potential molecular responses for each CIC case. This will improve our ability to systematically predict the impact of CIC and allow development of targeted strategies to monitor spatiotemporal patterns in CIC.

Case A describes a scenario where the biological production of a nutrient (the BPN) is restricted by the availability of another nutrient required for BPN production. This results in the limitation of one part of the community by the BPN because the production is limited by another nutrient. An increase in the nutrient limiting the biological production likely increases the supply of the BPN immediately, or with a short time-lag. In Case B, the biological production of the nutrient is limited by a factor that is produced by the consumer of the BPN, which is restricted because of the limited availability of the other co-limiting nutrient. An increased supply of the additional co-limiting nutrient might not directly relieve the limitation by the BPN as the production remains limited by the factor produced by the consumer. Finally, in Case C, producers and consumers are limited by the same resource, which inhibits the production of a separate BPN that limits the consumer growth (**Figure 1**). This case would behave like Case A when the supply of the limiting nutrient is increasing, but consumers and producers may compete for the same resource.

We use examples, supported by literature, to describe the three cases of CIC in **Figure 1**. In Case A of cobalamin CIC, cobalamin biosynthesis by the producers is restricted by cobalt availability. Case B describes a community that is already limited by one or more nutrients, such as iron and/or nitrogen, which results in insufficient production of dissolved organic carbon. This restricts bacterial growth and cobalamin synthesis, which results in restricted cobalamin supply to phytoplankton. Case C represents a scenario where both phytoplankton and cobalamin

FIGURE 1 | Cases of community interaction co-limitation modeled with cobalamin and nitrogen fixation. Solid arrows indicate that an increase in the resource or process where the arrow originates facilitates the process or production of the resource that the arrow is pointed toward. Dashed arrows indicate a potential influence of an increase in the resource (origin of arrow) on the process which resulted in its production (arrowhead) and thereby indicates potential resource consumption and/or competition. **Case A:** Biological production of a nutrient is restricted by the availability of nutrients required for the production of that nutrients. **Example 1:** Cobalt availability restricts cobalamin synthesis by heterotrophic bacteria which in turn limits phytoplankton by cobalamin. **Example 2:** Iron or phosphorus availability restricts nitrogen fixation by diazotrophs which limits fixed nitrogen supply to non-diazotroph growth. **Case B:** Biological production of a nutrient is limited by a factor which is produced by the consumer of the biological nutrient. **Example 1:** Nitrogen and/or iron co-limit phytoplankton growth which restricts dissolved organic carbon availability, limiting heterotrophic bacterial growth. This low bacterial growth then limits cobalamin production leaving the community co-limited by nitrogen or iron and cobalamin. **Example 2:** Iron or phosphorus availability restricts nitrogen fixation by diazotrophs which limits fixed nitrogen supply to non-diazotroph growth. This could result in a (co-)limitation of diazotrophs by DOC. **Case C:** Producers and consumers are limited by the same resource which further inhibits the biological production of a nutrient required by the consumers. **Example 1:** Both phytoplankton and heterotrophic bacteria are limited by iron (or another resource), which restricts cobalamin synthesis and could results in co-limitation of phytoplankton by cobalamin and iron. **Example 2:** Both diazotrophs or non-diazotrophs are limited by iron or phosphorus that constrains fixed nitrogen production which co-limits non-diazotrophs growth. *The diazotroph part of the community here could also represent or include close association between a diazotroph and an autotroph exchanging nutrients.

producers are limited by the same nutrient (e.g., iron), thereby restricting cobalamin synthesis and consequently leaving the phytoplankton community co-limited by cobalamin and iron.

Nitrogen fixation CIC can also be classified into the proposed cases (**Figure 1**). Case A describes a community in which the diazotroph is limited by iron or phosphorous, resulting in nitrogen limitation of the non-diazotrophic community, rendering the overall community co-limited by nitrogen and iron or phosphorous. In Case B, nitrogen fixation by diazotrophs is limited by biologically produced compounds, resulting in nitrogen limitation of the phytoplankton community. The cause of such a scenario could be preceding iron or phosphorous limitation of diazotrophs, restricting nitrogen supply to the community and thereby DOM availability for the nitrogen fixers. After an increase in iron or phosphorous supply, the diazotrophs remain limited by DOM, impeding a fast recovery from the nutrient limited state. Case C describes a community in which both diazotrophs and non-diazotrophs are limited by the same nutrient, i.e., iron, resulting in co-limitation of the non-diazotrophs by nitrogen and iron. The close association between haptophytes and diazotrophic cyanobacteria may be described by Case B when only considering the two symbionts. However, looking at this from the community perspective which the CIC definition aims to achieve, one might consider the close association of haptophyte/diazotroph as one part of the community which supplies fixed nitrogen to other members of the community through grazing, remineralization, or passive release. If we assume that the haptophyte/diazotroph association is independent of DOM supply from other members of the

community, the effect for the entire community would then classify as Case A or Case C depending on the mechanism of co-limiting nutrient (**Figure 1**).

We expect that other cases of CIC might become evident with additional research into auxotrophy and essential micronutrients, and that communities may experience a continuum, or a mix of more than one case occurring in the same community at any given time. Furthermore, there may also be links between CIC cases involving more than one BPN at the same time, for example, some cyanobacterial diazotrophs (UCYN-A) closely associated with haptophytes, are able to produce pseudocobalamin and might therefore be able to supply fixed nitrogen and cobalamin (after conversion) to their host (Koch et al., 2011; Muñoz-Marín et al., 2018; Barber-Lluch et al., 2019). Lesser explored community interactions (and therefore not included in **Table 1**) which could lead to CIC, might include exoenzymes which liberate inorganic nutrients for organisms other than those producing the enzymes (e.g., alkaline phosphatase; Luo et al., 2009). Phytoplankton might benefit from accessing these inorganic nutrients, rather than producing the enzymes required for the conversion themselves. Additionally, production of siderophores by select community members could lead to an increase in 'new' bioavailable iron by mobilizing particulate iron (Manck et al., 2022). The nutrient limiting the growth of siderophore producers could then possibly limit bioavailable iron delivery to the whole community, contributing to CIC. The details matter here though, as this siderophore-bound iron is not available to all community members equally (Hutchins et al., 1999; Sanchez et al., 2018; Sutak et al., 2020). Identifying microbial groups that rely on interactions

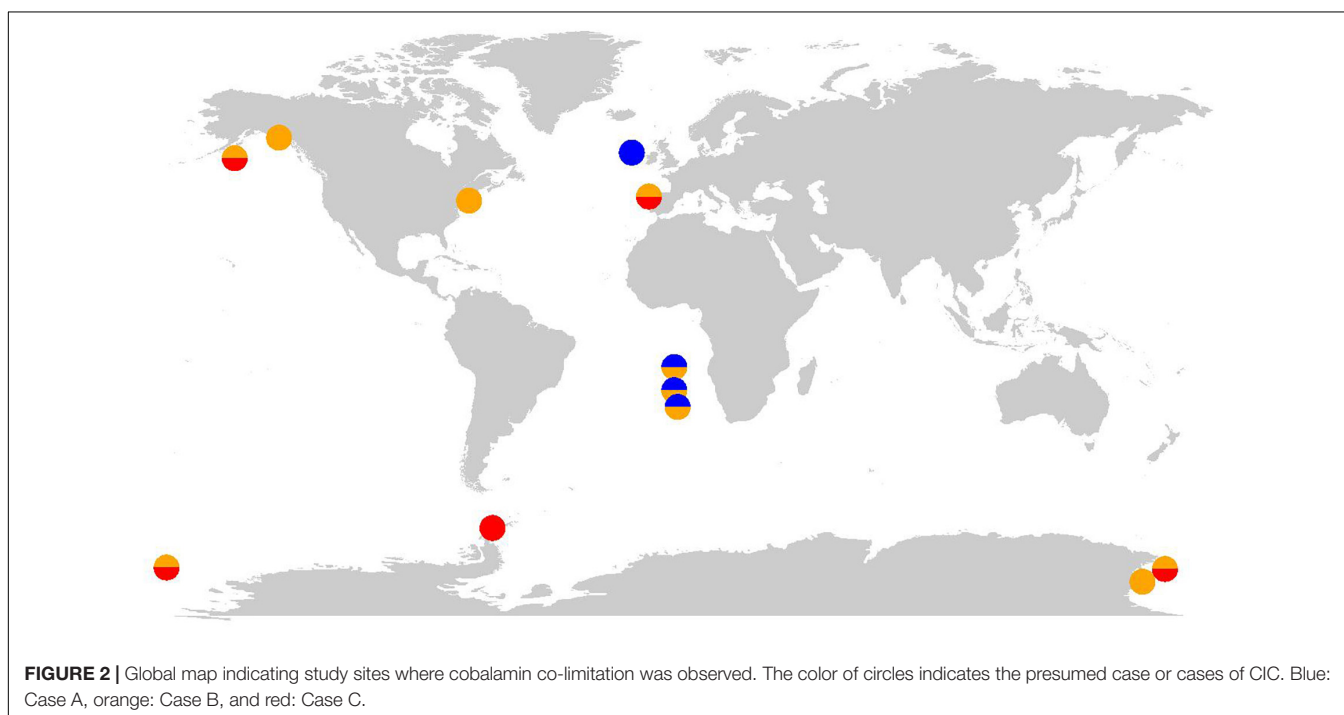to obtain BPNs will be important for investigations of additional mechanisms of CIC.

EVIDENCE OF COMMUNITY INTERACTION CO-LIMITATION IN THE LITERATURE

In the following, we describe nine studies that provide evidence that the *in situ* microbial community was experiencing at least one case of cobalamin CIC (**Figure 2**). Using the available data from these studies, we suggest the most likely case or cases of CIC (A-C) and highlight open questions which must be addressed to fully confirm the case type (**Table 2**). Generally, we suggest that the identification of CIC and underpinning mechanism requires demonstrating that the production of the BPN is limited by another resource and co-limiting phytoplankton growth. Traditionally, evidence for co-limitation has been gathered using ship-board bottle incubation bioassay experiments and evaluated based on phytoplankton growth responses to nutrient additions assessed by measuring chlorophyll *a*, photosynthetic efficiency (F_v/F_m), particulate organic carbon (POC) production, nutrient drawdown, and cell counts, often also assessing changes in community composition to some extent. However, bottle incubation bioassays alone are insufficient to identify the specific case of CIC (A-C) and describe any potential feedback loops of nutrient transfer in the community. Such identifications remain rare in the literature but may be accomplished using omics or other approaches that can resolve the response of different community members to nutrient manipulations (refer to section "Data gaps, uncertainties, and future directions" for expanded discussion).

Studies listed in **Table 2** revealed increases in phytoplankton growth after the addition of cobalamin alone or in combination with other nutrients. Although the magnitude of responses to cobalamin is often lower than for other limiting nutrients, some studies have found B_{12} additions alone cause a very similar or even larger response in chlorophyll *a* concentration compared to the other co-limiting nutrients (Koch et al., 2011; Barber-Lluch et al., 2019), or observed a doubling in chlorophyll *a* after combined addition of B_{12} with nitrogen or iron compared to adding nitrogen or iron alone (Bertrand et al., 2007; Gobler et al., 2007). Additionally, these studies demonstrated a strong shift in community composition with the addition of cobalamin and revealed seasonal variability in cobalamin co-limitation. Most of the studies do not provide enough data to definitively determine the specific case of CIC. However, a good first indicator of what might be limiting cobalamin production is the response of heterotrophic bacterial growth to nutrient addition. This indicator, however, needs to be treated with caution as only a subset of bacteria are cobalamin producers and a larger fraction appears to be cobalamin consumers (Soto et al., in review[1]; Shelton et al., 2019).

Some studies included in **Table 2** have additional limitations in terms of their ability to conclusively identify cases of CIC. For example, most studies did not quantify heterotrophic bacterial growth or did not assess the response to organic carbon addition, making it difficult to distinguish between Case B and C (Sañudo-Wilhelmy et al., 2006; Bertrand et al., 2007; Gobler et al., 2007; Koch et al., 2011). In the HNLC region in the

---

[1]Soto, M. A., Desai, D, LaRoche, J, and Bertrand, E. M. (in review). Cobalamin producers and prokaryotic consumers in the Northwest Atlantic. *Environ. Microbiol.*

FIGURE 2 | Global map indicating study sites where cobalamin co-limitation was observed. The color of circles indicates the presumed case or cases of CIC. Blue: Case A, orange: Case B, and red: Case C.

TABLE 2 | Instances of cobalamin community interaction co-limitation from literature.

Study site	Case of CIC	Evidence of co-limitation	Notes	Open questions	References
Sea ice edge of McMurdo sound	B	Addition of B_{12} and Fe led to increase in Chl a production and downregulation of iron and cobalamin deprivation indicators in phytoplankton and bacteria. Oceanospirillaceae ASP10-02A expressed organic compound acquisition genes and B_{12} production genes (see text).	No attempt to measure heterotrophic bacterial growth after nutrient additions.	Confirm limitation of B_{12} production by organic carbon availability?	Bertrand et al., 2015
Southern Ocean (Antarctic Peninsula)	C	Addition of B_{12}, B_1 and Fe together significantly enhanced phytoplankton growth. Addition of iron alone significantly increase bacterial production.	Addition of B-vitamins alone stimulated phytoplankton growth.	Were the bacteria stimulated by iron predominantly B_{12} consumers or producers?	Panzeca et al., 2006
Gulf of Alaska (coastal region)	B	$NO_3 + B_{12}$ addition resulted in the highest increase in $> 2 \, \mu M$ phytoplankton and heterotrophic bacterial growth.	Addition of cobalamin alone did not increase phytoplankton growth but caused a shift in community composition. Addition of N alone did not increase heterotrophic bacterial growth. Most B_{12} uptake was likely by heterotrophic bacteria.	Were the stimulated bacteria B_{12} consumers or producers?	Koch et al., 2011
Gulf of Alaska (HNLC region)	B or C	Addition of B_{12} and Fe together significantly enhanced phytoplankton and heterotrophic bacteria growth.	Addition of B_{12} or Fe alone stimulated phytoplankton and bacterial growth. Most B_{12} uptake was likely by heterotrophic bacteria.	Were the stimulated bacteria B_{12} consumers or producers? Were the bacteria co-limited by iron and B_{12} or by production of organic matter?	Koch et al., 2011
Long Island embayment	B or C	Addition of B_{12} and NO_3 together significantly enhanced phytoplankton growth.	No attempt to measure heterotrophic bacterial growth after nutrient additions. Addition of B_{12} alone stimulated $> 5 \, \mu M$ phytoplankton growth.	B_{12} production limited by organic carbon or nitrogen availability?	Sañudo-Wilhelmy et al., 2006; Gobler et al., 2007
Ross Sea	B or C	Fe + B_{12} amendment experiment resulted in an increase in the phytoplankton growth.	Addition of B_{12} alone did not significantly stimulate phytoplankton growth. Most B_{12} uptake was by large phytoplankton. Iron addition led to a higher cobalt consumption, but cobalt did not limit phytoplankton growth or cobalamin production	Response of heterotrophic bacteria to nutrient additions?	Bertrand et al., 2007
North Atlantic Ocean (Spanish Coast) (October)	C	B_{12} + inorganic nutrients amendment experiments increased phytoplankton growth more than inorganic nutrient addition alone.	Addition of B_{12} or Fe alone stimulated phytoplankton growth. Increase in bacterial biomass during inorganic nutrient addition. Decrease in bacterial biomass during B_{12} addition.	Were the stimulated bacteria B_{12} consumers or producers? What warrants a decrease in bacterial biomass during B_{12} addition?	Barber-Lluch et al., 2019
North Atlantic Ocean (Spanish Coast) (March)	B*	B_{12} + inorganic nutrient amendment experiments increased phytoplankton growth more than inorganic nutrient addition alone.	No evidence of enhanced heterotrophic bacterial growth after inorganic nutrient additions.	What is limiting B12 production in bacteria?	Barber-Lluch et al., 2019

(Continued)

TABLE 2 | (Continued)

Study site	Case of CIC	Evidence of co-limitation	Notes	Open questions	References
Eastern boundary of the South Atlantic gyre	A or B	Addition of nitrogen, iron and B_{12} or cobalt together significantly increased phytoplankton growth. This might suggest that B_{12} production was limited by cobalt.	No examination of heterotrophic bacterial growth after nutrient additions.	Is Co or B_{12} colimiting (or serially limiting) phytoplankton growth? Was cobalamin production limited by Co or DOC?	Browning et al., 2017
North Atlantic Ocean	A	Cobalt addition increased B_{12} production and phosphate and nitrate uptake. The concentration of dissolved B_{12} was positively correlated with Co concentrations, bacterial productivity, and phytoplankton biomass.	There was not an increase in phytoplankton biomass after Co addition	Change in community composition after nutrient additions?	Panzeca et al., 2008

Gulf of Alaska, an increase of phytoplankton and bacterial growth upon the addition of iron and B_{12} was recorded (Koch et al., 2011). However, with the data provided it is unclear if the observed bacterial growth was stimulated by the iron (indicating Case C) or by increased organic matter produced by actively growing phytoplankton (suggesting Case B). To determine the exact case of CIC present, researchers would have to demonstrate bacterial response to organic carbon addition or isolate the bacteria community independent of phytoplankton and then re-assess growth during iron addition. This is difficult, or impossible, to accomplish in complex communities. Other experimental approaches for determining the mechanism of CIC include the analysis of 'omics data. For example, heterotrophic bacterial growth was not measured in a McMurdo Sound study (Bertrand et al., 2015), but elevated expression of organic carbon acquisition genes in a cobalamin producer, in response to iron addition, suggested that cobalamin production was likely limited by organic carbon availability once iron limitation of primary producers was alleviated.

The study that performed nutrient incubations off the Spanish Coast (Barber-Lluch et al., 2019), revealed seasonal variability in cobalamin co-limitation. Specifically, this study demonstrated a shift from Case B in March to Case C in October. From the data provided we concluded that the bacterial community was not limited by inorganic nutrients in March (eliminating Case C), in contrast to October when bacterial biomass increased upon addition of inorganic nutrients (potentially indicating Case C). Since cobalt availability is usually elevated in coastal regions due to inputs from continental shelf sediments (Noble et al., 2012; Tagliabue et al., 2018), it is unlikely that cobalt is a limiting nutrient. Therefore, we suggest that cobalamin production was most likely limited by a factor produced by the phytoplankton community in March, e.g., DOC, representing Case B.

Evidence for Case A was present in two studies (Panzeca et al., 2008; Browning et al., 2017). At the Eastern Boundary of the South Atlantic gyre, additions of nitrogen and iron with either cobalt or B_{12} increased phytoplankton growth over the combined addition of nitrogen and iron alone (Browning et al., 2017).

Three different CIC interpretations are consistent with the available data. The first would suggest that cobalt is serially limiting phytoplankton growth and B_{12} addition only enhances growth through the indirect addition of cobalt contained in the B_{12} molecule, which must be re-packaged and used in ionic form in other enzymes. The second possibility would be that B_{12} production is limited by cobalt, and cobalt addition thereby enhances phytoplankton growth indirectly through the stimulation of B_{12} production, representing Case A. However, it is also possible that cobalt and B_{12} were both serially limiting phytoplankton growth and that cobalamin producers were limited by DOC, representing Case B. Clear identification of B_{12} stress of phytoplankton growth could have been achieved through the assessment of biomarkers that indicate cobalamin stress of phytoplankton (e.g., MetE), and the limiting factor for cobalamin production (cobalt or DOC) could have been assessed using multi-omic approaches. Panzeca et al. (2008) observed a correlation between cobalt and vitamin B_{12} concentrations, phytoplankton biomass and bacterial production in the North Atlantic Ocean. Although a clear confirmation of CIC is not available due to the lack of incubation experiments assessing phytoplankton biomass responses, it provides evidence for the presence of cobalt limitation of cobalamin production (Case A).

Despite numerous studies on marine nitrogen fixation, sufficient information from experiments that can identify CIC cases involving nitrogen fixation in the ocean are scarce. This is largely due to slow growth rates of many nitrogen fixers (Landolfi et al., 2015) and thereby slow response to nutrient additions, which makes a direct assessment of nutrient limitation of diazotrophs in bottle incubations challenging. Hence, only a few studies are available showing enhanced diazotroph growth after the addition of phosphorus, iron, or DOC while simultaneously demonstrating that phytoplankton growth was limited by nitrogen (Mills et al., 2004; Moisander et al., 2012). A study in the eastern tropical North Atlantic demonstrated phytoplankton growth to be primarily limited by nitrogen, whereas iron and phosphorous additions in the absence of added nitrogen did not influence phytoplankton

growth but significantly enhanced nitrogen fixation rates (Mills et al., 2004), hence suggesting the presence of Case A. However, the effect of DOC addition was not tested in this study and DOC may have been additionally limiting. The observed increase in nitrogen fixation after addition of iron and phosphorous in the absence of enhanced phytoplankton growth does not suggest a major limitation by DOC. Another study in the South Pacific demonstrated an increase in several nitrogen fixers after the addition of iron alone and in combination with organic carbon, at stations where phytoplankton growth was limited by nitrogen (Moisander et al., 2012), suggesting the occurrence of Case A and Case B of CIC. However, iron additions contained EDTA and hence they could not rule out that EDTA may have acted as an additional carbon source and enhanced nitrogen fixation instead of iron. Additionally, some studies suggest the presence of CIC involving nitrogen fixation by demonstrating the nutrient limitation of diazotroph growth in nitrogen deplete conditions using lab-based experiments, cell quotas, transcriptomic and proteomic data (using biomarkers of nutrient stress), and observed correlations between nitrogen fixers and nutrient availability (e.g., Berman-Frank et al., 2001; Sañudo-Wilhelmy et al., 2001; Moore et al., 2009; Chappell et al., 2012; Held et al., 2020). These studies focus on the nutrients limiting nitrogen fixation, but rarely simultaneously look at nutrient limitation of non-diazotrophs within the community. Therefore, they can demonstrate the likely importance of CIC in the context of nitrogen fixation but are unable to identify the underlying case of CIC.

As is evident from these examples, the identification of CIC and its associated case is a difficult task requiring assessments of nutritional status across the entire microbial community. However, as we will describe below, the ability to identify the underlying mechanisms behind instances of co-limitation may improve our ability to predict and monitor the impact CIC has on larger biogeochemical cycles in the present and future ocean.

DISCUSSION

Community Interaction Co-limitation Influence on Biogeochemical Cycles

A major goal in modern marine biogeochemistry is to develop a predictive understanding of the relationships between microbial communities and biogeochemistry (Boyd et al., 2015; Moore et al., 2018; Dutkiewicz et al., 2020; Van de Waal and Litchman, 2020). Predicting future changes in such processes as primary production and carbon export to the deep ocean is a challenging task due to the large number of controlling factors (e.g., nutrient supply, temperature) and large uncertainties in the future trajectories of these variables. However, it is widely acknowledged that community composition and diversity, particularly size class distributions and the abundance of different functional groups, play a large role in determining primary production, nutrient cycling, and aspects of the biological carbon pump (Finkel et al., 2010; Tréguer et al., 2018). The studies listed in section "Evidence of Community Interaction Co-limitation in the Literature" have documented instances where community productivity and

composition are constrained by the biological production of resources like cobalamin. While it can and has been argued that whatever limits the production of that biologically produced resource is the true limiting nutrient (Karl, 2000), without understanding the production, exchange and susceptibility to change of that biological produced nutrient, we risk remaining in the dark about important controls on microbial community productivity and limiting our capacity to predict responses to change. Using cobalamin as an example, we highlight the potential role CIC may play in global biogeochemical processes.

Cobalamin's impact on phytoplankton size distributions is one example of the biogeochemical consequences of BPNs. Phytoplankton cell size is often controlled by nutrient availability, with large phytoplankton such as diatoms typically dominating in nutrient-rich regions, and smaller phytoplankton species dominating in oligotrophic waters (Finkel et al., 2010). Effects of cell size on biogeochemical cycles are partly a result of differences in sinking velocity and impacts on grazer community (Finkel et al., 2010; Ward et al., 2012). For example, large cells sink faster than small cells and therefore may contribute more to the biological carbon pump (Ward et al., 2012). Several studies have suggested that high cobalamin availability favors the growth of larger phytoplankton ($>5 \mu m$) perhaps because diatoms and dinoflagellates are more likely than other groups to be cobalamin auxotrophs and have many larger representatives (Croft et al., 2005; Sañudo-Wilhelmy et al., 2006; Gobler et al., 2007; Bertrand et al., 2015). In contrast, cyanobacteria may gain competitive advantage in low cobalamin environments because they produce and use their own cobalamin-like molecule, pseudo-cobalamin (Bertrand and Allen, 2012; Helliwell et al., 2016; Heal et al., 2017). This suggests that cobalamin availability may drive changes in size classes and functional groups that could influence contribution to the biological carbon pump. Identifying the mechanisms and interactions that underpin how BPNs impact community composition could help further determine CIC's role in biogeochemical cycles.

Cobalamin CIC also influences global biogeochemical processes by controlling primary production, and, potentially, algal bloom development. A study at the Spanish Coast, which provides high temporal resolution data across all four seasons, cobalamin was shown to co-limit phytoplankton growth in combination with inorganic nutrients in March. We hypothesized earlier that this scenario likely represents Case B, where cobalamin production is limited by DOM. This period of co-limitation was followed by the transition to cobalamin being the primary limiting nutrient for phytoplankton growth in April at the onset of the spring bloom (most productive period of the year is observed in May), when inorganic nutrients are high and bacterial counts low (Barber-Lluch et al., 2019). This could suggest that although limitation by inorganic nutrients in April was partly relieved, phytoplankton remained cobalamin limited. In turn, the limited phytoplankton growth and DOM production continued to restrict cobalamin production, thereby hindering the community's ability to quickly recover from nutrient limitation. Although an argument could be made that non-auxotrophic phytoplankton growth may eventually relieve DOM limitation of cobalamin producers, it

is unclear how the enhanced non-auxotrophic phytoplankton growth, and subsequent DOM production, affect the timing and strength of cobalamin production. This suggests that the strength and mechanisms of the cobalamin feedback loop affected community productivity and succession over the span of weeks. Furthermore, cobalamin limitation in spring may result in a shift in phytoplankton community to the dominance of species with cobalamin-independent growth at the onset of the bloom. Perhaps due to higher resource use costs for cobalamin-independent growth (Bertrand et al., 2013), this may consequently result in a time-lag in maximal productivity and an increased nutrient consumption, hence decreasing the overall productivity of a bloom. Investigating the spatio-temporal patterns of CIC, especially during periods of high productivity, could uncover additional mechanisms of CIC and its influence on global primary production in the ocean, now and in the future.

Community Interaction Co-limitation in the Future Ocean

Considering the controls of CIC and biological produced nutrients could prove important when predicting the impact of on-going climatic change. Assessing the direct impact on CIC is challenging, due to large uncertainties in predicting the spatial extent and intensity of such environmental changes, and the limited understanding of potential biological responses to them. However, an improved understanding of the relationships between key environmental variables and BPNs might be a first step to anticipate the fate of CIC in the future ocean. Here we explore the possible impact predicted environmental conditions might have on cobalamin availability and requirements, and their potential consequences for CIC.

Cobalamin availability, and instances of cobalamin CIC, might be influenced by changes in light, nutrient supply and sea surface temperature (SST) (Bopp et al., 2001, 2013; Sarmiento et al., 2004a,b; Doney et al., 2012). An increase in irradiance is predicted to occur in many ocean regions due to enhanced stratification and sea-ice retreat (Grebmeier et al., 2010; Doney et al., 2012). This could have a negative impact on the availability of cobalamin by increasing its photodegradation (Juzeniene and Nizauskaite, 2013), thereby enhancing the potential for cobalamin limitation and CIC. One factor that could influence the patterns of CIC in the future ocean is the predicted changes in the supply of other essential nutrients (Tagliabue et al., 2017; Moore et al., 2018). For example, an increased supply of a limiting nutrient in a given region may result in an increase in productivity and a shift to another nutrient becoming limiting or co-limiting. Changes in SST are likely to have an impact on metabolic rates (Talling, 1955; Eppley, 1972) which could influence cobalamin production rates and competition for cobalamin, or other BPNs, within communities. Labile DOM availability has been suggested to be a more important factor for bacterial production than temperature in polar regions (Kirchman et al., 2009) indicating tightly coupled changes in metabolic rates of phytoplankton and bacteria. Other studies have suggested a larger increase in bacterial activity and remineralization rates compared to

photosynthesis rates with increasing temperature (Lopez-Urrutia et al., 2006) suggesting a larger increase in cobalamin production compared to consumption by phytoplankton, thereby reducing the potential for cobalamin limitation. However, enhanced bacterial growth could also result in increased competition for cobalamin, as auxotrophic bacterial growth will likely also positively correlate with temperature. Overall, the response of cobalamin CIC will be a result of the balance between the relative response of cobalamin producers, phytoplankton growth and (competing) bacterial cobalamin consumers, as well as how their nutrient production rates and cellular quotas/nutrient demand changes. Few relationships and feedback mechanisms of CIC are identified to date and a more detailed understanding of the effects on community interactions is required. Therefore, predicting the response of CIC to environmental changes might benefit from incorporating ecological and evolutionary theories that aim to understand microbial interactions and dependencies. Furthermore, considering evolutionary forces underpinning microbial interactions may streamline evaluation of possible adaptations of microbial communities to environmental change in the future.

Incorporating Ecological and Evolutionary Theory Into Community Interaction Co-limitation

The idea that microbial communities are more than the sum of the species present, due to the intricate web of interactions between community members, is foundational to the concept of CIC. Microbial interactions underpin productivity, composition, and resilience in microbial communities, factors that determine a community's contribution to larger ecological processes. To date, the field studies of nutrient co-limitation that lead us to hypothesize that CIC may play an important role in larger biogeochemical cycles are largely focused on bulk biogeochemical parameters and processes: by and large, the measurements and principles applied in such studies fail to examine the elaborate interactions in marine microbial communities. To attempt to elucidate the experience, and response, of microbial communities under nutrient co-limitation, additional steps must be taken that understand the potential ecological and evolutionary processes at play. Ultimately, this could help to streamline identification of key relationships and underpinning interactions and begin to systematically investigate the factors that influence them.

Microbial interactions are ubiquitous in the marine environment and are usually classified into categories such as mutualism, commensalism, and competition (Ramanan et al., 2016). However, these associations are extremely dynamic and can quickly switch from cooperation to competition, and vice versa, depending on an organism's metabolic status and nutrient availability (Seyedsayamdost et al., 2011; Hoek et al., 2016). These associations are not (entirely) random, and it has been shown that specific groups of bacteria and algae tend to co-occur in the marine environment (Zelezniak et al., 2015; Milici et al., 2016). Studies have discovered that phytoplankton, or "hosts," modulate their associated bacterial communities through excreted metabolites (Fu et al., 2020;

Shibl et al., 2020). This is beneficial since the close interactions a phytoplankton supports in its phycosphere could determine its access to resources and impact its physiology, productivity, and survival (Amin et al., 2009; Bolch et al., 2011, 2017). Therefore, clarifying the direction, plasticity, and intensity of microbial interactions during periods of BPN limitation are relevant factors that can be used to identify cases and predict consequences of CIC. A more detailed understanding of when and why close mutualistic interaction are successful strategies would be particularly beneficial in the context of resource limitation: it affects the energy put into nutrient acquisition, ultimately impacting an organism's productivity and fitness. Exchanging nutrients in a direct interaction could allow more consistent access to resources and therefore stability during periods of nutrient limitation. However, there are risks associated with this strategy as it relies on the fitness and presence of a partner and could be more energetically costly. In contrast, an organism that is scavenging nutrients from the surrounding environment may have more restricted access when a resource becomes scarce. The success of this organism relies on efficient acquisition strategies, as there would most likely be competition for the resource from all the consumers within the community. To draw an example from CIC, organisms that are in direct, mutualistic interaction exchanging cobalamin during times of limitation may out-compete scavengers, which would influence composition and species succession in the community. The prevalence of interaction-based nutrient acquisition strategies within a microbial community should be acknowledged when interpreting instances of CIC as it could identify organisms more vulnerable to CIC (e.g., scavengers) and factors that influence formation of cooperative interactions. Incorporating microbial interactions into resource limitation studies could uncover larger ecological or evolutionary patterns and perspectives regarding when and where these strategies are successful.

Community interaction co-limitation can be viewed as the imbalance of resource supply and demand between groups within a community, which is likely to be impacted by community composition. Therefore, enhanced insight into the ecological principles that control microbial community organization could enable better predictions of the patterns and mechanisms controlling CIC, and their susceptibility to change. Recently, computational modeling at a scale unattainable by culture studies alone, has identified important controls on community assembly (Pacheco et al., 2019). Microbial community assembly appears replicable and can be explained by a mix of environmental factors, nutrient availability, microbial interactions, and metabolite profiles of key species (Friedman et al., 2017; Louca et al., 2017; Goldford et al., 2018; Fu et al., 2020). Cross-feeding is a nutrient-centric type of microbial interaction and is becoming recognized as central control of microbial community assembly (Goldford et al., 2018; Pacheco et al., 2019). Cross-feeding is the transfer of metabolites between organisms that improve the fitness of recipient and is classified into specific types based on reciprocity of exchange (D'Souza et al., 2018; Smith et al., 2019). Until recently, it was assumed the cross-feeding had to have a metabolic cost for the producing

organisms, however, it has been demonstrated that metabolites which do not impact an organism's fitness when excreted (e.g., costless) enrich minimal nutrient environments and support interactions between species (Pacheco et al., 2019).

Analogies can be made between specific types of cross-feeding and cases of CIC, as the BPN is produced by one group of organisms and used by another. For example, Case A of CIC can be compared to a simple case of metabolite cross-feeding where one organism releases a metabolite that is used by another organism (Smith et al., 2019). However, unlike CIC, the original classification of metabolite cross-feeding defined the exchanged metabolite as a waste product that cannot be further metabolized by the producing bacteria. Case B is comparable to mutual, or two way, cross-feeding because of the exchange of both labile carbon and cobalamin between organisms. These analogies are most easily transferrable to a bipartite interaction scale, however, CIC adds an important dimension to cross feeding concepts, as it emphasizes that possibility that "public good" metabolites can limit production at a community level. Interdisciplinary research into cross-feeding and resource limitation could identify other metabolites, or their precursors, that control community production and assembly further identifying instances of CIC. Exploring the labile carbon requirements of cobalamin producers or nitrogen fixers is an exciting avenue for further research that could provide additional insight into Case B of CIC. Systematically integrating community assembly models and theories into research surrounding CIC (and vice versa) would be a powerful step forward in enabling predictions of the prevalence of and mechanisms behind CIC.

Ecological interactions between micro-organisms are the result of millions of years of co-evolution and thus ecological theories cannot be discussed without the appreciation of evolution. Evolutionary theories are being developed to describe how the ecological interactions between micro-organisms emerged and control microbial community productivity and composition. Recent studies have sought to better understand the evolution of microbial interactions using evolutionary game theory, examining the fitness response of individuals that are interacting within a community (Frey, 2010; Hummert et al., 2014). These are particularly powerful when coupled with genome-scale metabolic networks which allow growth predictions about interactions and metabolic exchange based on organism's genomes (Zomorrodi and Segrè, 2017). Such research is developing theoretical principles, based on evolutionary mechanisms, which could more accurately describe resulting future shifts in the form and function of microbial communities. Incorporating such theories into resource limitation research could provide a framework to anticipate the response of CIC and key microbial interactions in the modern and future ocean.

One evolutionary theory that could be particularly relevant in resource-limited environments is The Black Queen Hypothesis, which describes the evolution of obligate relationships in microbial communities (Morris et al., 2012; Mas et al., 2016). It hypothesizes that "mutualistic" obligate interactions can arise because of leaky production of a "common good," which could be a nutrient, by the "helper," which prompts loss of function in another organism, the "beneficiary" (Mas et al., 2016). This

produces an obligate interaction between the beneficiary and helper, as the beneficiary is no longer able to produce the common good nutrient for the fitness benefit of a reduced genome (Giovannoni et al., 2014; Mas et al., 2016). This theory is also particularly relevant as it considers the cost of production of a particular "public good" which could be relevant for understanding why such cases of CIC can occur. Investigating the cost of the BPN (production vs. acquisition) could help researchers understand which BPNs are likely to be a limiting nutrient in the environment and factors that influence them. The theory may shed some light on the occurrence and evolutionary benefit of microbial dependencies and nutrient exchange during times of resource limitation. Once the types (e.g., obligate or facultative) of key interactions underpinning instances of CIC are identified, the details of nutrient exchange and the associated feedback loop might become clearer, furthering investigation into factors that influence an interaction's stability and potentially helping to predict responses in an eco-evolutionary framework. Therefore, systematically incorporating developments in microbial evolution and ecology will support better synthesis, interpretation and prediction of key microbial interactions that underpin cases of CIC.

Data Gaps, Uncertainties, and Future Directions

While the presence of co-limitation in the ocean is known to be widespread (e.g., around 8% of the surface ocean is estimated to experience co-limitation by nitrogen and iron; Browning et al., 2017), we largely lack a detailed understanding of mechanisms behind co-limitation and its broader effect on global biogeochemical cycles. Studies investigating CIC are scarce to date, resulting in limited knowledge about the variety of biologically produced resources that can lead to CIC. Low spatial and temporal resolution in documented cases of CIC restricts our understanding of the spatial extent and seasonal distribution of CIC and its driving factors. However, understanding the mechanisms and susceptibility of CIC to change in future ocean directly relies on researcher's ability to detect and identify it. This is not straightforward, as it requires the combination of knowledge on different scales including individual organism's nutritional requirements, community dynamics, and global patterns of nutrient availability and species distribution.

Bottle incubation bioassay experiments offer a valuable insight into the oceanic distribution of nutrient limitations, despite their limitations (e.g., bottle-induced changes in community composition; failure to capture grazing pressure). These studies often do not fully resolve the type of co-limitation which exists and on their own don't allow researchers to draw conclusions about the community interactions at play. Identifying the biological nutrient producers and consumers in any given environment is a first step in assessing the supply and demand that constrains nutrient availability. However, these identifications will not always reveal the key interactions or community dynamics underpinning the CIC. As highlighted previously, it is important to understand the role bacteria play in nutrient availability and exchange in microbial communities. As such, an important avenue for further research into CIC is the recognition and incorporation of bacterial community response into resource limitation and nutrient addition bioassays.

Auxotrophy is widespread in the marine environment and could have major implications for microbial community interdependencies and nutrient requirements, yet it is still poorly understood (Johnson et al., 2020). To identify additional types of CIC, and monitor the cases of CIC presented, it is critical to understand the occurrence of auxotrophy in species as well as auxophore (metabolites required by auxotrophs) requirements and the factors that influence them. This includes the occurrence of precursor auxotrophy, a budding field of interest (Paerl et al., 2017, 2018a). Additionally, knowing when a community is dominated by auxotrophs might suggest that a community is more vulnerable to cases of CIC. Finally, having an ongoing atlas of common auxophores and instances of auxotrophies in specific primary producers could be of use for studies of CIC.

Sensitive molecular analysis techniques will be required to investigate the interactions and uncover the feedback loops that are critical for defining and monitoring cases of CIC. A successful analysis would uncover (1) which organisms are stressed for or limited by a BPN, (2) which organisms are producing that nutrient, (3) what is limiting the production/transformation of that nutrient and (4) what the nature of interactions between the relevant microbial groups. Although this task seems daunting, it is becoming more feasible with current advances in multi-layered omics analysis.

One approach to uncover the mechanisms of CIC present is by assessing gene or protein expression profiles through meta-transcriptomic or meta-proteomic analyses during nutrient incubation experiments in field studies. Gene expression analyses could allow the identification of organisms that are stressed for or produce a BPN. For example, through meta-transcriptomics, organisms involved in cobalamin cycle and the feedback loops that controlled cobalamin availability were identified in the Southern Ocean (Case B, cobalamin CIC) (Bertrand et al., 2015). Lab cultures can also be a powerful tool for assessing changes in gene and protein expression under varying growth conditions, including pure cultures and co-cultures of interacting microbial species, and can facilitate the identification of relevant target genes and proteins (Helliwell et al., 2018; Wu et al., 2019). Recently, by incorporating various meta-omics techniques, researchers have already made significant advances in the identification of key mechanisms and metabolites involved with microbial interactions that could lead to additional types of CIC (Amin et al., 2015; Bertrand et al., 2015; Durham et al., 2017; Helliwell et al., 2018; Shibl et al., 2020).

Meta-omics studies can also serve as the basis for developing targeted approaches for detecting and monitoring CIC, for instance through identifying critical metabolites and biomarker genes or protein abundances that indicate different nutrient limitation scenarios (Bertrand et al., 2015; Chappell et al., 2015; Heal et al., 2019; Ustick et al., 2021). Such targeted assays require previous knowledge about producers and consumers present and the specific genes of proteins expression patterns sensitive to certain nutrient limitations. However, they can be used more

broadly, as the analysis can be faster, and more sensitive and specific than meta-omics approaches. As such, they are a useful tool for assessing global nutrient limitation and CIC patterns when used alone or in combination with bioassay experiments. For example, the use of biomarker peptides for MetE and CBA1 expression (cobalamin stress markers) can be used to indicate regions of cobalamin stress or limitation in the field (Bertrand et al., 2012, 2013). Other biomarker peptide mapping studies have collected valuable data about the nitrogen, iron, and phosphorus status of marine cyanobacterial communities (Saito et al., 2015). These methods could also be used for identification of organisms responsible for biological nutrient production using proteins involved in biological nutrient synthesis (e.g., CobO protein for cobalamin production). However, there is still a big gap in our knowledge about suitable targets to detect nutrient stress for most marine phytoplankton for known limiting or co-limiting nutrients, particularly BPNs involved in CIC.

Furthermore, putting CIC in a global context requires incorporating such evidence found at the organismal and community level with larger oceanographic trends to identify potential spatio-temporal patterns and driving factors that control the underpinning mechanisms of limitation. To map and predict the consequences of CIC in the modern and future ocean, we must apply interdisciplinary methods and approaches capable of elucidating important feedback loops that control nutrient exchange and availability in microbial communities and key factors that influence them. Recent advances in experimental design and molecular approaches provide a hopeful outlook.

CONCLUSION

The interactions within a microbial community can control nutrient availability and can profoundly impact metabolic status of community members. Incorporating community interactions into research on resource (co-)limitation enables the identification of community level processes that affect larger biogeochemical cycles. Defining community interaction co-limitation, CIC, as fourth category of nutrient co-limitation allows researchers to systematically combine the tools and findings of microbial ecology and evolution with resource

limitation studies to obtain clearer descriptions of mechanisms of nutrient limitation in the environment. It also provides a consistent language and framework that can be used to describe the relevant processes and interactions, hopefully improving cross-disciplinary communication on these subjects. We suggest that the CIC framework will support the systematic examination of the role community interactions have in resource limitation which, in turn, will support more accurate estimates and predictions of the impact these processes have on global biogeochemical cycles.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Amin, S. A., Green, D. H., Hart, M. C., Küpper, F. C., Sunda, W. G., and Carrano, C. J. (2009). Photolysis of iron–siderophore chelates promotes bacterial–algal mutualism. *Proc. Natl. Acad. Sci. U. S. A.* 106, 17071–17076. doi: 10.1073/pnas.0905512106

Amin, S. A., Hmelo, L. R., van Tol, H. M., Durham, B. P., Carlson, L. T., Heal, K. R., et al. (2015). Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* 522, 98–101. doi: 10.1038/nature14488

Amin, S. A., Parker, M. S., and Armbrust, E. V. (2012). Interactions between Diatoms and Bacteria. *Microbiol. Mol. Biol. Rev.* 76, 667–684. doi: 10.1128/MMBR.00007-12

Banerjee, R., and Ragsdale, S. W. (2003). The Many Faces of Vitamin B 12: catalysis by Cobalamin-Dependent Enzymes. *Annu. Rev. Biochem.* 72, 209–247. doi: 10.1146/annurev.biochem.72.121801.161828

Barber-Lluch, E., Hernández-Ruiz, M., Prieto, A., Fernández, E., and Teira, E. (2019). Role of vitamin B12 in the microbial plankton response to nutrient enrichment. *Mar. Ecol. Prog. Ser.* 626, 29–42. doi: 10.3354/meps13077

Berman-Frank, I., Cullen, J. T., Shaked, Y., Sherrell, R. M., and Falkowski, P. G. (2001). Iron availability, cellular iron quotas, and nitrogen fixation in Trichodesmium. *Limnol. Oceanogr.* 46, 1249–1260. doi: 10.4319/lo.2001.46.6.1249

Bertrand, E. M., Allen, A. E., Dupont, C. L., Norden-Krichmar, T. M., Bai, J., Valas, R. E., et al. (2012). Influence of cobalamin scarcity on diatom molecular physiology and identification of a cobalamin acquisition protein. *Proc. Natl. Acad. Sci. U. S. A.* 109, E1762–E1771. doi: 10.1073/pnas.1201731109

Bertrand, E. M., and Allen, A. E. (2012). Influence of vitamin B auxotrophy on nitrogen metabolism in eukaryotic phytoplankton. *Front. Microbiol.* 3:375. doi: 10.3389/fmicb.2012.00375

Bertrand, E. M., McCrow, J. P., Moustafa, A., Zheng, H., McQuaid, J. B., Delmont, T. O., et al. (2015). Phytoplankton–bacterial interactions mediate micronutrient

colimitation at the coastal Antarctic sea ice edge. *Proc. Natl. Acad. Sci. U. S. A.* 112, 9938–9943. doi: 10.1073/pnas.1501615112

Bertrand, E. M., Moran, D. M., McIlvin, M. R., Hoffman, J. M., Allen, A. E., and Saito, M. A. (2013). Methionine synthase interreplacement in diatom cultures and communities: implications for the persistence of B12 use by eukaryotic phytoplankton. *Limnol. Oceanogr.* 58, 1431–1450. doi: 10.4319/lo.2013.58.4.1431

Bertrand, E. M., Saito, M. A., Rose, J. M., Riesselman, C. R., Lohan, M. C., Noble, A. E., et al. (2007). Vitamin B12 and iron colimitation of phytoplankton growth in the Ross Sea. *Limnol. Oceanogr.* 52, 1079–1093. doi: 10.4319/lo.2007.52.3.1079

Bolch, C. J. S., Bejoy, T. A., and Green, D. H. (2017). Bacterial Associates Modify Growth Dynamics of the Dinoflagellate Gymnodinium catenatum. *Front. Microbiol.* 8:670. doi: 10.3389/fmicb.2017.00670

Bolch, C. J. S., Subramanian, T. A., and Green, D. H. (2011). The Toxic Dinoflagellate Gymnodinium Catenatum (dinophyceae) Requires Marine Bacteria for Growth1. *J. Phycol.* 47, 1009–1022. doi: 10.1111/j.1529-8817.2011.01043.x

Bonnet, S., Dekaezemacker, J., Turk-Kubo, K. A., Moutin, T., Hamersley, R. M., Grosso, O., et al. (2013). Aphotic N2 Fixation in the Eastern Tropical South Pacific Ocean. *PLoS One* 8:e81265. doi: 10.1371/journal.pone.0081265

Bopp, L., Monfray, P., Aumont, O., Dufresne, J.-L., Treut, H. L., Madec, G., et al. (2001). Potential impact of climate change on marine export production. *Glob. Biogeochem. Cycles* 15, 81–99. doi: 10.1029/1999GB001256

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences* 10, 6225–6245. doi: 10.5194/bg-10-6225-2013

Boyd, P. W., Lennartz, S. T., Glover, D. M., and Doney, S. C. (2015). Biological ramifications of climate-change-mediated oceanic multi-stressors. *Nat. Clim. Change* 5, 71–79. doi: 10.1038/nclimate2441

Browning, T. J., Achterberg, E. P., Engel, A., and Mawji, E. (2021). Manganese co-limitation of phytoplankton growth and major nutrient drawdown in the Southern Ocean. *Nat. Commun.* 12:884. doi: 10.1038/s41467-021-21122-6

Browning, T. J., Achterberg, E. P., Rapp, I., Engel, A., Bertrand, E. M., Tagliabue, A., et al. (2017). Nutrient co-limitation at the boundary of an oceanic gyre. *Nature* 551, 242–246. doi: 10.1038/nature24063

Chappell, P. D., Moffett, J. W., Hynes, A. M., and Webb, E. A. (2012). Molecular evidence of iron limitation and availability in the global diazotroph Trichodesmium. *ISME J.* 6, 1728–1739. doi: 10.1038/ismej.2012.13

Chappell, P. D., Whitney, L. P., Wallace, J. R., Darer, A. I., Jean-Charles, S., and Jenkins, B. D. (2015). Genetic indicators of iron limitation in wild populations of Thalassiosira oceanica from the northeast Pacific Ocean. *ISME J.* 9, 592–602. doi: 10.1038/ismej.2014.171

Collier, J. L., Baker, K. M., and Bell, S. L. (2009). Diversity of urea-degrading microorganisms in open-ocean and estuarine planktonic communities. *Environ. Microbiol.* 11, 3118–3131. doi: 10.1111/j.1462-2920.2009.02016.x

Collier, J. L., Lovindeer, R., Xi, Y., Radway, J. C., and Armstrong, R. A. (2012). Differences in Growth and Physiology of Marine Synechococcus (cyanobacteria) on Nitrate Versus Ammonium Are Not Determined Solely by Nitrogen Source Redox State1. *J. Phycol.* 48, 106–116. doi: 10.1111/j.1529-8817.2011.01100.x

Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J., and Smith, A. G. (2005). Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* 438, 90–93. doi: 10.1038/nature04056

D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., and Kost, C. (2018). Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Nat. Prod. Rep.* 35, 455–488. doi: 10.1039/C8NP00009C

Danger, M., Daufresne, T., Lucas, F., Pissard, S., and Lacroix, G. (2008). Does Liebig's law of the minimum scale up from species to communities? *Oikos* 117, 1741–1751. doi: 10.1111/j.1600-0706.2008.16793.x

Doney, S. C., Ruckelshaus, M., Emmett Duffy, J., Barry, J. P., Chan, F., English, C. A., et al. (2012). Climate Change Impacts on Marine Ecosystems. *Annu. Rev. Mar. Sci.* 4, 11–37. doi: 10.1146/annurev-marine-041911-111611

Dowling, D. P., Croft, A. K., and Drennan, C. L. (2012). Radical Use of Rossmann and TIM Barrel Architectures for Controlling Coenzyme B12 Chemistry. *Annu. Rev. Biophys.* 41, 403–427. doi: 10.1146/annurev-biophys-050511-102225

Durham, B. P., Boysen, A. K., Carlson, L. T., Groussman, R. D., Heal, K. R., Cain, K. R., et al. (2019). Sulfonate-based networks between eukaryotic phytoplankton and heterotrophic bacteria in the surface ocean. *Nat. Microbiol.* 4, 1706–1715. doi: 10.1038/s41564-019-0507-5

Durham, B. P., Dearth, S. P., Sharma, S., Amin, S. A., Smith, C. B., Campagna, S. R., et al. (2017). Recognition cascade and metabolite transfer in a marine bacteria-phytoplankton model system. *Environ. Microbiol.* 19, 3500–3513. doi: 10.1111/1462-2920.13834

Dutkiewicz, S., Cermeno, P., Jahn, O., Follows, M. J., Hickman, A. E., Taniguchi, D. A. A., et al. (2020). Dimensions of marine phytoplankton diversity. *Biogeosciences* 17, 609–634. doi: 10.5194/bg-17-609-2020

Eppley, R. W. (1972). Temperature and phytoplankton growth in the sea. *Fish. Bull.* 70, 1063–1085.

Finkel, Z. V., Beardall, J., Flynn, K. J., Quigg, A., Rees, T. A. V., and Raven, J. A. (2010). Phytoplankton in a changing world: cell size and elemental stoichiometry. *J. Plankton Res.* 32, 119–137. doi: 10.1093/plankt/fbp098

Frey, E. (2010). Evolutionary game theory: theoretical concepts and applications to microbial communities. *Phys. Stat. Mech. Appl.* 389, 4265–4298. doi: 10.1016/j.physa.2010.02.047

Friedman, J., Higgins, L. M., and Gore, J. (2017). Community structure follows simple assembly rules in microbial microcosms. *Nat. Ecol. Evol.* 1, 1–7. doi: 10.1038/s41559-017-0109

Fu, H., Uchimiya, M., Gore, J., and Moran, M. A. (2020). Ecological drivers of bacterial community assembly in synthetic phycospheres. *Proc. Natl. Acad. Sci. U. S. A.* 117, 3656–3662. doi: 10.1073/pnas.1917265117

Garcia, N. S., Sexton, J., Riggins, T., Brown, J., Lomas, M. W., and Martiny, A. C. (2018). High Variability in Cellular Stoichiometry of Carbon, Nitrogen, and Phosphorus Within Classes of Marine Eukaryotic Phytoplankton Under Sufficient Nutrient Conditions. *Front. Microbiol.* 9:543. doi: 10.3389/fmicb.2018.00543

Giovannoni, S. J., Cameron Thrash, J., and Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *ISME J.* 8, 1553–1565. doi: 10.1038/ismej.2014.60

Gobler, C., Norman, C., Panzeca, C., Taylor, G., and Sañudo-Wilhelmy, S. (2007). Effect of B-vitamins (B1, B12) and inorganic nutrients on algal bloom dynamics in a coastal ecosystem. *Aquat. Microb. Ecol.* 49, 181–194. doi: 10.3354/ame01132

Goldford, J. E., Lu, N., Baji, D., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P., et al. (2018). Emergent simplicity in microbial community assembly. *Science* 361, 469–474. doi: 10.1126/science.aat1168

Grant, M. A., Kazamia, E., Cicuta, P., and Smith, A. G. (2014). Direct exchange of vitamin B12 is demonstrated by modelling the growth dynamics of algal–bacterial cocultures. *ISME J.* 8, 1418–1427. doi: 10.1038/ismej.2014.9

Grebmeier, J. M., Moore, S. E., Overland, J. E., Frey, K. E., and Gradinger, R. (2010). Biological Response to Recent Pacific Arctic Sea Ice Retreats. *Eos Trans. AGU* 91:161. doi: 10.1029/2010EO180001

Harpole, W. S., Ngai, J. T., Cleland, E. E., Seabloom, E. W., Borer, E. T., Bracken, M. E. S., et al. (2011). Nutrient co-limitation of primary producer communities. *Ecol. Lett.* 14, 852–862. doi: 10.1111/j.1461-0248.2011.01651.x

Heal, K. R., Carlson, L. T., Devol, A. H., Armbrust, E. V., Moffett, J. W., Stahl, D. A., et al. (2014). Determination of four forms of vitamin B 12 and other B vitamins in seawater by liquid chromatography/tandem mass spectrometry: analysis of forms of B 12 in seawater. *Rapid Commun. Mass Spectrom.* 28, 2398–2404. doi: 10.1002/rcm.7040

Heal, K. R., Kellogg, N. A., Carlson, L. T., Lionheart, R. M., and Ingalls, A. E. (2019). Metabolic Consequences of Cobalamin Scarcity in the Diatom Thalassiosira pseudonana as Revealed Through Metabolomics. *Protist* 170, 328–348. doi: 10.1016/j.protis.2019.05.004

Heal, K. R., Qin, W., Ribalet, F., Bertagnolli, A. D., Coyote-Maestas, W., Hmelo, L. R., et al. (2017). Two distinct pools of B 12 analogs reveal community interdependencies in the ocean. *Proc. Natl. Acad. Sci. U. S. A.* 114, 364–369. doi: 10.1073/pnas.1608462114

Held, N. A., Webb, E. A., McIlvin, M. M., Hutchins, D. A., Cohen, N. R., Moran, D. M., et al. (2020). Co-occurrence of Fe and P stress in natural populations of the marine diazotroph <i>Trichodesmium<i/>. *Biogeosciences* 17, 2537–2551. doi: 10.5194/bg-17-2537-2020

Helliwell, K. E. (2017). The roles of B vitamins in phytoplankton nutrition: new perspectives and prospects. *New Phytol.* 216, 62–68. doi: 10.1111/nph.14669

Helliwell, K. E., Lawrence, A. D., Holzer, A., Kudahl, U. J., Sasso, S., Kräutler, B., et al. (2016). Cyanobacteria and Eukaryotic Algae Use Different Chemical Variants of Vitamin B12. *Curr. Biol.* 26, 999–1008. doi: 10.1016/j.cub.2016.02.041

Helliwell, K. E., Pandhal, J., Cooper, M. B., Longworth, J., Kudahl, U. J., Russo, D. A., et al. (2018). Quantitative proteomics of a B12-dependent alga grown in coculture with bacteria reveals metabolic tradeoffs required for mutualism. *New Phytol.* 217, 599–612. doi: 10.1111/nph.14832

Helliwell, K. E., Wheeler, G. L., Leptos, K. C., Goldstein, R. E., and Smith, A. G. (2011). Insights into the Evolution of Vitamin B12 Auxotrophy from Sequenced Algal Genomes. *Mol. Biol. Evol.* 28, 2921–2933. doi: 10.1093/molbev/msr124

Ho, T.-Y., Quigg, A., Finkel, Z. V., Milligan, A. J., Wyman, K., Falkowski, P. G., et al. (2003). The Elemental Composition of Some Marine Phytoplankton1. *J. Phycol.* 39, 1145–1159. doi: 10.1111/j.0022-3646.2003.03-090.x

Hodgkin, D. C., Kamper, J., Mackay, M., Pickworth, J., Trueblood, K. N., and White, J. G. (1956). Structure of vitamin B12. *Nature* 178, 64–66. doi: 10.1038/178064a0

Hoek, T. A., Axelrod, K., Biancalani, T., Yurtsev, E. A., Liu, J., and Gore, J. (2016). Resource Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-Feeding Mutualism. *PLoS Biol.* 14:e1002540. doi: 10.1371/journal.pbio.1002540

Hummert, S., Bohl, K., Basanta, D., Deutsch, A., Werner, S., Theißen, G., et al. (2014). Evolutionary game theory: cells as players. *Mol. BioSyst.* 10, 3044–3065. doi: 10.1039/C3MB70602H

Hutchins, D. A., Witter, A. E., Butler, A., and Luther, G. W. (1999). Competition among marine phytoplankton for different chelated iron species. *Nature* 400, 858–861. doi: 10.1038/23680

Johnson, W. M., Alexander, H., Bier, R. L., Miller, D. R., Muscarella, M. E., Pitz, K. J., et al. (2020). Auxotrophic interactions: a stabilizing attribute of aquatic microbial communities? *FEMS Microbiol. Ecol.* 96:fiaa115. doi: 10.1093/femsec/fiaa115

Juzeniene, A., and Nizauskaite, Z. (2013). Photodegradation of cobalamins in aqueous solutions and in human blood. *J. Photochem. Photobiol. B Biol.* 122, 7–14. doi: 10.1016/j.jphotobiol.2013.03.001

Karl, D. M. (2000). Phosphorus, the staff of life. *Nature* 406, 31–33. doi: 10.1038/35017683

Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., et al. (2002). "Dinitrogen fixation in the world's oceans," in *The Nitrogen Cycle at Regional to Global Scales*, eds E. W. Boyer and R. W. Howarth (Dordrecht: Springer Netherlands), 47–98. doi: 10.1007/978-94-017-3405-9_2

Kazamia, E., Czesnick, H., Nguyen, T. T. V., Croft, M. T., Sherwood, E., Sasso, S., et al. (2012). Mutualistic interactions between vitamin B12-dependent algae and heterotrophic bacteria exhibit regulation. *Environ. Microbiol.* 14, 1466–1476. doi: 10.1111/j.1462-2920.2012.02733.x

Kirchman, D. L., Morán, X. A. G., and Ducklow, H. (2009). Microbial growth in the polar oceans — role of temperature and potential impact of climate change. *Nat. Rev. Microbiol.* 7, 451–459. doi: 10.1038/nrmicro2115

Koch, F., Hattenrath-Lehmann, T. K., Goleski, J. A., Sanudo-Wilhelmy, S., Fisher, N. S., and Gobler, C. J. (2012). Vitamin B1 and B12 Uptake and Cycling by Plankton Communities in Coastal Ecosystems. *Front. Microbiol.* 3:363. doi: 10.3389/fmicb.2012.00363

Koch, F., Marcoval, M. A., Panzeca, C., Bruland, K. W., Sañudo-Wilhelmy, S. A., and Gobler, C. J. (2011). The effect of vitamin B12 on phytoplankton growth and community structure in the Gulf of Alaska. *Limnol. Oceanogr.* 56, 1023–1034. doi: 10.4319/lo.2011.56.3.1023

Landolfi, A., Koeve, W., Dietze, H., Kähler, P., and Oschlies, A. (2015). A new perspective on environmental controls of marine nitrogen fixation. *Geophys. Res. Lett.* 42, 4482–4489. doi: 10.1002/2015GL063756

Lopez-Urrutia, A., San Martin, E., Harris, R. P., and Irigoien, X. (2006). Scaling the metabolic balance of the oceans. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8739–8744. doi: 10.1073/pnas.0601137103

Louca, S., Jacques, S. M. S., Pires, A. P. F., Leal, J. S., Srivastava, D. S., Parfrey, L. W., et al. (2017). High taxonomic variability despite stable functional structure across microbial communities. *Nat. Ecol. Evol.* 1:0015. doi: 10.1038/s41559-016-0015

Luo, H., Benner, R., Long, R. A., and Hu, J. (2009). Subcellular localization of marine bacterial alkaline phosphatases. *Proc. Natl. Acad. Sci. U. S. A.* 106, 21219–21223. doi: 10.1073/pnas.0907586106

Ma, A. T., Tyrell, B., and Beld, J. (2020). Specificity of cobamide remodeling, uptake and utilization in *Vibrio cholerae*. *Mol. Microbiol.* 113, 89–102. doi: 10.1111/mmi.14402

Manck, L. E., Park, J., Tully, B. J., Poire, A. M., Bundy, R. M., Dupont, C. L., et al. (2022). Petrobactin, a siderophore produced by Alteromonas, mediates community iron acquisition in the global ocean. *ISME J.* 16, 358–369. doi: 10.1038/s41396-021-01065-y

Martens, J.-H., Barg, H., Warren, M., and Jahn, D. (2002). Microbial production of vitamin B12. *Appl. Microbiol. Biotechnol.* 58, 275–285. doi: 10.1007/s00253-001-0902-7

Mas, A., Jamshidi, S., Lagadeuc, Y., Eveillard, D., and Vandenkoornhuyse, P. (2016). Beyond the Black Queen Hypothesis. *ISME J.* 10, 2085–2091. doi: 10.1038/ismej.2016.22

Milici, M., Deng, Z.-L., Tomasch, J., Decelle, J., Wos-Oxley, M. L., Wang, H., et al. (2016). Co-occurrence Analysis of Microbial Taxa in the Atlantic Ocean Reveals High Connectivity in the Free-Living Bacterioplankton. *Front. Microbiol.* 7:649. doi: 10.3389/fmicb.2016.00649

Mills, M. M., Ridame, C., Davey, M., La Roche, J., and Geider, R. J. (2004). Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature* 429, 292–294. doi: 10.1038/nature02550

Moisander, P. H., Zhang, R., Boyle, E. A., Hewson, I., Montoya, J. P., and Zehr, J. P. (2012). Analogous nutrient limitations in unicellular diazotrophs and Prochlorococcus in the South Pacific Ocean. *ISME J.* 6, 733–744. doi: 10.1038/ismej.2011.152

Moore, C. M., Mills, M. M., Achterberg, E. P., Geider, R. J., LaRoche, J., Lucas, M. I., et al. (2009). Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. *Nat. Geosci.* 2, 867–871. doi: 10.1038/ngeo667

Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., et al. (2013). Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* 6, 701–710. doi: 10.1038/ngeo1765

Moore, J. K., Fu, W., Primeau, F., Britten, G. L., Lindsay, K., Long, M., et al. (2018). Sustained climate warming drives declining marine biological productivity. *Science* 359, 1139–1143. doi: 10.1126/science.aao6379

Moore, L. R., Post, A. F., Rocap, G., and Chisholm, S. W. (2002). Utilization of different nitrogen sources by the marine cyanobacteria Prochlorococcus and Synechococcus. *Limnol. Oceanogr.* 47, 989–996. doi: 10.4319/lo.2002.47.4.0989

Morris, J. J., Lenski, R. E., and Zinser, E. R. (2012). The Black Queen Hypothesis: evolution of Dependencies through Adaptive Gene Loss. *mBio* 3, e00036–12. doi: 10.1128/mBio.00036-12

Muñoz-Marín, M., del, C., Shilova, I. N., Shi, T., Farnelid, H., Cabello, A. M., et al. (2018). The Transcriptional Cycle Is Suited to Daytime N2 Fixation in the Unicellular Cyanobacterium " Candidatus Atelocyanobacterium thalassa" (UCYN-A). *mBio* 10, e02495–18. doi: 10.1128/mBio.02495-18

Noble, A. E., Lamborg, C. H., Ohnemus, D. C., Lam, P. J., Goepfert, T. J., Measures, C. I., et al. (2012). Basin-scale inputs of cobalt, iron, and manganese from the Benguela-Angola front to the South Atlantic Ocean. *Limnol. Oceanogr.* 57, 989–1010. doi: 10.4319/lo.2012.57.4.0989

Pacheco, A. R., Moel, M., and Segrè, D. (2019). Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nat. Commun.* 10:103. doi: 10.1038/s41467-018-07946-9

Paerl, R. W., Bertrand, E. M., Rowland, E., Schatt, P., Mehiri, M., Niehaus, T. D., et al. (2018a). Carboxythiazole is a key microbial nutrient currency and critical component of thiamin biosynthesis. *Sci. Rep.* 8:5940. doi: 10.1038/s41598-018-24321-2

Paerl, R. W., Sundh, J., Tan, D., Svenningsen, S. L., Hylander, S., Pinhassi, J., et al. (2018b). Prevalent reliance of bacterioplankton on exogenous vitamin B1 and precursor availability. *Proc. Natl. Acad. Sci. U. S. A.* 115, E10447–E10456. doi: 10.1073/pnas.1806425115

Paerl, R. W., Bouget, F.-Y., Lozano, J.-C., Vergé, V., Schatt, P., Allen, E. E., et al. (2017). Use of plankton-derived vitamin B1 precursors, especially thiazole-related precursor, by key marine picoeukaryotic phytoplankton. *ISME J.* 11, 753–765. doi: 10.1038/ismej.2016.145

Panzeca, C., Beck, A. J., Leblanc, K., Taylor, G. T., Hutchins, D. A., and Sañudo-Wilhelmy, S. A. (2008). Potential cobalt limitation of vitamin B12 synthesis in

the North Atlantic Ocean. *Glob. Biogeochem. Cycles* 22:GB2029. doi: 10.1029/2007GB003124

Panzeca, C., Tovar-Sanchez, A., Agustí, S., Reche, I., Duarte, C. M., Taylor, G. T., et al. (2006). B vitamins as regulators of phytoplankton dynamics. *Eos Trans. Am. Geophys. Union* 87, 593–596. doi: 10.1029/2006EO520001

Rahav, E., Giannetto, M. J., and Bar-Zeev, E. (2016). Contribution of mono and polysaccharides to heterotrophic N2 fixation at the eastern Mediterranean coastline. *Sci. Rep.* 6:27858. doi: 10.1038/srep27858

Rahav, E., Herut, B., Mulholland, M., Belkin, N., Elifantz, H., and Berman-Frank, I. (2015). Heterotrophic and autotrophic contribution to dinitrogen fixation in the Gulf of Aqaba. *Mar. Ecol. Prog. Ser.* 522, 67–77. doi: 10.3354/meps11143

Ramanan, R., Kim, B.-H., Cho, D.-H., Oh, H.-M., and Kim, H.-S. (2016). Algae–bacteria interactions: evolution, ecology and emerging applications. *Biotechnol. Adv.* 34, 14–29. doi: 10.1016/j.biotechadv.2015.12.003

Saito, M. A., Dorsk, A., Post, A. F., McIlvin, M. R., Rappé, M. S., DiTullio, G. R., et al. (2015). Needles in the blue sea: sub-species specificity in targeted protein biomarker analyses within the vast oceanic microbial metaproteome. *Proteomics* 15, 3521–3531. doi: 10.1002/pmic.201400630

Saito, M. A., Goepfert, T. J., and Ritt, J. T. (2008). Some thoughts on the concept of colimitation: three definitions and the importance of bioavailability. *Limnol. Oceanogr.* 53, 276–290. doi: 10.4319/lo.2008.53.1.0276

Sanchez, N., Brown, E. A., Olsen, Y., Vadstein, O., Iriarte, J. L., Gonzalez, H. E., et al. (2018). Effect of Siderophore on Iron Availability in a Diatom and a Dinoflagellate Species: contrasting Response in Associated Bacteria. *Front. Mar. Sci.* 5:118. doi: 10.3389/fmars.2018.00118

Sañudo-Wilhelmy, S. A., Gobler, C. J., Okbamichael, M., and Taylor, G. T. (2006). Regulation of phytoplankton dynamics by vitamin B12. *Geophys. Res. Lett.* 33, 1–4. doi: 10.1029/2005GL025046

Sañudo-Wilhelmy, S. A., Gómez-Consarnau, L., Suffridge, C., and Webb, E. A. (2014). The Role of B Vitamins in Marine Biogeochemistry. *Annu. Rev. Mar. Sci.* 6, 339–367. doi: 10.1146/annurev-marine-120710-100912

Sañudo-Wilhelmy, S. A., Kustka, A. B., Gobler, C. J., Hutchins, D. A., Yang, M., Lwiza, K., et al. (2001). Phosphorus limitation of nitrogen fixation by Trichodesmium in the central Atlantic Ocean. *Nature* 411, 66–69. doi: 10.1038/35075041

Sarmiento, J. L., Gruber, N., Brzezinski, M. A., and Dunne, J. P. (2004a). High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature* 427, 56–60. doi: 10.1038/nature02127

Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., et al. (2004b). Response of ocean ecosystems to climate warming. *Glob. Biogeochem. Cycles* 18:GB3003. doi: 10.1029/2003GB002134

Seyedsayamdost, M. R., Case, R. J., Kolter, R., and Clardy, J. (2011). The Jekyll-and-Hyde chemistry of Phaeobacter gallaeciensis. *Nat. Chem.* 3, 331–335. doi: 10.1038/nchem.1002

Seymour, J. R., Amin, S. A., Raina, J.-B., and Stocker, R. (2017). Zooming in on the phycosphere: the ecological interface for phytoplankton-bacteria relationships. *Nat. Microbiol.* 2, 17065. doi: 10.1038/nmicrobiol.2017.65

Shelton, A. N., Seth, E. C., Mok, K. C., Han, A. W., Jackson, S. N., Haft, D. R., et al. (2019). Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *ISME J.* 13, 789–804. doi: 10.1038/s41396-018-0304-9

Shibl, A. A., Isaac, A., Ochsenkühn, M. A., Cárdenas, A., Fei, C., Behringer, G., et al. (2020). Diatom modulation of select bacteria through use of two unique secondary metabolites. *Proc. Natl. Acad. Sci. U. S. A.* 117, 27445–27455. doi: 10.1073/pnas.2012088117

Soto, M. A., Desai, D, LaRoche, J, and Bertrand, E. M. (in review). Cobalamin producers and prokaryotic consumers in the Northwest Atlantic. *Environ. Microbiol.*

Smith, N. W., Shorten, P. R., Altermann, E., Roy, N. C., and McNabb, W. C. (2019). The Classification and Evolution of Bacterial Cross-Feeding. *Front. Ecol. Evol.* 7:153. doi: 10.3389/fevo.2019.00153

Sohm, J. A., Webb, E. A., and Capone, D. G. (2011). Emerging patterns of marine nitrogen fixation. *Nat. Rev. Microbiol.* 9, 499–508. doi: 10.1038/nrmicro2594

Sokolovskaya, O. M., Shelton, A. N., and Taga, M. E. (2020). Sharing vitamins: cobamides unveil microbial interactions. *Science* 369:eaba0165. doi: 10.1126/science.aba0165

Sutak, R., Camadro, J.-M., and Lesuisse, E. (2020). Iron Uptake Mechanisms in Marine Phytoplankton. *Front. Microbiol.* 11:566691. doi: 10.3389/fmicb.2020.566691

Tagliabue, A., Bowie, A. R., Boyd, P. W., Buck, K. N., Johnson, K. S., and Saito, M. A. (2017). The integral role of iron in ocean biogeochemistry. *Nature* 543, 51–59. doi: 10.1038/nature21058

Tagliabue, A., Hawco, N. J., Bundy, R. M., Landing, W. M., Milne, A., Morton, P. L., et al. (2018). The Role of External Inputs and Internal Cycling in Shaping the Global Ocean Cobalt Distribution: insights From the First Cobalt Biogeochemical Model. *Glob. Biogeochem. Cycles* 32, 594–616. doi: 10.1002/2017GB005830

Talling, J. F. (1955). The Relative Growth Rates of Three Plankton Diatoms in Relation to Underwater Radiation and Temperature. *Ann. Bot.* 19, 329–341. doi: 10.1093/oxfordjournals.aob.a083432

Thompson, A. W., Foster, R. A., Krupke, A., Carter, B. J., Musat, N., Vaulot, D., et al. (2012). Unicellular Cyanobacterium Symbiotic with a Single-Celled Eukaryotic Alga. *Science* 337, 1546–1550. doi: 10.1126/science.1222700

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., et al. (2018). Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* 11, 27–37. doi: 10.1038/s41561-017-0028-x

Twining, B. S., and Baines, S. B. (2013). The Trace Metal Composition of Marine Phytoplankton. *Annu. Rev. Mar. Sci.* 5, 191–215. doi: 10.1146/annurev-marine-121211-172322

Twining, B. S., Antipova, O., Chappell, P. D., Cohen, N. R., Jacquot, J. E., Mann, E. L., et al. (2021). Taxonomic and nutrient controls on phytoplankton iron quotas in the ocean. *Limnol. Oceanogr. Lett.* 6, 96–106. doi: 10.1002/lol2.10179

Ustick, L. J., Larkin, A. A., Garcia, C. A., Garcia, N. S., Brock, M. L., Lee, J. A., et al. (2021). Metagenomic analysis reveals global-scale patterns of ocean nutrient limitation. *Science* 372, 287–291. doi: 10.1126/science.abe6301

Van de Waal, D. B., and Litchman, E. (2020). Multiple global change stressor effects on phytoplankton nutrient acquisition in a future ocean. *Philos. Trans. R. Soc. B* 375:20190706. doi: 10.1098/rstb.2019.0706

Ward, B. A., Dutkiewicz, S., Jahn, O., and Follows, M. J. (2012). A size-structured food-web model for the global ocean. *Limnol. Oceanogr.* 57, 1877–1891. doi: 10.4319/lo.2012.57.6.1877

Wu, M., McCain, J. S. P., Rowland, E., Middag, R., Sandgren, M., Allen, A. E., et al. (2019). Manganese and iron deficiency in Southern Ocean Phaeocystis antarctica populations revealed through taxon-specific protein indicators. *Nat. Commun.* 10:3582. doi: 10.1038/s41467-019-11426-z

Zehr, J. P., and Capone, D. G. (2020). Changing perspectives in marine nitrogen fixation. *Science* 368:eaay9514. doi: 10.1126/science.aay9514

Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D. R., Bork, P., and Patil, K. R. (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* 112, 6449–6454. doi: 10.1073/pnas.1421834112

Zomorrodi, A. R., and Segrè, D. (2017). Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. *Nat. Commun.* 8:1563. doi: 10.1038/s41467-017-01407-5

Defining the Realized Niche of the Two Major Clades of *Trichodesmium*: A Study on the West Florida Shelf

Kristina A. Confesor[1], Corday R. Selden[1,2], Kimberly E. Powell[1], Laura A. Donahue[3], Travis Mellett[4,5], Salvatore Caprara[4], Angela N. Knapp[6], Kristen N. Buck[4] and P. Dreux Chappell[1*]

[1] Old Dominion University, Department of Ocean and Earth Sciences, Norfolk, VA, United States, [2] Rutgers University, Department of Marine and Coastal Sciences, New Brunswick, NJ, United States, [3] Department of Biology Haverford College, Haverford, PA, United States, [4] University of South Florida, College of Marine Science, St. Petersburg, FL, United States, [5] University of Washington, School of Oceanography, Seattle, WA, United States, [6] Florida State University, Department of Earth, Ocean, and Atmospheric Sciences, Tallahassee, FL, United States

The cyanobacterium *Trichodesmium* plays an essential role supporting ocean productivity by relieving nitrogen limitation *via* dinitrogen (N_2) fixation. The two common *Trichodesmium* clades, *T. erythraeum* and *T. thiebautii*, are both observed in waters along the West Florida Shelf (WFS). We hypothesized that these taxa occupy distinct realized niches, where *T. thiebautii* is the more oceanic clade. Samples for DNA and water chemistry analyses were collected on three separate WFS expeditions (2015, 2018, and 2019) spanning multiple seasons; abundances of the single copy housekeeping gene *rnpB* from both clades were enumerated *via* quantitative PCR. We conducted a suite of statistical analyses to assess *Trichodesmium* clade abundances in the context of the physicochemical data. We observed a consistent coastal vs. open ocean separation of the two clades: *T. erythraeum* was found in shallow waters where the concentrations of dissolved iron (dFe) and the groundwater tracer Ba were significantly higher, while *T. thiebautii* abundance was positively correlated with water column depth. The Loop Current intrusion in 2015 with entrained Missisippi River water brought higher dFe and elevated abundance of both clades offshore of the 50 m isobath, suggesting that both clades are subject to Fe limitation on the outer shelf. Whereas, previous work has observed that *T. thiebautii* is more abundant than *T. erythraeum* in open ocean surface waters, this is the first study to examine *Trichodesmium* niche differentiation in a coastal environment. Understanding the environmental niches of these two key taxa bears important implications for their contributions to global nitrogen and carbon cycling and their response to global climate change.

Keywords: *Trichodesmium*, **diazotrophs, niche separation, coastal/ocean separation, West Florida Shelf, groundwater**

INTRODUCTION

Trichodesmium is an important genus of marine cyanobacteria that converts dinitrogen gas (N_2) into bioaccessible ammonia *via* N_2 fixation (Capone et al., 2005). Once fixed, new dissolved nitrogen (N) can be released, fueling local food webs (Capone, 2001) and, ultimately, playing an essential role in total ocean productivity by relieving N limitation (Carpenter and Capone, 2008). Indeed, *Trichodesmium*'s new N inputs are estimated at 60-80 Tg N yr (Bergman et al., 2013)—a large fraction of the ocean's total estimated N_2 fixation rate of 100-200 Tg N yr^{-1} (Karl et al., 2002). Thus, *Trichodesmium*'s N contributions to the biosphere impacts both the global N and carbon (C) cycles, and is thought to influence global carbon dioxide (CO_2) sequestration on hundred-to-thousand-year timescales (Gruber and Sarmiento, 1997; Haselkorn and Buikema, 1997; Falkowski, 1998).

The *Trichodesmium* genus is composed of six different species, grouped into two major clades (Hynes et al., 2011; Rouco et al., 2014). *Trichodesmium thiebautii* and *Trichodesmium erythraeum* are considered the representative species of the two different clades (Rouco et al., 2014), with *T. thiebautii* being the more abundant clade in open ocean regions (Hynes et al., 2009; Chappell et al., 2012; Rouco et al., 2014). The groups differ in their ecophysiology, including in their optimal growth temperatures and iron (Fe) stress responses (Breitbarth et al., 2007; Chappell and Webb, 2010). Nevertheless, most culture-based studies rely on IMS101, a lab strain of *T. erythraeum* first isolated off the North Carolina coast roughly 30 years ago (Prufert-Bebout et al., 1993). Work with IMS101 has suggested that N_2 fixation rates increase commensurately with atmospheric CO_2 (Hutchins et al., 2007; Levitan et al., 2007); however, the growth and N_2 fixation responses of *T. thiebautii* and *T. erythraeum* differ under rising CO_2 concentrations (Hutchins et al., 2013). Understanding how the environmental sensitivities and realized niches of the two clades differ is essential to leveraging culture-based work and predicting their response to climate change.

Trichodesmium is found at tropical latitudes of the ocean (Capone et al., 2005), including along the West Florida Shelf (WFS) (Lenes et al., 2001). Surface waters beyond the 50 m isobath offshore of Tampa Bay bear low dissolved Fe (dFe) concentrations (Lenes et al., 2001; Mellett and Buck, 2020) that are highly influenced by seasonal dust deposition (Mellett and Buck, 2020). Siderophore production in open ocean *Trichodesmium* colonies may enable the uptake of dust Fe *via* siderophore-mediated dissolution, making the dust Fe readily available to species with this capability (Basu et al., 2019). However, the *T. erythraeum* strain IMS101 is not known to produce siderophores (Basu et al., 2019), implying that differences in the capability to produce siderophores in *Trichodesmium* communities may lead to niche specialization related to dust deposition.

Low Fe has been shown to limit both N_2 fixation and growth of *Trichodesmium* (Berman-Frank et al., 2001; Kustka et al., 2003). *Trichodesmium* in the open ocean can be co-limited by phosphorus (P) and Fe (Mills et al., 2004; Basu et al., 2019), and has been shown to be more strongly P-limited in North Atlantic waters in comparison to the North Pacific (Sañudo-Wilhelmy et al., 2001; Sohm et al., 2008). Modeling work has shown that an increase in Fe availability can lead to N_2-fixation by *Trichodesmium* becoming P-limited rather than Fe-limited (Ye et al., 2012). Recently it has been found that some *Trichodesmium* species appear to not fix N at all, presumably as a result of evolutionary adaptations to resource availability (Delmont, 2021). Little is known about the physiology of the non-N_2-fixing *Trichodesmium* species, which have been identified solely through metagenomic analysis. Macronutrient concentrations at the WFS are negligible, and well below the half-saturation constant for phosphate uptake by *Trichodesmium* (Lenes et al., 2008), suggesting these trace macronutrient concentrations are likely to have little impact on clade distribution or inhibition of N_2 fixation by nitrate (Knapp et al., 2012).

Previous work examining clade distributions in the open ocean showed evidence of niche separation due to potential resource competition, where *T. thiebautii* distributions extended from the surface down to ≥80 m, while *T. erythraeum* was only observed in the mixed layer of the ocean (Rouco et al., 2016). Multiple studies have found that the *T. thiebautii* clade is more abundant than the *T. erythraeum* clade in open ocean waters (Hynes et al., 2009; Chappell et al., 2012; Rouco et al., 2014; Rouco et al., 2016). Whereas, N_2 fixation rates have been shown to positively correlate with dFe concentrations in *T. thiebautii*-dominated open ocean waters (Chappell et al., 2012), the niche preferences of the two clades and the environmental controls on their abundance remain poorly constrained in coastal areas. Here, we assess the realized niches of the primary *Trichodesmium* clades (*T. erythraeum* and *T. thiebautii*) along the WFS by comparing the abundance of a housekeeping gene (*rnpB*) diagnostic of clade identity to physical and chemical variables. By elucidating clade-level niche preferences, this work provides insight into the sensitivity of this biogeochemically important genus to climate forcings.

METHODS AND MATERIALS

Hydrographic Data

Samples were collected on the R/V Weatherbird II from June 18-21 in 2015 and from April 9-12 in 2019, as well as on the R/V Hogarth from February 27- March 2 in 2018 along the WFS (**Figure 1** and **Supplementary Image 1**). Surface salinity and temperature measurements were collected on all cruises using the ships' flow-through hydrographic systems (SeaBird). Hydrographic variables such as sea surface salinity and sea surface height (SSH) (**Figure 1**) for the calendar day and region that each cruise sampled were obtained using Daily CMEMS GLORYS12V1 global reanalysis (0.083° x 0.083° resolution) (Lien et al., 2021). SSH was used to outline the edge of the Loop Current for each sampling year, while low salinity indicated waters from the Mississippi river were entrained near the edge of the Loop Current in each sampling year. Samples for chemical analyses were obtained from the surface mixed layer as described below.
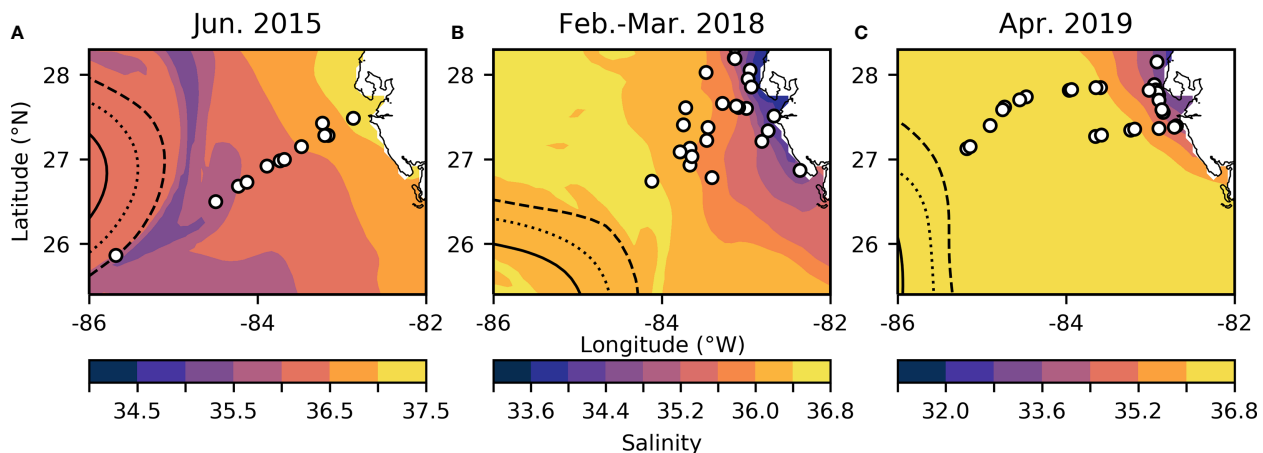
FIGURE 1 | WFS study region for June 2015 **(A)**, February-March 2018 **(B)**, and April 2019 **(C)**. Overlying sea surface salinity (note different color scales for each panel) extracted from the GLORYS12V1 reanalysis product (Lien et al., 2021). White dots indicate stations. Dashed, dotted, and solid lines indicate a SSH of 0 m, 0.2 m, and 0.4 m respectively, outlining the edge of the Loop Current.

Macronutrient and Trace Metal Analyses

Surface (~2 m) dFe, barium (Ba), and macronutrient samples were collected using a trace metal clean "towfish" system (Mellett and Buck, 2020). The specific sampling system and analytical details for trace metal and macronutrient concentration measurements are described in Mellett and Buck (2020). Briefly, trace metal samples were collected with a towfish sampling system and in-line filtered through 0.2 μm Acropak capsule filters into acid-cleaned low-density polyethylene bottles (dFe and Ba) or polypropylene tubes (silicic acid, nitrate+nitrite, and soluble reactive phosphorus, Si, N+N, and PO_4, respectively); samples for dissolved trace metals were acidified to pH ~1.8 with ultrapure hydrochloric acid (Optima HCl, Fisher; final concentration 0.024 M). Macronutrient samples were stored frozen (-20 °C) until analysis on a Lachat 8500 QuickChem analyzer using colorimetric methods (Parsons et al., 1984). Phosphate (PO_4) and N+N concentrations were mostly below the detection limit (**Supplementary Table 1**), which is consistent with long term nutrient concentration measurements from the region (Heil et al., 2014). We note that the methodology we used for analyzing macronutrients does not incorporate techniques optimized for low-level concentration analysis, so PO_4 and N+N were excluded in later statistical analyses. Dissolved Fe samples were UV-oxidized, preconcentrated onto a Nobias PA1 chelating resin using an automated seaFAST-pico system, and the resulting eluents analyzed using standard addition on a Thermo Scientific Element XR Inductively Coupled Plasma Mass Spectrometer in medium resolution and counting mode at the University of South Florida (Hollister et al., 2020). Dissolved Ba was measured after a 1:50 dilution in 5% ultrapure nitric acid and quantified by standard addition within 24 hours of the dilution preparation directly on the Element XR.

DNA Sample Collection

Surface DNA samples were collected either *via* the "towfish" system (Mellett and Buck, 2020) or from the ship's Niskin bottle rosette and filtered onto 0.2 μm polyethersulfone (PES) filters using a MasterFlex® peristaltic pumping system (Avantor™, Pennsylvania, USA). Up to 4 L of water was filtered, with lower volumes of water (1-2 L) collected in 2015. In 2018 and 2019, lower volumes (~1 L) were sometimes collected in the very nearshore waters because of filter clogging. In 2015, flat PES filters were stored with Qiagen® RLT Plus buffer (Qiagen, Germany), flash frozen in liquid N_2 at sea, and then stored at -80°C until analysis. On the 2018 and 2019 cruises, Sterivex® cartridge filters (MilliporeSigma, Burlington, MA) were preserved in RNA*later*™ (Life Technologies, Carlsbad, CA) at 4°C for approximately 12-18 hours before being secured in a dry shipper for transport and then storage at -80°C until analysis.

DNA Extraction

Filters were extracted using the Allprep RNA/DNA Mini Kit (Qiagen, Germany) following the manufacturer's protocol with the addition of bead-beating and homogenization using the QIAshredder® column (Qiagen, Germany). All extractions were performed in a HEPA filtered UV sterilized AC600 PCR workstation (AirClean® Systems, Creedmore, NC). For samples collected on Sterivex® filters, ethanol cleaned PVC pipe cutters were used to open the Sterivex® tube. The filter was then cut out with autoclave-sterilized scalpel blades into two parts, then placed into RLT+ Buffer tubes. DNA samples were eluted in 80 μL buffer EB and stored at -80°C until qPCR analysis.

Quantitative Polymerase Chain Reaction

An established qPCR procedure was used for distinguishing between the two primary *Trichodesmium* clades (Rouco et al., 2014).

Clade-specific primers targeted the single-copy housekeeping gene *rnpB*, which encodes for ribonuclease P and is frequently used in qPCR studies of cyanobacteria (Chappell and Webb, 2010; Rouco et al., 2014). Quantitative PCR amplification was performed using a Step One Plus Real Time PCR thermal cycler (Life Technologies, Carlsbad, CA). The only deviation from the original protocol (Chappell and Webb, 2010; Rouco et al., 2014) was that clade-specific gene abundances were absolutely quantified using standard curves prepared from *T. erythraeum* and *T. thiebautii rnpB* plasmids generated and quantified following the protocol of Chappell and Webb (2010). All standards, no-template controls, and samples were measured in triplicate with PowerSYBR® Green Mastermix fluorescent dye (Life Technologies, Carlsbad, CA) using 96-well plates. No-template control wells, used to monitor contamination, contained the qPCR master-mix, clade-specific primers, and Rnase-free water. Samples were run as 1:10 dilutions of the original extracted DNA with 2 μL of sample in a 20 μL reaction. Melt curves were incorporated to ensure that single products were successfully amplified (Rouco et al., 2014) and all qPCR efficiencies were above 93%.

Standard curves were created in triplicate with known gene abundances of serially diluted plasmid standards of each clade. The critical threshold, the point at which fluorescence intensity crosses the detectible level and corresponds to the initial abundance of DNA in samples (Kralik and Ricchi, 2017), was determined by the Step One Plus software and values from the triplicate qPCR reactions for each unknown were averaged and compared with the standard curve. Limits of detection and limits of quantification were determined for each sample as described in Selden et al. (2021). In subsequent statistical analyses, we chose a conservative approach whereby samples that were below detection limit (BDL, < 3 gene copies/PCR) were given the value of zero and samples that were detectable but unquantifiable (BQL, <10 gene copies/PCR) were given the value that was the limit of detection. Effective limit of quantification (ELOQ) are defined as 2000-3125 copies/L in 2015, 948-5000 copies/L in 2018, and 625 – 2000 copies/L in 2019.

Statistical Analyses

Kolmogorov-Smirnov tests were performed to test residuals for normality; however, residuals did not follow a normal distribution. Consequently, non-parametric tests (Spearman's Correlation and Kruskal-Wallis) using MATLAB, R2020a and ordination statistics [canonical correspondence analysis (CCA)] using the VEGAN package in RStudio (Dixon, 2003), were employed to compare clade abundances and correlations with environmental data. CCA was done with only one constrained axis. Environmental data evaluated for correlation with *Trichodesmium* spp. clades included dFe, Si, and Ba concentrations, as well as salinity and water column depth. Water column depths were obtained from ETOPO1 (bedrock) bathymetry data grid extraction (Amante and Eakins, 2009).

RESULTS AND DISCUSSION

Physiochemical Data

On the WFS, surface dFe concentrations were highest near the coast (mean = 3.84 ± 3.14 nM, **Table 1**) and declined towards the shelf-break where waters were more oceanic (mean = 0.70 ± 0.62 nM, **Table 1**, **Figures 2A–C** and **Supplementary Table 1**). Areas near the shelf-break and west of the 50 m isobath are defined as offshore, while areas on the mid-shelf with bottom depths <50 m are defined as inshore (**Figure 2**). Concentrations of dFe were negatively correlated with water column depth (Spearman: 5.07×10^{-9}, Rho = -0.736). The dFe concentrations were likely elevated in surface waters on the shelf due to proximity to continental sources, including riverine and submarine groundwater discharge (SGD), and shallow water column mixing with shelf sediments.

We used Ba and Si to trace continental runoff sources, including inputs from rivers and SGD (Shaw et al., 1998; Oehler et al., 2019). Continental runoff can deliver essential nutrients to macro- and microbiota living in coastal areas (Charette et al., 2013; Wang et al., 2018), and SGD in particular has been previously observed to be important on the inner WFS (<50 m) (Hu et al., 2006). As expected, Ba and Si were more abundant at stations with shallower bottom depths (Ba-Spearman: 1.38×10^{-12}, Rho=-0.817, Si- Spearman: 1.48×10^{-8}, Rho=-0.711, **Table 1**) and were also positively correlated with one another (Spearman: 3.48×10^{-07}, Rho=0.659), supporting the hypothesis that continental runoff was an important source of these elements on the inner shelf at the time of this study.

Based on salinity and SSH, we collected samples at or near the Loop Current edge on all three cruises (**Figure 1**) (Mellett and Buck, 2020). The Loop Current originates from warm Caribbean waters, enters the Gulf of Mexico through the Yucatan Channel and exits *via* the Florida Strait (Morrison et al., 1983). This water mass has been shown to be low in trace metals and macronutrients, but serves as a physical vector for bringing elevated dFe concentrations to the outer WFS region, by entrainment of Mississippi River plume water along the edge of the Loop that is advected south to the WFS (Mellett and Buck,

TABLE 1 | Average values and standard deviations of physiochemical data from 2019 (first two columns) and from all three cruises combined (last two columns).

Measurement	Offshore 2019	Inshore 2019	Offshore Total	Inshore Total
dFe (nM)	0.400 ± 0.15	2.81 ± 1.50	0.70 ± 0.62	3.84 ± 3.14
Ba (nM)	46.2 ± 5.7	59.1 ± 7.0	48.9 ± 3.8	61.1 ± 9.7
Si (uM)	1.07 ± 0.20	2.02 ± 1.08	0.95 ± 0.65	2.31 ± 1.76
Salinity	36.4 ± 0.1	35.0 ± 0.8	36.5 ± 0.2	35.4 ± 0.9

Values from 2015 & 2018 are reported in Mellet and Buck (Mellett and Buck, 2020). Offshore is defined as west of the 50 m isobath, and inshore as all measurements east of the 50 m isobath.
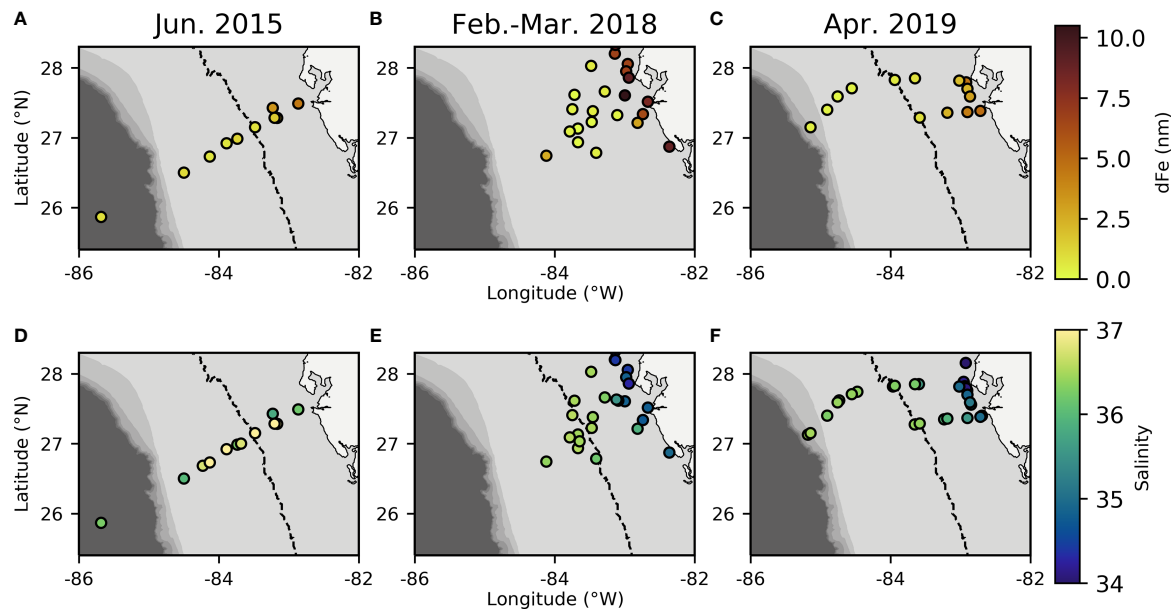
FIGURE 2 | Dissolved Fe concentration **(A–C)** and salinity **(D–F)** measurements from each sampling year plotted over bathymetry extracted from ETOPO1 bedrock (Amante and Eakins, 2009). Dashed line indicates the 50m isobath which is used to distinguish between offshore and inshore samples and indicates where there are known SGD inputs inshore of this area (Hu et al., 2006).

2020). We observed evidence of this entrainment across all three of our cruises (**Figures 1**, **2**), though the effect on offshore dFe was most pronounced in 2018 against the backdrop of much lower wintertime surface dFe (**Figure 2B**).

Trichodesmium Clade Gene Abundance in Coastal vs. Offshore Samples

Our results support a distinct niche distribution between the two *Trichodesmium* clades along the WFS, where *T. erythraeum* dominates the inner shelf (Kruskal-Wallis: p=0.0242, Chisq=5.08, df=66, std.dev=200,427) and *T. thiebautii* dominates the outer shelf (Kruskal-Wallis: p=0.000439, Chi-sq=12.4, df=66, std.dev=214,000) (**Figure 3**). Absolute gene abundances of *T. thiebautii* consistently averaged between 1.0×10^5-1.0×10^6 gene copies/L offshore of the 50 m isobath (**Figures 3A–C**), comparable with the average cells per L seen in prior studies at further offshore surface waters of the Atlantic (Rouco et al., 2014). In many inshore stations, *T. thiebautii* was below detection limits (**Supplementary Table 1**). *T. erythraeum*, on the other hand, had average ranges of 1.0×10^5-1.0×10^6 gene copies/L inshore of the 50 m isobath (**Figures 3G–I**). In many offshore stations, *T. erythraeum* was not detectable (**Supplementary Table 1**). Notably, *T. erythraeum* was not detected inshore during the 2015 cruise (**Figure 3G**), but we ascribe this anomaly to a higher detection limit for that specific sampling event that precluded quantification (**Figure 3G**, **Supplementary Table 1**).

Altogether, our results support prior studies suggesting that *T. thiebautii* is the more oceanic clade (Chappell et al., 2012). It is important to note that on each cruise, stations sampled close to the Loop Current had elevated abundances of both *Trichodesmium*

clades, which may have contributed to a higher mean abundance at offshore stations in 2015 (**Figure 3**). This may have resulted from the edge of the Loop Current delivering dFe and other nutrients to the shelf from entrainment of Mississippi River plume water as well as simultaneous dust deposition offshore in this season. The known SGD inputs inshore of the 50 m isobath (Hu et al., 2006) may also contribute to the differences in gene abundances and perceived environmental niches.

Before evaluating physicochemical drivers of clade distributions, we focused on water column depth as a way to distinguish between inshore and offshore samples. Water column depth and gene abundances were compared using multiple non-parametric tests (Spearman correlation and Kruskal-Wallis tests) to ensure that results were consistent regardless of analytical tool. Water column depth and *T. thiebautii* abundance were positively correlated (Spearman: p=0.0009, Rho=0.4063) while bottom depth and *T. erythraeum* abundance were negatively correlated (Spearman: p=0.0123, Rho=-0.3114). *T. erythraeum* was most often observed in waters <50 m depth. Based on these initial analyses, we distinguished stations deeper than 50 m as 'offshore' and those shallower as 'inshore'. This distinction was used in subsequent statistical analyses. Additionally, prior work has indicated that WFS waters <50 m are more influenced by SGD (Hu et al., 2006), which is consistent with the elevated dFe and groundwater tracers observed inshore of the 50 m isobath (**Figures 2**, **3**, **Table 1**).

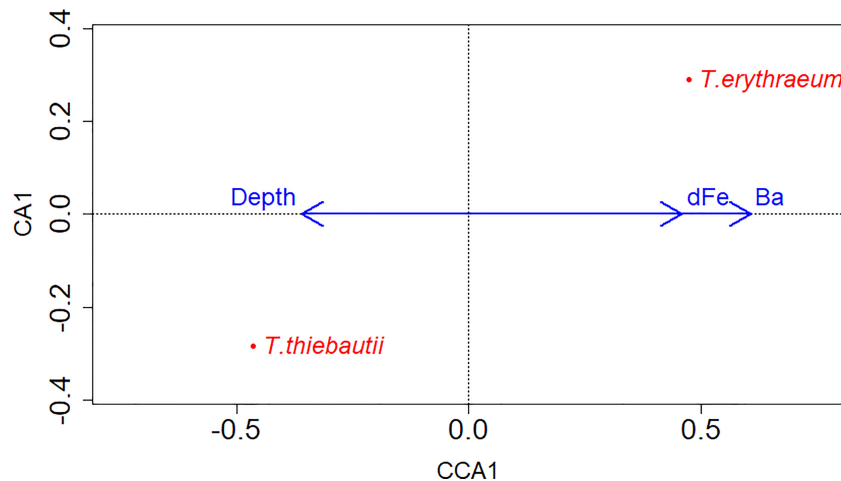## Gene Abundance Correlations With Physiochemical Data

To identify potential drivers of niche differentiation between the clades that might explain differences between inshore and offshore
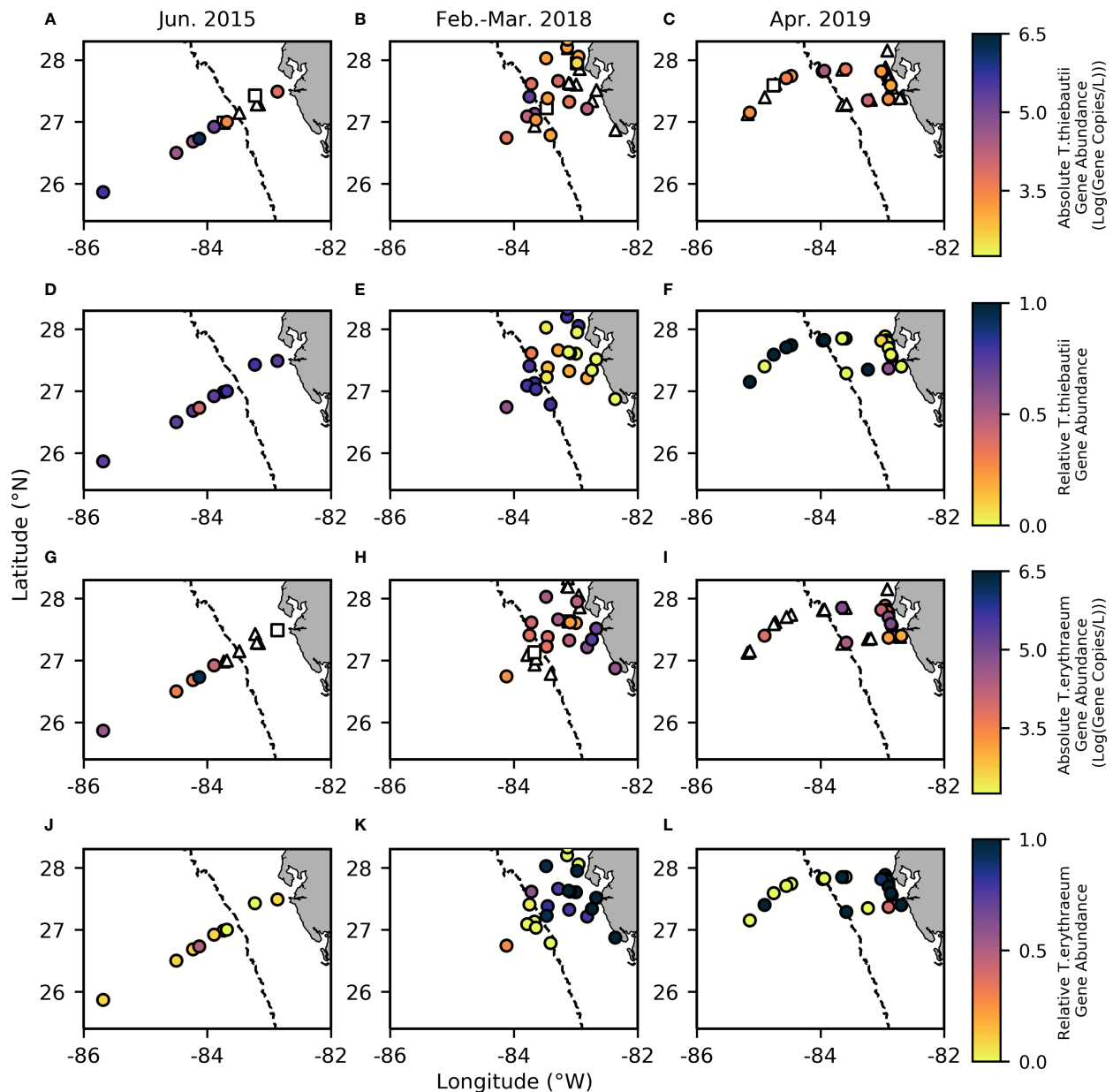
FIGURE 3 | Absolute gene abundance for *T. thiebautii* **(A–C)** and *T. erythraeum* **(G–I)**, as well as relative gene abundance for *T. thiebautii* **(D–F)** and *T. erythraeum* **(J–L)** with contours reflecting bathymetry extracted from ETOPO1 bedrock (Amante and Eakins, 2009). In all panels, the dashed line indicates the 50 m isobath, which is used to distinguish between offshore and inshore samples and indicates where there are known SGD inputs inshore of this area (Hu et al., 2006). Triangles represent *Trichodesmium* spp. abundances below the detection limit (BDL), while squares represent data that is below the quantification limit (BQL). (Effective limit of quantification (ELOQ): 2000-3125 copies/L in 2015, 948-5000 copies/L in 2018, and 625 – 2000 copies/L in 2019).

distributions, we examined correlations between each clade's gene abundance with a number of physicochemical variables. Spearman correlations were employed to evaluate clade abundances with physiochemical data, while CCA was used to evaluate associations between both clades and environmental parameters. Using CCA, significant correlations were found between the two clades and dFe and Ba concentrations, as well as water column depth (**Figure 4**). Significant positive correlations were found between

salinity and *T. thiebautii* abundance (Spearman Correlation (right tail): p= 0.0003, Rho=0.4079), consistent with this being the more oceanic clade as salinity increased offshore (**Figures 2D–F**). In the CCA analysis, there was no significant correlation between the abundance of both clades with salinity; rather, water column depth was significantly correlated with the abundance of both the clades (**Figure 4**). This is consistent with *T. thiebautii* being the more oceanic clade as water column depth increased, while

FIGURE 4 | CCA model of gene abundances of both clades in relation to physiochemical data and water column depth (extracted from ETOPO1 bedrock (Amante and Eakins, 2009). Clades are labeled on the data points, while blue arrows show the x-axis directionality of physiochemical data and environmental parameters. There was only one constrained axis as there were only two clades of interest. Significant environmental parameters and the two clades included dFe (nM) (p-value= 0.001, Chi-square=0.116), Water column depth (m) (p-value=0.001,Chi-square=0.058), and Ba (nM) (p-value=0.001, Chi-square=0.490).

T.erythraeum is the more coastal clade associated with shallower water columns.

Niche differentiation may also be related to ecological competition associated with dFe availability. In support of this hypothesis, *T. thiebautii* was negatively correlated with dFe concentration (Spearman Correlation, **Table 2**). CCA, but not Spearman, revealed a correlation between dFe and clade abundances (**Figure 4**), where *T. erythraeum* was positively correlated with dFe, but negative correlated with *T. thiebautii*. Prior studies have shown that in Fe-limited open ocean waters, *T. thiebautii* is the dominant *Trichodesmium* clade (Chappell and Webb, 2010); our results suggest that when dFe is higher inshore, the oceanic *Trichodesmium* clade representative, *T. thiebautii*, is outcompeted by *T. erythraeum*.

We proffer that this trend is driven by competition with *T. erythraeum* and other phytoplankton inshore, meaning that *T. thiebautii* is outcompeted by *T. erythraeum* when dFe from SGD and/or other sources is elevated inshore. This is supported by the observations that *T. thiebautii* was also negatively correlated with Ba and Si concentrations (Spearman Correlation, **Table 2**),

while dFe and Ba were positively correlated with *T. erythraeum* (CCA, **Figure 4**), in particular during the Loop Current intrusion in 2015. The Loop Current intrusion resulted in enhanced offshore *T. erythraeum* abundance in 2015 (**Figure 3G** and **Supplementary Table 1**). However, we do not see *T. erythraeum* abundance elevated offshore in 2018 (**Figure 3E** and **Supplementary Table 1**), even though there was elevated dFe offshore as a result of the Loop Current intrusion (**Figure 2B** and **Supplementary Table 1**). This suggests that *T. erythraeum* is also Fe-limited offshore of the 50 m isobath on the WFS, potentially due to the lack of siderophore production mechanisms that open ocean *Trichodesmium* colonies have (Basu et al., 2019), except in situations where entrainment of dFe from continental runoff sources is carried by Loop Current to offshore locations.

CONCLUSION

Our findings show that the two main *Trichodesmium* clades, *T. thiebautii* and *T. erythraeum*, occupy distinct realized niches

TABLE 2 | Spearman individual correlations between *Trichodesmium* clades and chemical concentrations.

Element concentration correlations	p-value	Rho
dFe & *T. thiebautii* (left-tail)*	0.0047	-0.3794
Ba & *T. thiebautii* (left-tail)*	0.0133	-0.3201
Si & *T. thiebautii* (left-tail)*	0.0082	-0.3446
Salinity & *T. thiebautii* (right-tail)*	0.000149	0.4284
Water Column Depth & *T. thiebautii* (right-tail)*	0.0000302	0.4698
dFe & *T. erythraeum* (right-tail)	0.368	0.0511
Ba & *T. erythraeum* (right-tail)*	0.0172	0.3059
Si & *T. erythraeum* (right-tail)	0.113	0.1779
Salinity & *T. erythraeum* (left-tail)	0.1486	-0.1293
Water Column Depth & *T. erythraeum* (left-tail)*	0.0200	-0.2516

Asterisks indicate a significant p-value.

on the WFS. *T. thiebautii* is the more oceanic clade: its abundance was elevated in samples collected at deeper stations (>50 m), with elevated salinity and lower dFe. *T. erythraeum* is more coastal: it was most abundant at shallow stations with significant continental runoff inputs and higher dFe, and appears to outcompete *T. thiebautii* in coastal regions. These distinct niche occupations are likely due to resource competition between the two clades at stations with higher dFe. More work is needed to identify the metabolic pathways that distinguish the two clades and lead to the observed distinctions in distribution patterns. The intrusion of the Loop Current and associated entrainment of Missisippi River water brought higher dFe and elevates gene abundances of the two clades offshore of the 50 m isobath, suggesting that both clades are subject to Fe limitation on the outer shelf. Understanding the environmental niches of these two key taxa bears important implications for their contributions to global N and C cycling and their response to global climate change.

DATA AVAILABILITY STATEMENT

The datasets analyzed for this study is available for direct download (**Supplementary Table 1**) and are archived in the Biological and Chemical Oceanography Data Management Office (https://www.bco-dmo.org/project/814733).

AUTHOR CONTRIBUTIONS

ANK, KNB, and PDC designed the sampling program. CRS, KAC, and PDC designed the *Trichodesmium* analysis plan. All authors were involved in sample collection and/or analysis. KAC, CRS, and PDC wrote the initial manuscript draft. All other authors contributed to manuscript editing and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.821655/full#supplementary-material

REFERENCES

Amante, C., and Eakins, B. W. (2009). *ETOPO1 1 ArcMinute Global Relief Model: Procedures, Data Sources and Analysis* (NOAA Technical Memorandum NMFS F/SPO: National Marine Fisheries Service, NOAA). Boulder, Colorado.

Basu, S., Gledhill, M., De Beer, D., Prabhu Matondkar, S. G., and Shaked, Y. (2019). Colonies of Marine Cyanobacteria *Trichodesmium* Interact With Associated Bacteria to Acquire Iron From Dust. *Commun. Biol.* 2 (1), 284. doi: 10.1038/s420030190534z

Bergman, B., Sandh, G., Lin, S., Larsson, J., and Carpenter, E. J. (2013). *Trichodesmium* a Widespread Marine Cyanobacterium With Unusual Nitrogen Fixation Properties. *FEMS Microbiol. Rev.* 37 (3), 286–302. doi: 10.1111/j.15746976.2012.00352.x

Berman Frank, I., Cullen, J. T., Shaked, Y., Sherrell, R. M., and Falkowski, P. G. (2001). Iron Availability, Cellular Iron Quotas, and Nitrogen Fixation in *Trichodesmium*. *Limnol. Oceanogr.* 46 (6), 1249–1260. doi: 10.4319/lo.2001.46.6.1249

Breitbarth, E., Oschlies, A., and Laroche, J. (2007). Physiological Constraints on the Global Distribution of *Trichodesmium*-Effect of Temperature on Diazotrophy. *Biogeosciences* 4 (1), 53–61. doi: 10.5194/bg4532007

Capone, D. G. (2001). Marine Nitrogen Fixation: What's the Fuss? *Curr. Opin. Microbiol.* 4 (3), 341–348. doi: 10.1016/s1369-5274(00)00215-0

Capone, D. G., Burns, J. A., Montoya, J. P., Subramaniam, A., Mahaffey, C., Gunderson, T., et al. (2005). Nitrogen Fixation by *Trichodesmium* Spp.: An Important Source of New Nitrogen to the Tropical and Subtropical North Atlantic Ocean. *Global Biogeochem. Cycles* 19 (2), 1–14. doi: 10.1029/2004GB002331

Carpenter, E. J., and Capone, D. G. (2008). *Nitrogen Fixation in the Marine Environment* (New York: Elsevier), 141–198.

Chappell, P. D., Moffett, J. W., Hynes, A. M., and Webb, E. A. (2012). Molecular Evidence of Iron Limitation and Availability in the Global Diazotroph *Trichodesmium*. *ISME J.* 6 (9), 1728–1739. doi: 10.1038/ismej.2012.13

Chappell, P. D., and Webb, E. A. (2010). A Molecular Assessment of the Iron Stress Response in the Two Phylogenetic Clades of *Trichodesmium*. *Environ. Microbiol.* 12 (1), 13–27. doi: 10.1111/j.14622920.2009.02026.x

Charette, M. A., Henderson, P. B., Breier, C. F., and Liu, Q. (2013). Submarine Ground water Discharge in a River Dominated Florida Estuary. *Mar. Chem.* 156, 3–17. doi: 10.1016/j.marchem.2013.04.001

Delmont, T. O. (2021). Discovery of Nondiazotrophic *Trichodesmium*; Species Abundant and Widespread in the Open Ocean. *Proc. Natl. Acad. Sci.* 118 (46), e2112355118. doi: 10.1073/pnas.2112355118

Dixon, P. (2003). VEGAN, a Package of R Functions for Community Ecology. *J. Veg. Sci.* 14 (6), 927–930. doi: 10.1111/j.16541103.2003.tb02228.x

Falkowski, P. G. (1998). *The Oceanic Photosynthetic Engine: Origins, Evolution, and Role in Global Biogeochemical Cycles* (Netherlands: Springer), 3941–3947.

Gruber, N., and Sarmiento, J. L. (1997). Global Patterns of Marine Nitrogen Fixation and Denitrification. *Global Biogeochem. Cycles* 11 (2), 235–266. doi: 10.1029/97GB00077

Haselkorn, R., and Buikema, W. J. (1997). *Heterocyst Differentiation and Nitrogen Fixation in Cyanobacteria* (Berlin Heidelberg: Springer), 163–166.

Heil, C. A., Dixon, L. K., Hall, E., Garrett, M., Lenes, J. M., Oneil, J. M., et al (2014). Blooms of *Karenia Brevis* on the West Florida Shelf: Nutrient Sources and Potential Management Strategies Based on a MultiYear Regional Study. *Harmful Algae* 38, 127–140. doi: 10.1016/j.hal.2014.07.016

Hollister, A. P., Kerr, M., Malki, K., Muhlbach, E., Robert, M., Tilney, C. L., et al. (2020). Regeneration of Macronutrients and Trace Metals During Phytoplankton Decay: An Experimental Study. *Limnol. Oceanogr.* 65 (8), 1936–1960. doi: 10.1002/lno.11429

Hu, C., Muller Karger, F. E., and Swarzenski, P. W. (2006). Hurricanes, Submarine Groundwater Discharge, and Florida's Red Tides. *Geophys. Res. Lett.* 33 (11), 1–4. doi: 10.1029/2005GL025449

Hutchins, D. A., Fu, F.X., Webb, E. A., Walworth, N., and Tagliabue, A. (2013). Taxon Specific Response of Marine Nitrogen Fixers to Elevated Carbon Dioxide Concentrations. *Nat. Geosci.* 6 (9), 790–795. doi: 10.1038/ngeo1858

Hutchins, D. A., Fu, F. X., Zhang, Y., Warner, M. E., Feng, Y., Portune, K., et al (2007). CO2 Control of *Trichodesmium* N_2 Fixation, Photosynthesis, Growth Rates, and Elemental Ratios: Implications for Past, Present, and Future Ocean Biogeochemistry. *Limnol. Oceanogr.* 52 (4), 1293–1304. doi: 10.4319/lo.2007.52.4.1293

Hynes, A. M. (2009). Diversity of the marine Cyanobacterium *Trichodesmium*: Characterization of the Woods Hole Culture Collection and Quantification of Field Populations. *Dissertation.* Cambridge (MA): MIT/WHOI Joint Program.

Hynes, A. M., Webb, E. A., Doney, S. C., and Waterbury, J. B. (2011). Comparison of Cultured *Trichodesmium* (Cyanophyceae) With Species Characterized From the Field. *J. Phycol.* 48 (1), 196–210. doi: 10.1111/j.1529-8817.2011.01096.x

Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., et al. (2002). Dinitrogen Fixation in the World's Oceans. *Biogeochemistry* 57 (1), 47–98. doi: 10.1023/A:1015798105851

Knapp, A. N., Dekaezemacker, J., Bonnet, S., Sohm, J. A., and Capone, D. G. (2012). Sensitivity of *Trichodesmium Erythraeum* and *Crocosphaera Watsonii* Abundance and N_2 Fixation Rates to Varying NO_3^- and PO_4^{3-} Concentrations in Batch Cultures. *Aquat. Microb. Ecol.* 66, 223–236. doi: 10.3354/ame01577

Kralik, P., and Ricchi, M. (2017). A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and Everything. *Front. Microbiol.* 8, 108–108. doi: 10.3389/fmicb.2017.00108

Kustka, A. B., Saudo-Wilhelmy, S. A., Carpenter, E. J., Capone, D., Burns, J., and Sunda, W. G. (2003). Iron Requirements for Dinitrogen and Ammonium Supported Growth in Cultures of *Trichodesmium* (IMS 101): Comparison With Nitrogen Fixation Rates and Iron: Carbon Ratios of Field Populations. *Limnol. Oceanogr.* 48 (5), 1869–1884. doi: 10.4319/lo.2003.48.5.1869

Lenes, J. M., Darrow, B. P., Cattrall, C., Heil, C. A., Callahan, M., Vargo, G. A., et al. (2001). Iron Fertilization and the *Trichodesmium* Response on the West Florida Shelf. *Limnol. Oceanogr.* 46 (6), 1261–1277. doi: 10.4319/lo.2001.46.6.1261

Lenes, J. M., Darrow, B. A., Walsh, J. J., Prospero, J. M., He, R., Weisberg, R. H., et al (2008). Saharan Dust and Phosphatic Fidelity: A Three Dimensional Biogeochemical Model of *Trichodesmium* as a Nutrient Source for Red Tides on the West Florida Shelf. *Continental Shelf Res.* 28 (9), 1091–1115. doi: 10.1016/j.csr.2008.02.009

Levitan, O., Rosenberg, G., Setlik, I., Setlikova, E., Grigel, J., Klepetar, J., et al. (2007). Elevated CO_2 Enhances Nitrogen Fixation and Growth in the Marine Cyanobacterium *Trichodesmium*. *Global Change Biol.* 13 (2), 531–538. doi: 10.1111/j.13652486.2006.01314.x

Lien, V. S., Nilsen, J. E., Perivoliotis, L., Sotiropoulou, M., Denaxa, D., Ehrhart, S., et al. (2021). *BioGeoChemical Product Provided by the Copernicus Marine Service* (Göttingen, Germany: Copernicus GmbH).

Mellett, T., and Buck, K. N. (2020). Spatial and Temporal Variability of Trace Metals (Fe, Cu, Mn, Zn, Co, Ni, Cd, Pb), Iron and Copper Speciation, and Electroactive Fe Binding Humic Substances in Surface Waters of the Eastern Gulf of Mexico. *Mar. Chem.* 227, 103891. doi: 10.1016/j.marchem.2020.103891

Mills, M., Ridame, C., Davey, M., Laroche, J., and Geider, R. (2004). Iron and Phosphorus Co-Limit Nitrogen Fixation in the Eastern Tropical North Atlantic. *Nature* 429, 292–294. doi: 10.1038/nature02550

Morrison, J. M., Merrell, W. J., Key, R. M., and Key, T. C. (1983). Property Distributions and Deep Chemical Measurements Within the Western Gulf of Mexico. *J. Geophys. Res.* 88 (C4), 2601. doi: 10.1029/JC088iC04p02601

Oehler, T., Tamborski, J., Rahman, S., Moosdorf, N., Ahrens, J., Mori, C., et al. (2019). DSi as a Tracer for Submarine Groundwater Discharge. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00563

Parsons, T. R., Maita, Y., and Lalli, C. M. (1984). *A Manual of Chemical & Biological Methods for Seawater Analysis.* Eds. T. R. Parsons, Y. Maita and C. M. Lalli (Amsterdam: Pergamon), xiii–xxiv.

Prufert-Bebout, L., Paerl, H., and Lassen, C. (1993). Growth, Nitrogen Fixation, and Spectral Attenuation in Cultivated *Trichodesmium* Species. *Appl. Environ. Microbiol.* 59, 1367–1375. doi: 10.1128/aem.59.5.13671375.1993

Rouco, M., Haley, S. T., Alexander, H., Wilson, S. T., Karl, D. M., and Dyhrman, S. T. (2016). Variable Depth Distribution of *Trichodesmium* Clades in the North Pacific Ocean. *Environ. Microbiol. Rep.* 8 (6), 1058–1066. doi: 10.1111/17582229.12488

Rouco, M., Warren, H. J., McGillicuddy, D. J., Waterbury, J. B., and Dyhrman, S. T. (2014). *Trichodesmium* Sp. Clade Distributions in the Western North Atlantic Ocean. *Limnol. Oceanogr.* 59 (6), 1899–1909. doi: 10.4319/lo.2014.59.6.1899

Sanudo-Wilhelmy, S. A., Kustka, A. B., Gobler, C. J., Hutchins, D. A., Yang, M., Lwiza, K., et al (2001). Phosphorus Limitation of Nitrogen Fixation by *Trichodesmium* in the Central Atlantic Ocean. *Nature* 411 (6833), 66–69. doi: 10.1038/35075041

Selden, C. R., Chappell, P. D., Clayton, S., Macas Tapia, A., Bernhardt, P. W., and Mulholland, M. R. (2021). A Coastal N_2 Fixation Hotspot at the Cape Hatteras Front: Elucidating Spatial Heterogeneity in Diazotroph Activity *via* Supervised Machine Learning. *Limnol. Oceanogr.* 66 (5), 1832–1849. doi: 10.1002/lno.11727

Shaw, T. J., Moore, W. S., Kloepfer, J., and Sochaski, M. A. (1998). The Flux of Barium to the Coastal Waters of the Southeastern USA: The Importance of Submarine Groundwater Discharge. *Geochim. Cosmochim. Acta* 62 (18), 3047–3054. doi: 10.1016/S00167037(98)00218X

Sohm, J. A., Mahaffey, C., and Capone, D. G. (2008). Assessment of Relative Phosphorus Limitation of *Trichodesmium* Spp. in the North Pacific, North Atlantic, and the North Coast of Australia. *Limnol. Oceanogr.* 53 (6), 2495–2502. doi: 10.4319/lo.2008.53.6.2495

Wang, X., Li, H., Zheng, C., Yang, J., Zhang, Y., Zhang, M., et al. (2018). Submarine Groundwater Discharge as an Important Nutrient Source Influencing Nutrient Structure in Coastal Water of Daya Bay, China. *Geochim. Cosmochim. Acta* 225, 52–65. doi: 10.1016/j.gca.2018.01.029

Ye, Y., Volker, C., Bracher, A., Taylor, B., and Wolf-Gladrow, D. A. (2012). Environmental Controls on N_2 Fixation by *Trichodesmium* in the Tropical Eastern North Atlantic Ocean-A Model-Based Study. *Deep Sea Res. Part I: Oceanogr. Res. Papers* 64, 104–117. doi: 10.1016/j.dsr.2012.01.004

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership