ECO-EVOLUTIONARY PROCESSES WITHIN MICROBIAL COMMUNITIES

EDITED BY: Maria Elena Martino, Alejandro Couce, Andrea Quagliariello and Ricardo S. Ramiro

PUBLISHED IN: Frontiers in Microbiology and Frontiers in Genetics

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-595-3
DOI 10.3389/978-2-83250-595-3

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

ECO-EVOLUTIONARY PROCESSES WITHIN MICROBIAL COMMUNITIES

Topic Editors:
Maria Elena Martino, University of Padua, Italy
Alejandro Couce, Polytechnic University of Madrid, Spain
Andrea Quagliariello, University of Padua, Italy
Ricardo S. Ramiro, New University of Lisbon, Portugal

Citation: Martino, M. E., Couce, A., Quagliariello, A., Ramiro, R. S., eds. (2022). Eco-evolutionary Processes Within Microbial Communities. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-595-3

Table of Contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 11 October 2022
DOI 10.3389/fmicb.2022.1045045

OPEN ACCESS

EDITED AND REVIEWED BY
Ludmila Chistoserdova,
University of Washington,
United States

*CORRESPONDENCE
Maria Elena Martino
mariaelena.martino@unipd.it

SPECIALTY SECTION
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

RECEIVED 15 September 2022
ACCEPTED 20 September 2022
PUBLISHED 11 October 2022

CITATION
Quagliariello A, Ramiro RS, Couce A and Martino ME (2022) Editorial: Integrating multi-scale approaches for predicting microbiome ecology and evolution.
Front. Microbiol. 13:1045045.
doi: 10.3389/fmicb.2022.1045045

Editorial: Integrating multi-scale approaches for predicting microbiome ecology and evolution

Andrea Quagliariello[1], Ricardo S. Ramiro[2], Alejandro Couce[3] and Maria Elena Martino[1]*

[1]Department of Comparative Biomedicine and Food Science, University of Padua, Padua, Italy, [2]InnovPlantProtect, Elvas, Portugal, [3]Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM), Madrid, Spain

KEYWORDS

microbial ecology, microbial evolution, microbial communities, microbial adaptation, comparative genomics

Editorial on the Research Topic
Integrating multi-scale approaches for predicting microbiome ecology and evolution

An intricate balance of factors determines the evolution of microbes living in communities. Ever-changing environmental conditions and complex microbial interactions combine in non-linear ways to hamper the identification of the general principles and basic mechanisms governing the evolution of microbial communities (Little et al., 2008). However, recent work has started to disentangle the predominant factors underpinning these processes, mainly under stable laboratory conditions. In addition, a large body of research has been focusing on the development and integration of experimental and bioinformatic approaches toward the understanding of microbial evolution. The combination of meta-omics approaches, high-throughput sequencing, mathematical modeling and experimental approaches is shedding light into genetic mechanisms underlying strain- and species-specific adaptation within microbial communities evolving in complex and heterogeneous environments.

To illustrate these recent advances in the 'Ecology and Evolution within Microbial Communities' field, we have selected nine articles for this Research Topic, which we believe focus on key avenues in this research field.

An important advancement in microbiome research has been the use of Experimental Evolution (EE), which largely facilitated the monitoring of bacterial genetic evolution and the understanding of adaptive evolution processes. EE corresponds to the study of the evolutionary modifications occurring within microbial populations in response to defined and controlled environmental stresses imposed by the experimenter (Kawecki et al., 2012; Barrick and Lenski, 2013; McDonald, 2019). Such experimental model makes it possible to constantly monitor the adaptation of populations to the surrounding environment by observing real-time evolution for ten, hundreds or even thousands of generations and allowing the detection of both phenotypic and genetic changes among individuals in the populations. Coupling EE with mass sequencing

and comparative genomics allows us to unravel the genetic mechanisms underlying adaptation in bacteria at the genomic level. In this light, EE studies have strongly contributed to recognize the genetic changes that distinguish an evolved clone from its original ancestor strain in a wide range of experimental conditions and to reveal the functional networks involved in microbial adaptation (Segrè et al., 2006; Kawecki et al., 2012; Bailey and Bataillon, 2016). Historically, EE assays have mostly focused on low complexity systems, while few studies have been conducted on complex systems using natural microbial communities. The review by Manriquez et al.presents an outlook on microbial EE studies done with systems of increasing complexity (from single species, to synthetic communities and natural communities), with a particular focus on studies between plants and plant-associated microorganisms (e.g., focal bacterial pathogens adapting to different hosts; molecular mechanisms leading to microbe/plant beneficial interaction; bulk soil communities in response to artificial selection applied on plant phenotypes). In this regard, the authors point out the current need to distinguish to what extent plant phenotypic traits are determined by the environmental pressure and the microbial ecological and evolutionary dynamics.

One evolutionary advantage afforded by living in communities is the possibility of receiving useful external genes *via* horizontal gene transfer (HGT). However, the maintenance of mobile genetic elements in microbial communities depends on the delicate balance between the benefits and costs carrying these elements entails. The study by Cyriaque et al. challenges the simple expectation that the spreading success of plasmids carrying beneficial genes (in this case, heavy metal resistance genes—MRG) is greater in the presence of strong selective pressures in favor of those genes (i.e., high concentration of lead). Contrary to the expectations, the authors find that high concentrations of lead elicit physiological responses that inhibit conjugation and partition systems, thus reducing plasmid dispersal. Indeed, the spread of MRG-carrying plasmids was best at intermediate metal concentrations. Overall, this study paints a complex picture in which the success of mobile genetic elements hinges on the specific physiological and mutational propensities of the constituents of a community, suggesting that the opportunity for HGT-driven evolution may vary widely across microbial communities.

If the previous study shows that environmental stresses can translate into non-linear selective pressures, the study by Pinto et al.investigates how taking into consideration microbiome composition adds an extra layer of complexity to the genotype to fitness map. In particular, they studied the selective advantage constituted by the wellknown *lac* operon in *Escherichia coli* when colonizing the mammalian gut. By directly competing lac+ and lac– *E. coli* strains in the mouse gut, the authors found that, in the presence of lactose, the *lac* operon confers a selective advantage that can be as high as 11%, while being close to neutral in the absence of this sugar. Such fitness advantage directly depends on microbiota composition and on population density. In particular, as the number of species increases, so does the opportunity for competition (i.e., the ability to consume lactose represents an advantage as other nutrients get depleted) as well as opportunity for synergism, with the absolute success of *E. coli* reflecting the net result between these two forms of ecological interactions. In this regard, the authors found that the population size of *E. coli* substantially increases in the triple-colonization, in comparison to mono-colonization, suggesting that, in this scenario, synergism dominates. Conversely, in the presence of a complex microbiota (several tens of species), its absolute success decreases and the most likely scenario is one dominated by competition. The adaptive value of specific operons is thus primarily dependent on the microbial diversity of a given community and the specificity of interactions among species. Yet, the extent to which interactions with coexisting species can alter microbial evolution and diversification remains elusive. Hypotheses range from coexistence impeding diversification due to niche exclusion, to coexistence boosting diversification due to niche construction

On this subject, the work by Chu et al. documents one example of niche exclusion: the presence of multiple species reduces the evolutionary success of the dominant morphotype that takes over *Pseudomonas fluorescens* populations in a classic model of adaptive radiation. In particular, the authors show that the effect is mostly driven but just one of the coexisting species (*Ochrobactrum* sp.), suggesting that previous reports on the positive relationship between species richness and niche exclusion is not a general effect, but simply the product of an increased likelihood of sampling at least one species with strong competitive abilities. These results may help to understand the inconsistent effects of interspecific competitors on diversification of resident species across different studies and suggest that the presence of niche-specific competitors represents a potential explanation.

However, on a broader scale, Bajić et al. make the case that the process of niche construction, by which organisms alter their environments creating novel evolutionary opportunities, plays a major role on how microbial communities assemble, diversify and evolve. Specifically, the authors argue that niche construction may provide support to genotypes that are not viable in the original niche, but that can act as "stepping-stones" to previously inaccessible regions of the genotypic space. However, a systematic, direct exploration of these possibilities is beyond current experimental capabilities. To address this issue, the authors point at the recent advances in genome-scale metabolic models (e.g., platforms such as "Computation of Microbial Ecosystems in Time and Space"—COMETS) that are now able to capture complex phenomena observed in microbial communities, such as niche construction. The authors suggest that these genome-scale metabolic models offer a unique opportunity to understand biological processes at lower levels of

organization without isolating them from the eco-evolutionary processes in which they are embedded. Yet, while metabolic models can predict the release of some compounds, the origin of complex metabolic mixtures is still poorly understood within microbial communities.

Understanding how microbial diversity scales through space, between connected ecosystems, is a key question for microbial ecology. This gains additional relevance when considering the microbiota connectivity among humans and across the collection of ecosystems within the human digestive tract, which play a crucial role in human health (Donaldson et al., 2016). Species-area relationship (SAR) models have long been a key tool for ecologists to analyse the scaling of biodiversity across space. In their study, Chen et al. apply a recently developed extension of the classical SAR, named the diversity-area relationship (DAR) to understand how the human digestive tract (DT) microbiota biodiversity changes both within (across DT sites) and across human adults. Their analysis shows that the main differences between individuals in terms of microbial diversity were at the level of taxa presence/absence (i.e., species richness) rather than at the level of taxa abundance (e.g., Shannon entropy or Simpson index). When comparisons were made at the DT site level, differences between sites were maintained across diversity metrics and were clearly stronger than between individuals, as would be expected given the strong biological differences between DT sites. The study reveals how inter and intra-individual variation shapes microbial biodiversity in healthy adults and opens the possibility of using the diversity and the ecological structure of the microbiota as biomarkers of disease, rather than the presence of specific microbes.

The study of microbial ecology and evolution is currently largely dependent on *in silico* approaches. Comparative genome analyses are at the core of microbiome research and are crucial to decipher the evolutionary relationships between species, giving also clinical clues in case of bacterial pathogens. By comparing the genomes of 290 strains of *Clostridium perfringens*, Jaakkola et al. disclose two distinct sub-groups of *C. perfringens* strains with different virulence characteristics, spore heat resistance properties, and, presumably, ecological niche. The authors reveal that spore heat resistance of *C. perfringens* strains is likely affected by multiple genes and the capacity to produce heat-resistant spores has developed primarily within one subgroup of chromosomal enterotoxin-carrying strains. Studies based on comparative genomics are thus instrumental to elucidate the population structure and ecology of bacterial species and communities. Coupling taxonomic comparison with microbial function represents a further step in understanding microbiome evolution.

In this regard, Chen et al. examined 938 annotated water metagenomes obtained worldwide to investigate the connection between function and taxonomy, and to identify the key drivers of water metagenomes function or taxonomic composition at a global scale. The authors find that microbial functions are significantly correlated to taxonomy in aquatic ecosystems, with salinity being the predominant driver of microbial functional and taxonomic diversities. Classification into six water biomes resulted in greater variation in taxonomic compositions than functional profiles, highlighting significant functional redundancy at a global level.

Finally, Kumar et al. consider how environmental pressure shapes microbial evolution, by providing an extensive review on the mechanisms of marine fungi adaptation to climate change. The authors show that environmental stressors cause a direct alteration of cellular and molecular pathways in *Hortaea werneckii* and *Aspergillus terreus*, which ultimately leads to fungal adaptation to environmental change. The authors claim that understanding the mechanisms and the extent to which microbes adapt to environmental fluctuations, particularly in the era of climate change, is key to help researchers to develop transgenic organisms that are able to grow in such environments and whose flexibility will reduce the risk of climate change impacts on environments and organisms.

Altogether, these studies showcase the need for a multi-pronged approach to gain insights into the complex relationship among the environment, microbiota composition and evolutionary outcomes. Taken together, there is reason to hope that integrating different approaches (experimental evolution, comparative metagenomics, metabolic models and functional associations) will pave the way to a better understanding of microbial evolution. This will also allow predicting evolution in many important real life scenarios, such as antibiotic resistance evolution, human microbiome functioning under development, aging and disease, or soil microbiome stability and crop yields under accelerating rates of global change.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

AQ was supported by the Supporting Talent in Research @ University of Padua – STARS Grants 2019. AC was supported by the Spanish Ministry of Science, Innovation, and Universities (PID2019-110992GA-I00) and the regional government of Madrid (2019-T1/BIO-12882). RR was supported by the European Social Fund (ALT20-05-3559-FSE-000036) and the Fundação para a Ciência e a Tecnologia (FCT) based on RCM 23/8 March 2018. MM was supported by the Supporting Talents

@Univesrity of Padua - STARS Grant 2017 and by the Italian Ministry for Universities and Research (MUR).

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.

References

Bailey, S. F., and Bataillon, T. (2016). Can the experimental evolution programme help us elucidate the genetic basis of adaptation in nature? *Mol. Ecol.* 25, 203–218. doi: 10.1111/mec.13378

Barrick, J., and Lenski, R. (2013). Genome dynamics during experimental evolution. *Nat. Rev. Genets.* 14, 827–839 doi: 10.1038/nrg3564

Donaldson, G. P., Lee, S. M., and Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol.* 14, 20–32. doi: 10.1038/nrmicro3552

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M. C. (2012). Experimental evolution. *Trends Ecol. Evol.* 27, 547–560. doi: 10.1016/j.tree.2012.06.001

Little, A. E. F., Robinson, C. J., Peterson, S. B., Raffa, K. F., and Handelsman, J. (2008). Rules of engagement: interspecies interactions that regulate microbial communities. *Annu. Rev. Microbiol.* 62, 375–401. doi: 10.1146/annurev.micro.030608.101423

McDonald, M. J. (2019). Microbial Experimental Evolution – a proving ground for evolutionary theory and a tool for discovery. *EMBO Rep.* 20, e46992. doi: 10.15252/embr.201846992

Segrè, A. V., Murray, A. W., Leu, J. Y. (2006). High-resolution mutation mapping reveals parallel experimental evolution in yeast. *PLoS Biol.* 4, e256. doi: 10.1371/journal.pbio.0040256

frontiers in Microbiology

REVIEW
published: 07 May 2021
doi: 10.3389/fmicb.2021.619122

Experimental Evolution in Plant-Microbe Systems: A Tool for Deciphering the Functioning and Evolution of Plant-Associated Microbial Communities

*Beatriz Manriquez, Daniel Muller and Claire Prigent-Combaret**

UMR 5557 Ecologie Microbienne, VetAgro Sup, CNRS, INRAE, University of Lyon, Université Claude Bernard Lyon 1, Villeurbanne, France

In natural environments, microbial communities must constantly adapt to stressful environmental conditions. The genetic and phenotypic mechanisms underlying the adaptive response of microbial communities to new (and often complex) environments can be tackled with a combination of experimental evolution and next generation sequencing. This combination allows to analyse the real-time evolution of microbial populations in response to imposed environmental factors or during the interaction with a host, by screening for phenotypic and genotypic changes over a multitude of identical experimental cycles. Experimental evolution (EE) coupled with comparative genomics has indeed facilitated the monitoring of bacterial genetic evolution and the understanding of adaptive evolution processes. Basically, EE studies had long been done on single strains, allowing to reveal the dynamics and genetic targets of natural selection and to uncover the correlation between genetic and phenotypic adaptive changes. However, species are always evolving in relation with other species and have to adapt not only to the environment itself but also to the biotic environment dynamically shaped by the other species. Nowadays, there is a growing interest to apply EE on microbial communities evolving under natural environments. In this paper, we provide a non-exhaustive review of microbial EE studies done with systems of increasing complexity (from single species, to synthetic communities and natural communities) and with a particular focus on studies between plants and plant-associated microorganisms. We highlight some of the mechanisms controlling the functioning of microbial species and their adaptive responses to environment changes and emphasize the importance of considering bacterial communities and complex environments in EE studies.

Keywords: experimental evolution, synthetic community, interaction network, microbiota, holobiont, evolutionary adaptation

OPEN ACCESS

Edited by:
Ricardo S. Ramiro, Gulbenkian Institute of Science (IGC), Portugal

Reviewed by:
Israel Pagan, Polytechnic University of Madrid, Spain
Inês Fragata, University of Lisbon, Portugal
Teppo Hiltunen, University of Turku, Finland

***Correspondence:**
Claire Prigent-Combaret
claire.prigent-combaret@univ-lyon1.fr

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *19 October 2020*
Accepted: *29 March 2021*
Published: *07 May 2021*

Citation:
Manriquez B, Muller D and Prigent-Combaret C (2021) Experimental Evolution in Plant-Microbe Systems: A Tool for Deciphering the Functioning and Evolution of Plant-Associated Microbial Communities. Front. Microbiol. 12:619122. doi: 10.3389/fmicb.2021.619122

INTRODUCTION

In nature, microorganisms are living inside complex microbial communities (i.e., microbiota) where they evolve under constant interaction with sympatric microbial populations (Brockhurst et al., 2003; Brockhurst and Koskella, 2013; Guan et al., 2013). While tremendous progress has been achieved regarding the description of natural microbial community composition, understanding

their functioning, their structure dynamics and how they may evolve is still in its infancy (Rainey and Quistad, 2020). Myriads of interactions (e.g., cooperation, competition, or predation) occur between microbial community members which are influenced by all the biotic and abiotic factors the microbial community faced up. In turn, the functioning of microbial communities affects their environment leading to further changes in biotic interactions and in evolutionary dynamics of its members. Microbial communities are thus dynamical systems.

The environmental persistence and functioning of microbial communities are influenced not only by their species richness but also by their functional diversity. There are thus two important aspects of microbial communities: their taxonomic structure (diversity and abundance of each individual population within the community) and how they function (community behavior and activities) (Little et al., 2008). Microbial communities are composed of various functional groups which could be defined as all populations doing the same function (Bouffaud et al., 2016). The same population could belong to several functional groups, which leads to interconnected functional networks within a single microbial community (Bruto et al., 2014; Renoud et al., 2020). Reconstructing the structure of functional and metabolic networks within microbial communities and understanding how this structure might evolve over time is largely unknown even when imposing stable laboratory environmental conditions. However, there is a great expectation placed on experimental evolution (EE) studies to unravel the evolutionary dynamics of populations within microbial community, although EE studies on natural microbial community are so far uncommon.

Experimental evolution corresponds to the study of the evolutionary modifications occurring on populations in response to environmental conditions imposed by the experimenter (Kawecki et al., 2012). It made it possible to monitor the real-time adaptation of populations to their environment (biotic and abiotic) by observing evolution in real-time for ten, hundreds or even thousands of generations and detecting phenotypic or genetic changes between individuals in the populations. The use of experimental replicates allows to disentangle the contribution of chance events (such as drift and founder effects) and the contribution of the imposed selective pressure. Thus, unlike carrying out genomic analysis on existing natural organisms and then interpreting their evolution, EE makes it possible to transform evolutionary genetics into a prospective undertaking, and to decipher the genetic bases of adaptation (Van den Bergh et al., 2018).

The first experiments with continuous-cultures were employed in the early 1950s with microorganisms and focused on describing dominant mutant phenotypes favored during extended growth (reviewed in Adams and Rosenzweig, 2014). In the decades that followed, scientists attempted to correlate the observed phenotypic changes with gene mutations or duplications, revealing that genetic adaptation is the basis of evolutionary adaptation. Genetic changes happen rapidly, within the first hundred cell generations, for a bacterial monoculture grown in a unique and limited resource environment (Lenski and Travisano, 1994; Cooper and Lenski, 2000; Wiser et al., 2013; Kram et al., 2017; Lenski, 2017). Bacteria are powerful candidates for EE as they offer short generation time and large population size, so multiple mutations can be present simultaneously. Moreover, bacteria are easy to track (i.e., isolation, numeration) and store, this allows to compare competitive fitness between evolved and ancestral genotypes. A regular collection and conservation of samples during EE and further genomic, genetic and/or phenotypic analyses could be done to decipher the evolution process (Jerison and Desai, 2015; Marchetti et al., 2017). Indeed, genomic and molecular data are currently available for many bacterial species, as well as techniques for their precise genetic analysis and manipulation (Vacheron et al., 2018; Batstone et al., 2020; Li et al., 2020; Scheuerl et al., 2020; Rodríguez-Verdugo and Ackermann, 2021).

Nowadays, mass sequencing allows to unravel the genetic mechanisms underlying adaptation in bacteria at the genomic level (Kawecki et al., 2012; Bailey and Bataillon, 2016). EE can now be used to answer more complex ecological questions, such as evolutionary adaptation during biotic interactions between a bacterial population and a host or between bacterial populations within synthetic or natural microbial communities (Brockhurst and Koskella, 2013; **Table 1** and **Figure 1**). It is also a powerful tool to decipher the underlying mechanisms of virus-bacteria co-evolution (Paterson et al., 2010; Scanlan et al., 2011), adaptation of pathogens to humans (Yang et al., 2011) or evolution of animal gut microbiota (Lescat et al., 2017; Tso et al., 2018) but the latter topics are not discussed in the present review. Here, a particular focus was given, but not limited, to EE studies done with plant-microbe systems. We compare systems with increasing complexity (**Table 1**). First, using selected examples, we describe EE studies involving a single microbial species evolving under low complexity environmental conditions. These studies allowed to infer the genomic mechanisms underlying bacterial adaptation to new environments. Second, we analyse EE studies using more complex systems, such as those investigating the interaction between single microorganism and the plant. These EE studies shed light on the role of the eco-evolutionary feedbacks during microbe-plant interactions. Third, we examine the importance of considering synthetic or natural bacterial communities (or microbiota) and complex environments in EE studies. Finally, we discuss future avenues of EE studies and point out the gaps must be bridged to analyze complex systems with the same detailed analysis of genomics adaptation than simple systems (**Figure 1**).

EXPERIMENTAL EVOLUTION TO DECIPHER THE ADAPTIVE EVOLUTION OF SINGLE MICROBIAL POPULATION IN SYNTHETIC SYSTEMS OF LOW COMPLEXITY

Experimental evolution has mostly been carried out on very simple conditions: using a single organism evolving in a low complexity system (Barrick et al., 2009). These EE experiments have provided essential knowledge on the molecular evolution and adaptive changes of microorganisms, particularly bacteria,

TABLE 1 | Selected examples of experimental evolution studies (with a focus on interactions with plants at the microcosm scale), that have contributed to advances in genomic evolution, horizontal gene transfer, and plant host–microbe interactions.

			Simple synthetic media	Complex and/or fluctuating media	Microcosms with plant host
Single clone	Wild-type 	**Single strain evolution** • Genetic/genomic evolution • Transcriptional changes • Adaptive radiation	• *Escherichia coli* single clone evolution (Lenski et al., 1991; Lenski and Travisano, 1994; Lenski, 2017) • Impact of the inoculum size (Garoff et al., 2020)	• Adaptive radiation of *Pseudomonas fluorescens* (Buckling et al., 2003; Barrett et al., 2005; Flohr et al., 2013; Koza et al., 2017) • Long-term evolution of *E. coli* (Kram et al., 2017) • Importance of horizontal gene transfer in *Helicobacter pylori* evolution (Woods et al., 2020) • *Saccharomyces cerevisiae* showing a transition to multicellularity life form (Ratcliff et al., 2012) • *Myxococcus xanthus* evolution (Rendueles and Velicer, 2017)	• *Ralstonia solanacearum* pathogen adaptation to plants (Guidot et al., 2014; Gopalan-Nair et al., 2020) • *Pseudomonas protegens* adaptation to Arabidospis (Li et al., 2020)
	Mutant/variant 	**Single strain evolution** • Genetic/genomic evolution • Transcriptional changes • Adaptive radiation	• *E. coli* single clone evolutions (reviewed in Lenski, 2017)	• Adaptive radiation of *P. fluorescens* SBW25 Δ*panB* (Barrett et al., 2005) • *Myxococcus xanthus* evolution (Rendueles and Velicer, 2020)	• Modified *Ralstonia solanacearum* evolving in a plant symbiont (Marchetti et al., 2010; Doin de Moura et al., 2020)
Synthetic community		**Several known species co-evolution** • Population dynamics • Genetic/genomic evolution of each species • Biotic interactions evolution • Gene transfer	• Bacterial–prey co-evolution (Nair et al., 2019; Scheuerl et al., 2019) • Yeast mutualists co-evolution (Vidal and Segraves, 2020) • Experimental co-evolution (Castledine et al., 2020) • Impact of the diversity and composition of the synthetic community on the evolution of a given species (Osmond and de Mazancourt, 2013; Jousset et al., 2016; Calcagno et al., 2017) • Neutral and selective dynamics in a synthetic microbial community (Cira et al., 2018)	• Fluctuating environment destabilizing bacterial interactions (Rodríguez-Verdugo and Ackermann, 2021) • *Acinetobacter–Pseudomonas putida* interaction, evolution in a biofilm (Hansen et al., 2007) • Impact on biotic interaction network (Lawrence et al., 2012; Cairns et al., 2018a) • Importance of spatial structure (Cairns et al., 2018b) and toxic molecules (Piccardi et al., 2019) • Community context affecting evolutionary dynamics (Fiegna et al., 2015; Scheuerl et al., 2020)	• *Pseudomonas syringae* adaptation on its natural host or on a distant one in the presence or absence of bacteriophages (Meaden and Koskella, 2017) • Enhanced cooperation with different plant genotypes of *Ensifer meliloti* in competition with a nitrogen-fixing defective cheater (Batstone et al., 2020)
Natural community		**Several species co-evolution** • Population dynamics in the future: single-cell genomic evolution and horizontal gene transfer?	No studies in the literature but could allow, with meta-omics and single-cell sequencing, analysis of: • Population dynamics • Genetic mechanisms underlying species-specific adaptation within microbiota • Genetic changes, at the cell level, within a community Selection of adapted and stable microbiota applicable to agriculture or medicine issues		• Adaptation of microbiota to host plants (Morella et al., 2020) • Evolution of rhizosphere microbiota to influence plant growth parameters (flowering, biomass) (Swenson et al., 2000; Panke-Buisse et al., 2015; Lu et al., 2018)

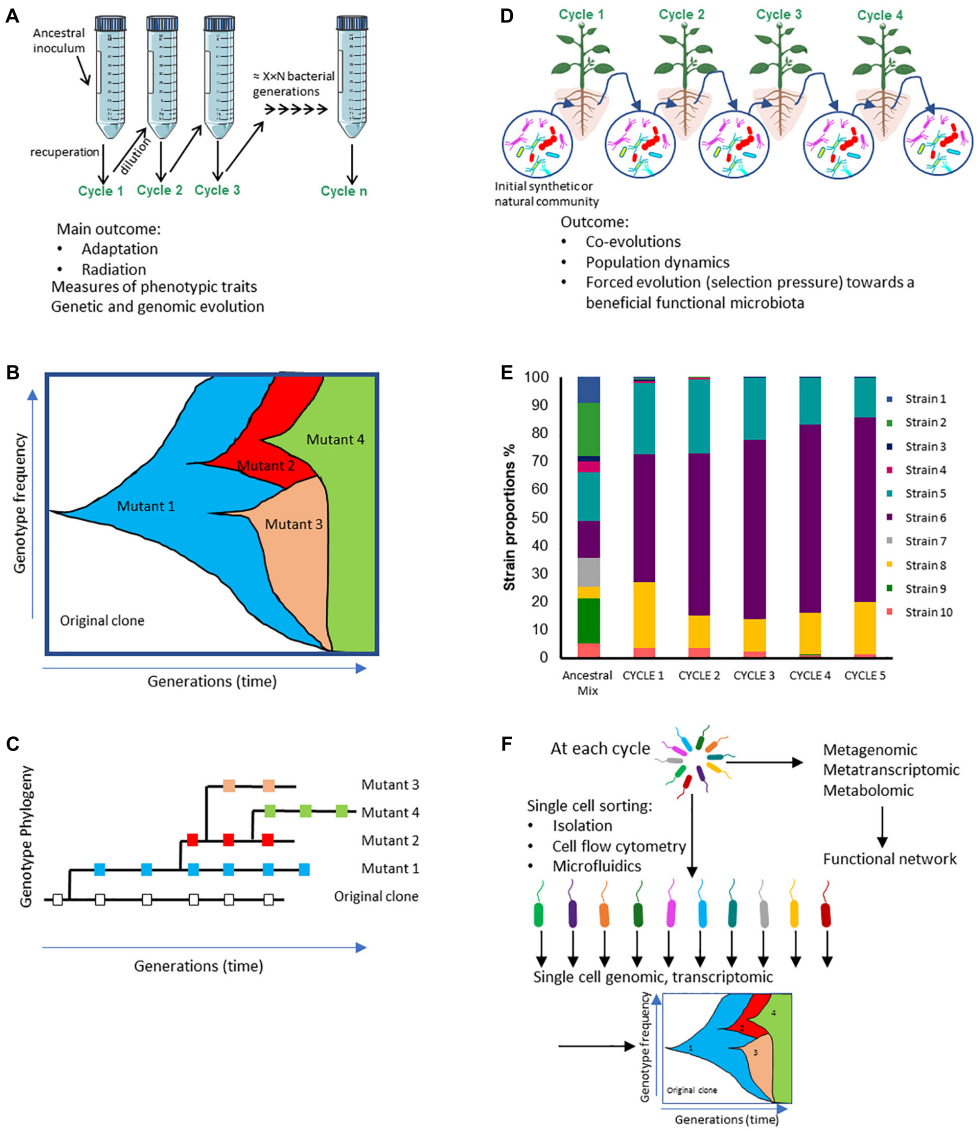


FIGURE 1 | Comparison of the potential outcomes from experimental evolution (EE) studies starting with **(A)** a single clone or **(D)** a microbial community. **(A)** Example of an experimental evolutionary design based on a single clone serially cultivated in a simple medium, a stressful or a changing environment; those experiments enable to track phenotypic and genetic changes over time. **(B)** Theoretical Muller diagram depicting the distinct genotypes frequencies over microbial generations that could be observed in EE study described in A. **(C)** Theoretical phylogenetic tree built from genotypes that can be sampled from different time points throughout the EE described in A and sampled at the same generation times. **(D)** Example of an experimental evolutionary design based on an inoculum made up of several microorganisms (synthetic or natural communities); those experiments make it possible to monitor changes in population levels **(E)**, and genetic, transcriptomic or metabolomic changes over time **(F)**. **(E)** Theoretical bar graph illustrating the composition of a microbial community showing differences in taxa among both the inoculum (ancestral mix) and different experimental cycles. **(F)** At each cycle, approaches could be implemented at the scale of the whole community (i.e., meta-omics) or of individual cells after cell-sorting (single-cell sequencing, transcriptomic).

to experimenter-imposed conditions and allowed to uncover the correlation between genetic and phenotypic changes (**Figure 1**).

Bacterial Adaption and Key Evolutionary Driving Forces

Adaptation of populations mostly happens through mutations that can improve the performance of organisms in their environment and therefore improve their fitness (Lang et al., 2011). Distinct environmental conditions will impose different selection pressures and will determine whether or not a mutation would be beneficial or detrimental in a given niche at a given time. If a beneficial mutation happens on a single clone it will need a certain amount of time to rise to the population level (fixation). According to Fisher's model, the time required for a mutation to fix is inversely proportional to its beneficial effect (Fisher, 1930; Tenaillon, 2014). However, stochastic extinction of a mutation, even if it

may confer some fitness advantages, or increase in frequency of a deleterious mutation can exist. So called "genetic drift" also plays an important role in fixation of mutation in a population and represents the role of chance in evolution (Kayser et al., 2018; McDonald, 2019). The spatial distribution of cells will also influence the genetic drift's force and competition between mutated clones (Kayser et al., 2018). However, bacteria reproduce asexually, so the different genotypes that rise in the population do not recombine and the corresponding mutations they harbor may not reach fixation within the population. This phenomenon is known as clonal interference and influences the dynamics of evolution and adaptation as clones with beneficial mutations will interfere with each other's spread in the population (Gerrish and Lenski, 1998; McDonald, 2019). However, motility and dispersal may help competitive populations to coexist in a same system through active segregation and spatial exclusion (Gude et al., 2020).

Parallel and Convergent Evolution of Single Microbial Species in Simple Environments

Given that evolution in bacteria happens mostly by random mutations, an important question arises from EE studies: are they repeatable? Do same genotypic and phenotypic changes arise in an evolved microbial population submitted to same environmental changes? In order to better understand evolutionary adaptation, genotype-to-phenotype correlations need to be more fully investigated. This can be done by focusing on patterns of parallel and convergent evolution.

Parallel evolution is defined as the independent evolution of similar phenotypic traits in lineages closely related to each other and involving changes in orthologous genes, whereas convergent evolution concerns non-related phylogenetic lineages and changes in non-homologous genes (Bai et al., 2015; Pickersgill, 2018).

Many EE studies have shown that a single clonal ancestor can give rise to independently evolved populations sharing similar traits under the same environmental conditions (Schluter et al., 2004; Colosimo, 2005; Zhang et al., 2005). Lenski's experimental evolution study is composed of 12 replicate populations of *E. coli* B, which are currently still evolving on a glucose-limited minimal salts medium (Lenski et al., 1991; Lenski, 2017; Consuegra et al., 2021; **Table 1**). DNA microarray analyses of two evolved populations showed similar changes in the transcription of 59 genes after 20,000 generations; the genetic bases of these changes were investigated and relevant parallel mutations were found on many of the independently evolved populations (Cooper et al., 2003).

In simple organisms, biological functions are mostly encoded by single genes, whereas, in higher forms of life, more complex regulatory networks are involved. In bacteria and yeast, complex functions are often related to modules of genes (Hartwell et al., 1999). Reflecting this, parallel evolution will not always mean changes in the same genes but rather similar changes in related gene modules. Herring et al. (2006) explored the parallel changes in metabolic and regulatory networks that arose in five *E. coli* populations that evolved separately. They proved, combining mutant approach and whole-genome resequencing of five clones of *E. coli*, that 13 different spontaneous mutations were responsible for improved fitness during adaptation to a glycerol synthetic growth medium (Herring et al., 2006). This study provides a clear example of how different changes can have similar phenotypic effects (**Table 1**). Dissimilar genetic changes can lead to parallel phenotypes, meaning that functional connections within and between genetic modules can be established by pairing experimental evolution to whole genome sequencing (Segrè et al., 2006; Yeh et al., 2006).

However, clonal populations evolving in an identical environment could reach different fitness rates and may diverge (Lenski et al., 1991). This divergence reveals that there may be more than one adaptive strategies under the same environmental conditions (Elena and Lenski, 2003). On the other hand, we could expect that experiments ran with genetically different organisms led to different fitness. Travisano et al. (1995) evaluated fitness on maltose of several *E. coli* lines that had evolved in glucose for 2,000 generations. Lines that started with the lower fitness in maltose improved faster, but all the lines tend eventually to converge to a similar fitness on maltose evidencing convergent evolution.

EE of Single Microbial Species in Variable and Complex Environments

In the past, EE studies were mostly done in simple environments, where the concentration of one essential source of carbon available to all individuals controls the population growth rate (Hillesland and Stahl, 2010; Yu et al., 2017).

This selects for different types of genotypes throughout experiment transfers. Heavy consumers and fast-growing microbes show increased fitness in the beginning of the experiment. When the resource starts to be depleted, genotypes that are able to survive with low resources and metabolic by-products increase in frequency. Finally, when there are almost no resources, there is an increase in prevalence of genotypes that are able to either metabolize toxic by-products (e.g., Blount et al., 2012) or are able to grow in very small amount of resources. This may lead to the evolution of two types of genotypes: generalist (able to consume a large range of the available resources) or specialist (able to grow faster than generalists, but in a shorter range of resources). This ensures the maintenance of polymorphism in population, with all genotypes present in the population, but with their prevalence varying until the cycle is started again with a new medium transfer. In complex environments that contain several resources, cell populations may become wide specialists able to consume various carbon substrates.

In the laboratory, because of the complexity of evolutionary processes, most of the EE assays have been carried out in a constant environment, using chemostats or continuous culture systems, so the growth conditions would remain the same for the whole experiment (Treves et al., 1998; Maharjan et al., 2006; Gresham and Hong, 2015). Recently, Westphal et al. (2018) used a batch culture system to study the impact of four changing environments on the adaptation of *E. coli*. In these systems there is no addition of nutrients, therefore the cultures experience fluctuation in nutrient availability. Therefore, as nutrients are consumed, waste products are released and

less energetically favorable metabolisms become important for survival. Their data revealed that different environments may select different mutations; this emphasizes the importance of performing experimental evolution in complex and ever-changing environments. To take these findings one step further, Kram et al. (2017) carried out an EE assay with *E. coli* in a complex and variable environment (i.e., in a rich medium LB and with serial passages every 4 days). This scheme allows cells to go through all phases of growth and to adapt to different stresses (nutrient limitation, oxidative stress and pH variation). This experiment showed that after only 30 generations, evolved populations presented changes in growth rates but also adaptive mutations allowing the cells to cope with the varying stresses arising during the culture (**Table 1**). They also evidenced parallel changes in evolved populations (Kram et al., 2017).

The importance of environmental complexity and evolution of niche width have been studied by Barrett et al. (2005) who compared the evolution of more than one hundred replicate lines of the bacterium *Pseudomonas fluorescens,* over ∼900 generations in 15 environments of different complexity (**Table 1**). To do this, the authors used a synthetic growth medium that contained 1–8 single carbon substrates or specific combinations between them, and compared the genetic evolution of *P. fluorescens* lineages in simple and complex environmental growth media. In complex medium, the *Pseudomonas* lines evolved into several co-existing genotypes (adaptive radiation), exhibiting greater fitness for a wider range of carbon sources (but not for all), than the lines that evolved in simple environments. A higher fitness variance within populations selected in complex environments was thus observed. Indeed, lineages evolved in simple environment specialized in consuming a single carbon substrate, while those evolved in complex media were able to consume several substrates (but not all), without any appreciable loss of functions or apparent fitness costs. These results suggest that evolution in complex environments will lead to the emergence of imperfect generalist overlapping lines, adapted to a certain range of substrates but not to all (Barrett et al., 2005).

In nature, communities grow under conditions where many substrates are available, supporting a great number of consumer strategies for microorganisms. However, the availability of these substrates can be heterogeneous (both in space and time), leading to fluctuating selection for different genotypes within microbial populations, or to varying species within the microbial community. The presence of biotic interactions can also be a cause of increased environmental complexity (Brockhurst and Koskella, 2013). However, this environmental component was largely overlooked until more recently. A great example of how biotic interactions provide spatial and temporal heterogeneity is the plant rhizosphere (Kuzyakov and Razavi, 2019; **Figure 2** and **Box 1**).

EXPERIMENTAL EVOLUTION WITH COMPLEX PLANT-MICROBE SYSTEMS

Plant select microorganisms in the rhizosphere through their exudates (**Figure 2**). Interactions between plant and microorganisms can be beneficial, neutral or deleterious and several studies have used EE to unravel the relationships between the plant and plant-associated microorganisms (**Table 1**). Here we review how these EE assays have helped to better understand plant-microorganism's interactions, first for plant pathogens and secondly for plant-beneficial bacteria.

Plant Pathogens

Plant pathogens are pervasive, and their management is important for agriculture and food supplies, with direct impacts in human health and welfare. Thus, understanding the mechanisms underlying host co-evolution can help on devising new strategies to eradicate these types of pathogens. It is known that pathogen and plant defenses co-evolve, which translates into an arms race, where host and pathogen constantly evolve by mirroring the response of one another (Brockhurst and Koskella, 2013). Deciphering the genetic bases of pathogen adaptation is thus critical to understand disease emergence and acquisition of novel traits by pathogens when colonizing hosts (Toft and Andersson, 2010; Gandon et al., 2013).

EE approaches have been widely used to study the ecology of plant pathogens and dynamics of their adaptation when they interact with new host. Here, a focus is made on *in planta* EE experiments done with bacterial plant pathogens (**Table 1**), but others studies have also focused on fungal (Gilbert and Parker, 2010) or virus plant pathogens (Bedhomme et al., 2012).

Indeed, in an attempt to experimentally study the genetic basis of adaptation to new host, Guidot et al. (2014) performed EE with *Ralstonia solanacearum*, a plant pathogen with a continuously broadening host spectrum. A single clone of the model strain GMI1000 was inoculated on three native host plants (tomato, eggplant and pelargonium) where the pathogen causes disease and two distant plants (cabbage and bean) where it grows asymptomatically. The pathogen was transferred serially to the same plant line, in order to maintain the pathogen on the same host for 300 bacterial generations (26 serial passages). Although evolved strains showed an increase in competitiveness (pathogenicity) in both host and non-host plants, this increase was greater in the non-host plant (cabbage and beans), to which the pathogen was originally not adapted. This rapid evolution when colonizing a distant host plant tends then to decrease until reaching an optimum. This is known as "diminishing returns epistasis" that is often observed during EE studies with almost invariantly a reduction of adaptation speed and of mutations' fitness gain during the adaptation (Couce and Tenaillon, 2015).

Whole-genome analysis and comparison of the ancestral GMI1000 strain with nine evolved clones (three from the tomato host and six from the bean) highlighted that only few genes contribute to adaptation to a specific host. In particular, the transcriptional regulator encoding gene *efpR* was identified as important for the adaptation of *R. solanacearum* to bean. Evolved clones harboring *efpR* mutations had a greater competitiveness compared to the wild type clone during co-infection of bean plants (Guidot et al., 2014). EfpR was thereafter identified as a global catabolic repressor and regulator of virulence traits, whose

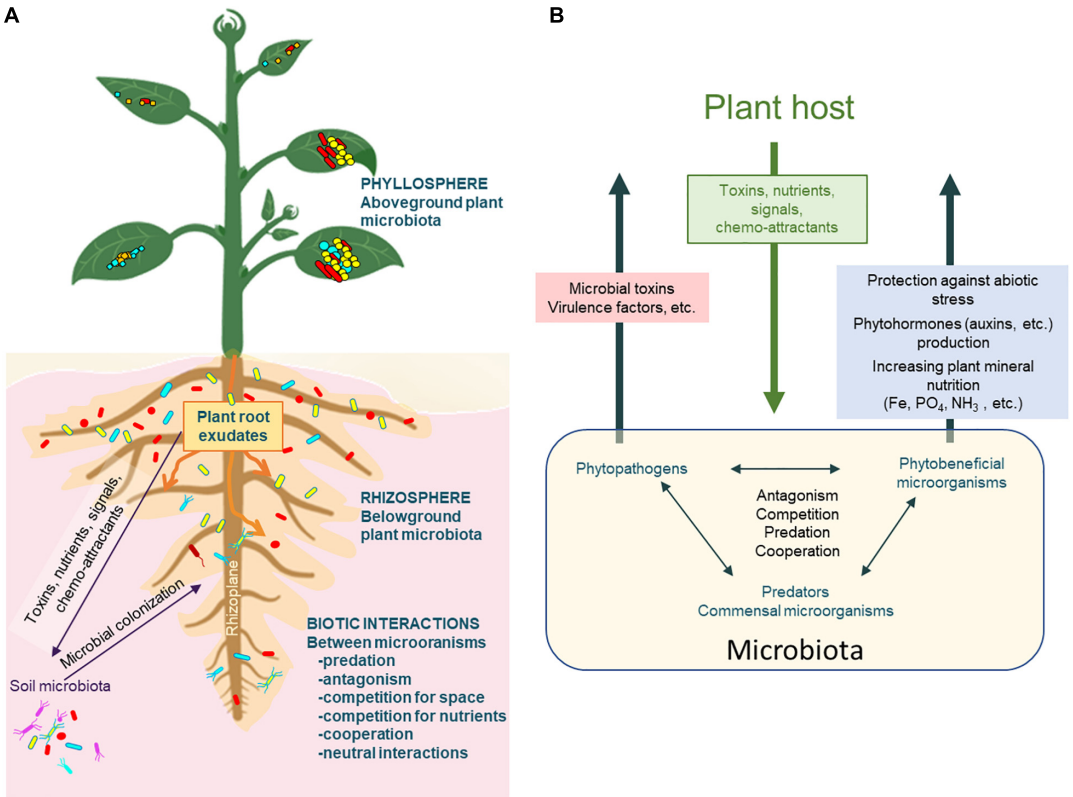


FIGURE 2 | Relationships between plant and microbial components. **(A)** Plants interact with a wide diversity of microorganisms (microbiota) in the different compartments of the plant system. The rhizosphere corresponds to the soil zone where the roots impact the soil organization and microbial functioning, while the rhizoplane is the interface between the root surface and the soil. The phyllosphere corresponds to the surface of all the aerial organs of the plant. Root exudates attract or repel soil microorganisms toward roots and increase the growth of myriads of microorganisms that will interact between each other (positive, negative or neutral biotic interactions). **(B)** Major biotic interactions in the rhizosphere include plant-microorganism interactions and microorganism/microorganism interactions, involving the exchange of different classes of molecules (hormones, toxins, virulence factors, signals, etc.).

BOX 1 | The plant rhizosphere, a rich and heterogeneous environment. Many soil microorganisms are capable of colonizing the plant rhizosphere, the volume of soil under the influence of root activities. In the rhizosphere plants exert a selection on soil microorganisms by secreting around their roots a wide range of compounds known as plant exudates (**Figure 2**). About 5–21% of the total carbon fixed by photosynthesis has been reported to be exuded at the root level and to influence the composition of the rhizosphere microbiota (Derrien et al., 2004; Vandenkoornhuyse et al., 2015). The exudates contain, among other things, nutrients but also signaling and chemo-attracting molecules. Along the root system, the exuded molecules are not evenly distributed. This leads to a heterogeneous root and rhizoplane colonization by selected microorganisms able to use the nutrients locally available. Some of these microorganisms will cooperate with the host plant through diverse mechanisms, like increasing the plant's mineral and nitrogen nutrition, synthesizing phytohormones influencing plant phenotypes or stimulating the induced systemic response (Vacheron et al., 2013; **Figure 2**). The nature of the exudates will impact the composition and functions of the root-associated microbiota which in turn will affect the physiology of the plant, generating new exudates (Vacheron et al., 2013; Chaparro et al., 2014). Thus, the rhizosphere is a discontinuous environment both in space and time, rich in nutrients, locally increasing the growth of highly diverse microbial populations that are ultimately in constant competition for nutrients. These competitive conditions make the rhizosphere a hotspot for horizontal gene transfer (HGT, van Elsas et al., 2003), mutations or gene arrangements, and will drive bacterial adaptation and acquisition of new genes (Wisniewski-Dyé et al., 2011). The latter will result in population genotype and phenotype changes, in the modification of the structure and function of the community and hence, will have profound effects on the long-term evolution of communities. Nevertheless, understanding how rhizosphere community members may evolve, at a species, sub-species or even individual cell level, is still a complex issue. Right now, rhizosphere natural communities and plant rhizosphere systems have rarely been included in EE studies. But, thanks to the development of single cells sequencing, considering such complex and heterogeneous living system in EE studies is becoming feasible and will allow to obtain a comprehensive detailed view of rhizosphere communities functioning and driving forces (**Figure 1**).

mutations allow an enhanced metabolic versatility and adaptation to host vascular tissues of new hosts (Perrier et al., 2016). This study highlighted the importance of combining EE with whole-genome sequencing in order to unravel the genetic basis of pathogen adaptation. More recently, transcriptomic analyses combined with genomic sequencing evidenced that epigenetic modifications also occur during the adaptive evolution of *R. solanacearum* to a tomato resistant cultivar, allowing expression changes of the EfpR regulon (Gopalan-Nair et al., 2020).

Another EE study focused on understanding how the prior evolutionary history of a pathogen affected subsequent evolution

in a new host. To this end, *Pseudomonas syringae*, a well-known plant pathogen, has been used to infect a distant host (*Arabidopsis*) or a native host (tomato) over several infection cycles, in the presence or absence of phages. Bacteriophages impose evolutionary trade-offs on the bacteria they infect, triggering modifications of bacterial surface receptors that subsequently may impact *P. syringae* interactions with distinct hosts. Selection on *Arabidopsis* (through serial passaging) leads to a larger increase in pathogen growth rate on both hosts, than selection on tomato. These results point out that a given association between a plant and a pathogen may affect the growth rate of the pathogen on other plants it will subsequently infect (Meaden and Koskella, 2017).

Other EE studies done with *Agrobacterium* have highlighted the appearance of cheaters in pathogen populations that will benefit of the production of virulence factors by the rest of the population without producing them themselves (Tannières et al., 2017). Indeed, the production and regulation of virulence factors are usually associated with high costs (energy or fitness) and are finely regulated by quorum sensing signals in *Agrobacterium*. Tannières et al. (2017) showed that cheater mutants that minimize the costs of expressing quorum sensing regulated functions spread during EE.

Non-pathogenic Plant Associated Microorganisms

EE studies have also investigated the genetic mechanisms underlying adaptation of bacteria that present mutualistic or synergistic plant-microbe interactions.

Fabaceae-rhizobium interactions are known as highly specific interactions between a bacterial symbiont and its host plant. In this case, the plant rewards cooperative symbionts in detriment of less mutualistic microbes but without necessarily punishing the latter (Batstone et al., 2017). Using a year-long EE experiment between *Ensifer meliloti* and *Medicago truncatula,* Batstone et al. (2020) have shown that local and recent adaptation of the symbiont to a plant genotype increases cooperation, independently of host selection.

Another EE experiment has investigated which bacterial genes facilitate symbiosis between bacteria and plant using a legume symbiont and a non-symbiotic bacterium. To do this, *R. solanacearum*, hosting the *Cupriavidus taiwanensis* symbiotic plasmid of *Mimosa pudica*, was repeatedly inoculated on the plant host (Marchetti et al., 2010; **Table 1**). After a series of *Mimosa pudica* infection cycles, an evolved symbiont well adapted to the host (i.e., able to induce nodulation and infect nodules, however, not to fix nitrogen) was obtained. Sequencing of intermediate and final forms revealed that the symbiosis-adaptive mutations happened in global regulatory proteins, leading to a reworking of the regulatory systems in *R. solanacearum*. These adaptive mutations included the inactivation of the type III secretion systems (the main virulence factor of *R. solanacearum*; Genin and Denny, 2012) and modifications on the expression of *efpR* (Guan et al., 2013; Capela et al., 2017). Genomic analyses revealed that is not only the *efpR* gene that is mutated but also its upstream region. Altogether these mutations led to metabolic and transcriptomic changes, allowing mutualistic interaction with the plant (Capela et al., 2017; Marchetti et al., 2017; Doin de Moura et al., 2020). These findings highlight how EE can help to identify the evolutionary pathways that drive the evolution of symbiotic functions in rhizobia, by understanding how adaptation modifies the regulatory systems that control virulence and determine the ecological functions of bacteria.

Other plant-beneficial bacteria are involved in less specific interactions with plant. These bacteria known as Plant Growth-Promoting Rhizobacteria (PGPR) colonize the plant rhizosphere (**Figure 2**). EE assays carried out *in planta* with plant-inoculated PGPR are uncommon. To our knowledge there is only one open resource work dealing with EE during plant-PGPR interaction (Li et al., 2020). In this work, a well-known biocontrol PGPR *Pseudomonas protegens* CHA0 was grown on the roots of *Arabidopsis thaliana* (on five independent Col-0 replicates) in gnotobiotic conditions, during 6 months (i.e., six one-month plant growth cycles). The bacterium was shown to evolve to a more mutualistic relationships with its plant, acquiring an improved competitiveness for root exudates, a better ability to tolerate plant-secreted antimicrobial compounds and a stronger positive effect on the plant performance. Different mutations in the key two-component regulator system GacS/GacA were recorded, conferring higher competitiveness to evolved CHA0 clones compared to the ancestral form in the presence of *A. thaliana* plants.

EXPERIMENTAL EVOLUTION STUDIES WITH SYNTHETIC AND NATURAL MICROBIAL COMMUNITIES

A large portion of studies on EE and microbial adaptation focuses on one single species, considering the genetic diversity within an evolving population rather than diversity between species within communities (e.g., Wiser et al., 2013; Kram et al., 2017; Lenski, 2017; Garoff et al., 2020). However, microorganisms are indeed in constant interaction with other organisms in their environment, so that any population in nature evolves with the other sympatric microbial populations.

Species interactions can influence how species evolve and adapt to environmental changes (Gorter et al., 2020). Unraveling the interactions that take place within a community is essential to understand how a community carries a function and how it will respond to perturbations and may evolve. For example, within a microbial community, some species can use the waste products generated by others (Lawrence et al., 2012; Seth and Taga, 2014; Piccardi et al., 2019). These interactions can become so relevant that a given microorganism can present a lower growth rate or even may not grow when cultivated alone, compared to when cultivated with another microorganism (Hillesland and Stahl, 2010) or within a community (Smith and Schuster, 2019). However, adaptation of one population to a new environment could be favored or counteracted by sympatric populations (Castledine et al., 2020). This process of reciprocal adaptation by interacting species is defined as co-evolution (Buckling and Rainey, 2002). Sympatric populations could be

locally maladapted and may show trait convergence for the use of same resource and these competitive interactions are specific to the co-evolved community members (Castledine et al., 2020).

EE With Synthetic Community in Environments of Increasing Complexity

Evolving together creates adaptive co-evolutionary dependencies: "it takes all the running you can do to stay in the same place," meaning that species have to constantly adapt to the other evolving species in order to survive/maintain their fitness (i.e., "Red Queen hypothesis"; Van Valen, 1977; Strotz et al., 2018). Species that co-evolved increased competitive interactions between them. Coevolution between competitors is expected to change species abundances within the community and affect subsequent community evolution (Nair et al., 2019; Scheuerl et al., 2019, 2020; **Figure 1**).

Thus, co-evolution experiments focused initially on microorganisms sharing negative interactions (i.e., predation, antagonism) (Nair et al., 2019; Scheuerl et al., 2019). Nowadays, co-evolution experiments also focus on mutualistic interactions, because these interactions are widespread in nature (Chomicki et al., 2020).

Several EE studies investigate the evolution of pairwise interactions between organisms (Nair et al., 2019; Scheuerl et al., 2019; Vidal and Segraves, 2020; Rodríguez-Verdugo and Ackermann, 2021). Here again, EE can be performed in constant environment (controlled microcosms with stable nutritive and physicochemical parameters), or in a fluctuating or complex environment (**Table 1**). Changes in the environment strongly constrain the existing biotic interactions between microorganisms. Rodríguez-Verdugo and Ackermann (2021) compared the evolution of co-cultures of *Acinetobacter johnsonii* and *Pseudomonas putida* in constant environment (the same nutrient medium at each cycle) to EE conducted in fluctuating environment (alternation of nutrient sources between each cycle). They showed that, the two species coexisted over 200 generations in the constant environment, whereas in the fluctuating environment, the extinction of one of the two partners was observed in half of the repetition, suggesting that the fluctuating environment destabilizes positive pairwise interactions.

However, natural microbiota are generally more complex and host a multitude of species of microorganisms, that will directly impact the evolutionary trajectory of each population (Jousset et al., 2016; Cira et al., 2018; Hall et al., 2020; Scheuerl et al., 2020), for example by suppressing competitors (Osmond and de Mazancourt, 2013), or by generating new niches (Calcagno et al., 2017). Species have thus to adapt not only to the environment itself but also to the biotic environment dynamically shaped by the other species (Ratzke and Gore, 2018; Scheuerl et al., 2020). Sharing the same ecological niche implies sharing some resources and this inevitably promotes competition within the group. Synthetic community approaches aim to mimic natural microbiota (Vorholt et al., 2017) and can help to better understand the functioning of biotic interaction network within natural communities (**Figure 1**). Cairns et al. (2018a) serially transferred a synthetic community of 33 bacterial strains on a complex liquid media. Over half of the strains from different species were lost in 16 days, after which the evolved community was relatively undisturbed until the end of the experiment (48 days). Within the evolved community, 14 strains co-existed with the predominance of three strains. The evolved synthetic community shared high diversity at different levels (i.e., taxonomic, metabolic, and functional levels) (**Table 1**).

Adaption to other competitive species may imply the production of antimicrobial metabolites, but a trade-off can arise due to the substantial energy-cost of their production (Yan et al., 2018). Some compounds can be costly to produce for one individual, but beneficial for all the members of the community. Microbes have thus developed multicellular cooperative behaviors, like biofilm formation and quorum sensing, along with nutrition acquisition, and the outcome of these interactions are referred as public goods (Besset-Manzoni et al., 2018; Smith and Schuster, 2019). Public goods take many forms from large proteins to small metabolites and can be actively or passively secreted, but one of their main features is that their benefit increases with population density.

Various EE studies comparing synthetic microbiota whose complexity increases (diversity and/or richness), increasing also the complexity of biotic interactions (competition for resources, cooperation, etc.), show that species adaptation largely depends on the community in which these species co-evolve. Thus, factors such as diversity or richness must be considered in microbial adaptation and evolution (Fiegna et al., 2015; Scheuerl et al., 2020). Another factor to consider in microbiota evolution is the impact of harsh conditions which may stimulate competition or cooperation. In a EE performed with synthetic microbiota, Scheuerl et al. (2020) adapted a continuous culture protocol and replaced the fresh medium addition at the end of each evolutionary cycle with the addition of only 10% of fresh medium, mimicking more natural conditions, and inducing competition and adaptation to recalcitrant carbon sources. This EE study revealed that the adaptation of bacteria to new environment is influenced by interspecies interactions.

Experimental evolution studies also enable to investigate the outcomes of interspecific interactions. For example, it is expected that plant root microbiota will be subjected to large environmental changes during plant development, which may lead to subsequent adaptation of the microbial community (**Box 1**). Specifically, a large number of natural compounds are exuded in the soil surrounding the roots, especially near the young parts, which modifies the physio-chemistry of the soil (oxygenation, pH, etc.), generating a stressful environment for the microbiota. Such stressful environment can result in competition or cooperation between members of the community. Using a synthetic community composed of four bacterial species, Piccardi et al. (2019) showed that the interactions between species evolved differently if the environmental stress level is low or high (stress gradient). Indeed, in a mild-stress environment, species evolved competitive interactions whereas, in harsh and toxic conditions, members of the community evolved cooperation or neutral behavior.

Unfortunately, cooperative behavior is not the only evolutionarily stable strategy as cheaters may appear and invade the community. Cheaters are non-cooperative individuals that benefit from the public goods without producing them. By not paying the cost of their production, cheaters could have more energy allocated to their growth, and therefore their relative fitness increases. The underlying mechanisms of cheaters' loss of function often involve a selective gene loss to optimize their adaptation to the environment (i.e., the "Black Queen Hypothesis," Mas et al., 2016). This evolutionary strategy has great impact on long term interactions because the fitness of cheaters depends on the public goods provided by the cooperating microbes. The stability of the community can only be maintained if the proportion of cheaters remains low within the community. Ultimately, cheaters will reduce the effective population size of the cooperating microbes, reducing the rate of public goods production, but also the rate at which beneficial mutations would arise in the community and the species would be able to adapt to a novel environment. The evolution of cooperation is a tricky issue since all microorganisms will tend, according to the natural selection theory, to maximize their own fitness. However, cooperative behavior is ubiquitous even though selfish interests have always been a source of conflict (Sachs and Wilcox, 2006; Burt and Trivers, 2009; Oliveira et al., 2014).

Batstone et al. (2020) investigated the prevalence of cheaters and non-cheaters in a simplified plant root community by performing EE with two strains of *Ensifer meliloti*. Both of these strains were able to receive carbon from the plant, but one strain lost the ability to fix N_2 to feed back to the plant (cheater), and another strain maintained that function (cooperative). After 1 year, the authors observed that the frequency of these two strains varied according to the level of coevolution with the plant. On one hand, in the initial stages of the EE, the cheater presented a twofold fitness advantage in host colonization, but by the end of the experiment it was extinct in the five plant genotypes tested. On the other hand, the N_2-fixing bacteria, which are less efficient in the first cycles of EE, became dominant at the end of EE. The interaction of evolved clones with the five *Medicago* lines showed that evolved clones achieved a higher fitness and provided greater benefits on the genotype with which they shared evolutionary history. Overall, this study suggests that cheaters are not able to outcompete cooperative genotypes, once they co-evolved with their host (Batstone et al., 2020).

Despite these recent studies, we still lack information on the eco-evolutionary dynamics behind cooperative and mutualistic behaviors. Moreover, deeper insight into biotic interactions within microbial communities is thus crucial to understand adaptation to novel environments. It can help in deciphering the dynamics of living systems and to predict responses to anthropogenic changes in the natural environment (Winder and Schindler, 2004; Davis et al., 2005; Berg et al., 2010).

EE With Natural Microbial Communities in Complex Environments

Numerous studies have taken an interest in plant-associated microbiota, illustrating the ability of the microorganisms to positively influence plant health and developmental traits such as disease resistance, herbivory, abiotic stress tolerance and growth (Mendes et al., 2011; Grunseich et al., 2020). During plant development, plant-microbiota interactions can change, which may in turn impact plant development (Swenson et al., 2000). For instance, an artificial microcosm selection carried on *A. thaliana* showed that plant biomass levels can be modified by soil microorganisms. Several *A. thaliana* seeds were inoculated with non-sterile soil batch and let grow for 35 days, which corresponds to one microcosm cycle. After each cycle, the plants presenting the highest and lowest biomasses were selected. The soil from these plants was retrieved and used to inoculate the next batch of plants. The experiment was carried for 16 microcosm cycles and both artificial selections (high or low biomass) were analyzed. *Arabidopsis* inoculated with the soil community from "high biomass" plants presented indeed higher biomasses than those inoculated with the soil community from "low biomass" plants. After 13 cycles, a soil analysis revealed different soil characteristics for high and low biomasses, that may reflect differences in the biotic components of the corresponding soils (Swenson et al., 2000; **Figure 1**).

Plant flowering is another trait that can be modulated by plant-associated microorganisms and the contributing microbiota's populations can be experimentally selected and enriched from one microcosm cycle to the next. In a similar way to Swenson and collaborators' experiment, several microcosms of *A. thaliana* Col0 were created with seeds placed on sterile soil. In this experiment, microcosms were selected for either an early or late flowering phenotype. Plants were harvested and the soil retrieved as soon as all the plants of the microcosm flowered, therefore, the duration of the microcosm cycle depended on the flowering time. After ten cycles on *A. thaliana* Col0, microorganisms were retrieved and inoculated on other *A. thaliana* genotypes and on a related crucifer, *Brassica rapa*. Plant-associated microorganisms induced on these plants the same phenotype (early flowering) as the one induced on the genotype (Col0) used for selecting the evolved community (Panke-Buisse et al., 2015). Rather than imposing no selective pressure on the plant host, in these studies, researchers selected a particular plant trait (biomass or flowering) to select plant-associated microorganisms contributing to this plant trait (**Table 1**). In a similar approach, Lu et al. (2018) selected evolved microbiota capable of either stimulating precocious flowering or delaying it. By studying the molecular interactions between root exudates and the microbiota, a new network of molecular interactions was established, linking the production of auxin phytohormone (i.e., indole-3-acetic acid) from tryptophan by the microbiota, the nitrogen cycling and the timing of flowering in the host plant. EE on natural communities thus enabled to document novel metabolic networks in which soil microbiota influenced plant flowering time, thus shedding light on the key role of soil microbiota on plant functioning.

To better understand the holobiont's mechanistic functioning, an EE study was recently done without taking into consideration improved plant health, growth or development phenotype as a selective pressure outcome of the evolutive process (Morella et al., 2020). The authors collected the leaf microbiota (phyllosphere) of field-grown tomato plants in order to spray it on fresh tomato

plants. After 10 days of growth, the phyllosphere microbiota were sampled again, and used to inoculate a new round of plants. A total of four passages was done. They compare the impact of five genotypes differing in disease resistance genes in the selection of the tomato's phyllosphere microbiota. They evidenced a strong selection of a stable microbiota adapted to this ecological niche, with a significant driven effect of the tomato genotype. Contrariwise to the initial leaf microbiota that was unstable, the evolved microbiota became well adapted to its host and robust to the invasion of the initial community (Morella et al., 2020).

In the rhizosphere, networks of within-microbiota interactions may also shape the evolution and stability of the community as a whole and the observed effects on the plant, affecting its development, health, and response to abiotic and biotic stresses (**Figure 2**). Deciphering the principles that underlying ecological and evolutionary properties of microbial communities can allow to build predictive models of ecological dynamics of microbial communities.

CONCLUSION AND FUTURE AVENUES FOR EE STUDIES

In natural conditions, microbes often coevolve within an interspecific community that may interact with other non-microbial organisms, such as plants. Understanding what drives all the different types of interactions with and within microbial communities, and, importantly, how they co-evolved, will allow to draw predictions on how microbial communities and their hosts respond to environmental changes.

EE assays have historically mostly focused on low complexity systems, allowing to understand the evolutionary adaptation of microorganisms to stressful conditions. EE studies combined with whole-genome sequencing have allowed us to understand the genetic bases separating an evolved clone from its original ancestor strain in a wide range of situations and to reveal the genetic and functional networks involved in microbial adaptation (Segrè et al., 2006; Bailey and Bataillon, 2016; **Figure 1**). To date, there have been very few EE studies done on more complex systems using natural microbial communities (**Table 1**). Indeed, whole genome sequencing of many individual clones from one population or from several populations evolving together is a big technological barrier to solve (**Figure 1**).

The combination of recent technological advances on meta-omic approaches, cell sorting and single-cell sequencing, will soon allow to investigate more deeply the genetic mechanisms underlying species-specific adaptation within microbial communities evolving in complex and heterogeneous environments, like the rhizosphere (**Box 1** and **Table 1**). Genetic changes, random genetic drift and natural selection operate on each community member leading to the fixation of mutations, hence altering the genetic composition of populations and indirectly affecting species interactions that dictate community ecology. The intersection of ecology and evolution is key to understand microbial communities.

AUTHOR CONTRIBUTIONS

BM wrote the first version of the manuscript and figures. BM, DM, and CP-C corrected and improved the manuscript. All authors have read and agreed to the submitted version of the manuscript.

FUNDING

BM was supported by a Ph.D. fellowship from the French Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche. The authors acknowledge funding from the French National Programme EC2CO "Structuring initiative Continental and Coastal Ecosphere"—Microbien 2018–2019.

REFERENCES

Adams, J., and Rosenzweig, F. (2014). Experimental microbial evolution: history and conceptual underpinnings. *Genomics* 104, 393–398. doi: 10.1016/j.ygeno.2014.10.004

Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., et al. (2015). Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* 528, 364–369. doi: 10.1038/nature16192

Bailey, S. F., and Bataillon, T. (2016). Can the experimental evolution programme help us elucidate the genetic basis of adaptation in nature? *Mol. Ecol.* 25, 203–218. doi: 10.1111/mec.13378

Barrett, R. D. H., MacLean, R. C., and Bell, G. (2005). Experimental evolution of *Pseudomonas fluorescens* in simple and complex environments. *Am. Nat.* 166, 470–480. doi: 10.1086/444440

Barrick, J. E., Yun, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., et al. (2009). Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461, 1243–1247. doi: 10.1038/nature08480

Batstone, R. T., Dutton, E. M., Wang, D., Yang, M., and Frederickson, M. E. (2017). The evolution of symbiont preference traits in the model legume *Medicago truncatula*. *New Phytol.* 213, 1850–1861. doi: 10.1111/nph.14308

Batstone, R. T., O'Brien, A. M., Harrison, T. L., and Frederickson, M. E. (2020). Experimental evolution makes microbes more cooperative with their local host genotype. *Science* 370, 476–478. doi: 10.1126/science.abb7222

Bedhomme, S., Lafforgue, G., and Elena, S. F. (2012). Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations. *Mol. Biol. Evol.* 29, 1481–1492. doi: 10.1093/molbev/msr314

Berg, M. P., Kiers, E. T., Driessen, G., van der Heijden, M., Kooi, B. W., Kuenen, F., et al. (2010). Adapt or disperse: understanding species persistence in a changing world. *Glob. Change Biol.* 16, 587–598. doi: 10.1111/j.1365-2486.2009.02014.x

Besset-Manzoni, Y., Rieusset, L., Joly, P., Comte, C., and Prigent-Combaret, C. (2018). Exploiting rhizosphere microbial cooperation for developing sustainable agriculture strategies. *Environ. Sci. Pollut. Res.* 25, 29953–29970. doi: 10.1007/s11356-017-1152-2

Blount, Z., Barrick, J., Davidson, C., and Lenski, R. E. (2012). Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 489, 513–518. doi: 10.1038/nature11514

Bouffaud, M. L., Renoud, S., Moënne-Loccoz, Y., and Muller, D. (2016). Is plant evolutionary history impacting recruitment of diazotrophs and nifH expression in the rhizosphere? *Sci. Rep.* 6:21690. doi: 10.1038/srep21690

Brockhurst, M. A., and Koskella, B. (2013). Experimental coevolution of species interactions. *Trends Ecol. Evol.* 28, 367–375. doi: 10.1016/j.tree.2013.02.009

Brockhurst, M. A., Morgan, A. D., Rainey, P. B., and Buckling, A. (2003). Population mixing accelerates coevolution: population mixing accelerates coevolution. *Ecol. Lett.* 6, 975–979.

Bruto, M., Prigent-Combaret, C., Muller, D., and Moënne-Loccoz, Y. (2014). Analysis of genes contributing to plant-beneficial functions in plant growth-promoting rhizobacteria and related *Proteobacteria*. *Sci. Rep.* 4:6261.

Buckling, A., and Rainey, P. B. (2002). Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. R. Soc. Lond. B.* 269, 931–936. doi: 10.1098/rspb.2001.1945

Buckling, A., Wills, M. A., and Colegrave, N. (2003). Adaptation limits diversification of experimental bacterial populations. *Science* 302, 2107–2109. doi: 10.1126/science.1088848

Burt, A., and Trivers, R. (2009). *Genes in Conflict: The Biology of Selfish Genetic Elements*. Cambridge, MA: Harvard University Press.

Cairns, J., Jokela, R., Hultman, J., Tamminen, M., Virta, M., and Hiltunen, T. (2018a). Construction and characterization of synthetic bacterial community for experimental ecology and evolution. *Front. Genet.* 9:312. doi: 10.3389/fgene.2018.00312

Cairns, J., Ruokolainen, L., Hultman, J., Tamminen, M., Virta, M., and Hiltunen, T. (2018b). Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Commun. Biol.* 1:35. doi: 10.1038/s42003-018-0041-7

Calcagno, V., Jarne, P., Loreau, M., Mouquet, N., and David, P. (2017). Diversity spurs diversification in ecological communities. *Nat. Commun.* 8:15810.

Capela, D., Marchetti, M., Clérissi, C., Perrier, A., Guetta, D., Gris, C., et al. (2017). Recruitment of a lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen experimentally evolved into a legume symbiont. *Mol. Biol. Evol.* 34, 2503–2521. doi: 10.1093/molbev/msx165

Castledine, M., Padfield, D., and Buckling, A. (2020). Experimental (co)evolution in a multi-species microbial community results in local maladaptation. *Ecol. Lett.* 23, 1673–1681. doi: 10.1111/ele.13599

Chaparro, J. M., Badri, D. V., and Vivanco, J. M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* 8, 790–803.

Chomicki, G., Kiers, E. T., and Renner, S. S. (2020). The evolution of mutualistic dependence. *Annu. Rev. Ecol. Evol. Syst.* 51, 409–432.

Cira, N. J., Pearce, M. T., and Quake, S. R. (2018). Neutral and selective dynamics in a synthetic microbial community. *Proc. Natl. Acad. Sci. U.S.A.* 115, E9842–E9848. doi: 10.1073/pnas.1808118115

Colosimo, P. F. (2005). Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307, 1928–1933. doi: 10.1126/science.1107239

Consuegra, J., Gaffé, J., Lenski, R. E., Hindré, T., Barrick, J. E., Tenaillon, O., et al. (2021). Insertion-sequence-mediated mutations both promote and constrain evolvability during a long-term experiment with bacteria. *Nat. Commun.* 12:980. doi: 10.1038/s41467-021-21210-7

Cooper, T. F., Rozen, D. E., and Lenski, R. E. (2003). Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1072–1077. doi: 10.1073/pnas.0334340100

Cooper, V. S., and Lenski, R. E. (2000). The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* 407, 736–739. doi: 10.1038/35037572

Couce, A., and Tenaillon, O. A. (2015). The rule of declining adaptability in microbial evolution experiments. *Front. Genet.* 6:99. doi: 10.3389/fgene.2015.00099

Davis, M. B., Shaw, R. G., and Etterson, J. R. (2005). Evolutionary responses to changing climate. *Ecology* 86, 1704–1714. doi: 10.1890/03-0788

Derrien, D., Marol, C., and Balesdent, J. (2004). The dynamics of neutral sugars in the rhizosphere of wheat. An approach by 13C pulse-labelling and GC/C/IRMS. *Plant Soil* 267, 243–253. doi: 10.1007/s11104-005-5348-8

Doin de Moura, G. G., Remigi, P., Masson-Boivin, C., and Capela, D. (2020). Experimental evolution of legume symbionts: What have we learnt? *Genes* 11:339. doi: 10.3390/genes11030339

Elena, S. F., and Lenski, R. E. (2003). Microbial genetics: evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4, 457–469. doi: 10.1038/nrg1088

Fiegna, F., Moreno-Letelier, A., Bell, T., and Barraclough, T. G. (2015). Evolution of species interactions determines microbial community productivity in new environments. *ISME J.* 9, 1235–1245. doi: 10.1038/ismej.2014.215

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

Flohr, R. C., Blom, C. J., Rainey, P. B., and Beaumont, H. J. (2013). Founder niche constrains evolutionary adaptive radiation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 20663–20668. doi: 10.1073/pnas.1310310110

Gandon, S., Hochberg, M. E., Holt, R. D., and Day, T. (2013). What limits the evolutionary emergence of pathogens? *Philos. Trans. R. Soc. B* 368:20120086. doi: 10.1098/rstb.2012.0086

Garoff, L., Pietsch, F., Huseby, D. L., Lilja, T., Brandis, G., and Hughes, D. (2020). Population bottlenecks strongly influence the evolutionary trajectory to fluoroquinolone resistance in *Escherichia coli*. *Mol. Biol. Evol.* 37, 1637–1646. doi: 10.1093/molbev/msaa032

Genin, S., and Denny, T. P. (2012). Pathogenomics of the *Ralstonia solanacearum* species complex. *Annu. Rev. Phytopat.* 50, 67–89. doi: 10.1146/annurev-phyto-081211-173000

Gerrish, P. J., and Lenski, R. E. (1998). The fate of competing beneficial mutations in an asexual population. *Mutation Evol.* 7, 12–144. doi: 10.1007/978-94-011-5210-5_12

Gilbert, G. S., and Parker, I. M. (2010). Rapid evolution in a plant-pathogen interaction and the consequences for introduced host species: pathogen evolution and novel hosts. *Evol. Appl.* 3, 144–156. doi: 10.1111/j.1752-4571.2009.00107.x

Gopalan-Nair, R., Jardinaud, M. F., Legrand, L., Landry, D., Barlet, X., Lopez-Roques, C., et al. (2020). Convergent rewiring of the virulence regulatory network promotes adaptation of *Ralstonia solanacearum* on resistant tomato. *Mol; Biol. Evol.* doi: 10.1093/molbev/msaa320 [Epub ahead of print].

Gorter, F. A., Manhart, M., and Ackermann, M. (2020). Understanding the evolution of interspecies interactions in microbial communities. *Philos. Trans. R. Soc. B.* 375:20190256. doi: 10.1098/rstb.2019.0256

Gresham, D., and Hong, J. (2015). The functional basis of adaptive evolution in chemostats. *FEMS Microbiol. Rev.* 39, 2–16. doi: 10.1111/1574-6976.12082

Grunseich, J. M., Thompson, M. N., Aguirre, N. M., and Helms, A. M. (2020). The role of plant-associated microbes in mediating host-plant selection by insect herbivores. *Plants* 9:6. doi: 10.3390/plants9010006

Guan, S. H., Gris, C., Cruveiller, S., Pouzet, C., Tasse, L., Leru, A., et al. (2013). Experimental evolution of nodule intracellular infection in legume symbionts. *ISME J.* 7, 1367–1377. doi: 10.1038/ismej.2013.24

Gude, S., Pinçe, E., Taute, K. M., Seinen, A. B., Shimizu, T. S., and Tans, S. J. (2020). Bacterial coexistence driven by motility and spatial competition. *Nature* 578, 588–592. doi: 10.1038/s41586-020-2033-2

Guidot, A., Jiang, W., Ferdy, J.-B., Thébaud, C., Barberis, P., Gouzy, J., et al. (2014). Multihost experimental evolution of the pathogen *Ralstonia solanacearum* unveils genes involved in adaptation to plants. *Mol. Biol. Evol.* 31, 2913–2928. doi: 10.1093/molbev/msu229

Hall, A. R., Ashby, B., Bascompte, J., and King, K. C. (2020). Measuring coevolutionary dynamics in species-rich communities. *Trends Ecol. Evol.* 35, 539–550. doi: 10.1016/j.tree.2020.02.002

Hansen, S. K., Haagensen, J. A., Gjermansen, M., Jørgensen, T. M., Tolker-Nielsen, T., and Molin, S. (2007). Characterization of a *Pseudomonas putida* rough variant evolved in a mixed-species biofilm with *Acinetobacter* sp. strain C6. *J. Bacteriol.* 189, 4932–4943. doi: 10.1128/JB.00041-07

Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. (1999). From molecular to modular cell biology. *Nature* 402, C47–C52. doi: 10.1038/35011540

Herring, C. D., Raghunathan, A., Honisch, C., Patel, T., Applebee, M. K., Joyce, A. R., et al. (2006). Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat. Genet.* 38, 1406–1412. doi: 10.1038/ng1906

Hillesland, K. L., and Stahl, D. A. (2010). Rapid evolution of stability and productivity at the origin of a microbial mutualism. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2124–2129. doi: 10.1073/pnas.0908456107

Jousset, A., Eisenhauer, N., Merker, M., Mouquet, N., and Scheu, S. (2016). High functional diversity stimulates diversification in experimental microbial communities. *Sci. Adv.* 2, e1600124. doi: 10.1126/sciadv.1600124

Jerison, E. R., and Desai, M. M. (2015). Genomic investigations of evolutionary dynamics and epistasis in microbial evolution experiments. *Curr. Opin. Genet Dev.* 35, 33–39. doi: 10.1016/j.gde.2015.08.008

Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M. C. (2012). Experimental evolution. *Trends Ecol. Evol.* 27, 547–560. doi: 10.1016/j.tree.2012.06.001

Kayser, J., Schreck, C. F., Yu, Q., Gralka, M., and Hallatschek, O. (2018). Emergence of evolutionary driving forces in pattern-forming microbial populations. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373, 20170106. doi: 10.1098/rstb.2017.0106

Koza, A., Kusmierska, A., McLaughlin, K., Moshynets, O., and Spiers, A. J. (2017). Adaptive radiation of *Pseudomonas fluorescens* SBW25 in experimental microcosms provides an understanding of the evolutionary ecology and molecular biology of A-L interface biofilm formation. *FEMS Microbiol. Lett.* 364:fnx109. doi: 10.1093/femsle/fnx109

Kram, K. E., Geiger, C., Ismail, W. M., Lee, H., Tang, H., Foster, P. L., et al. (2017). Adaptation of *Escherichia coli* to long-term Serial passage in complex medium: evidence of parallel evolution. *Systems* 2:e00192-16. doi: 10.1128/mSystems.00192-16

Kuzyakov, Y., and Razavi, B. S. (2019). Rhizosphere size and shape: temporal dynamics and spatial stationarity. *Soil Biol. Biochem.* 135, 343–360. doi: 10.1016/j.soilbio.2019.05.011

Lang, G. I., Botstein, D., and Desai, M. M. (2011). Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* 188, 647–661. doi: 10.1534/genetics.111.128942

Lawrence, D., Fiegna, F., Behrends, V., Bundy, J. G., Phillimore, A. B., Bell, T., et al. (2012). Species interactions alter evolutionary responses to a novel environment. *PLoS Biol.* 10:e1001330. doi: 10.1371/journal.pbio.1001330

Lenski, R. E. (2017). Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *ISME J.* 11, 2181–2194. doi: 10.1038/ismej.2017.69

Lenski, R. E., Rose, M. R., Simpson, S. C., and Tadler, S. C. (1991). Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am. Nat.* 138, 1315–1341. doi: 10.1086/285289

Lenski, R. E., and Travisano, M. (1994). Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc. Natl. Acad. Sci. U.S.A.* 91, 6808–6814. doi: 10.1073/pnas.91.15.6808

Lescat, M., Launay, A., Ghalayini, M., Magnan, M., Glodt, J., Pintard, C., et al. (2017). Using long-term experimental evolution to uncover the patterns and determinants of molecular evolution of an *Escherichia coli* natural isolate in the streptomycin-treated mouse gut. *Mol. Ecol.* 26, 1802–1817. doi: 10.1111/mec.13851

Li, E., Jiang, H., Pieterse, C. M. J., Jousset, A., Bakker, P. A. H. M., and de Jonge, R. (2020). Experimental evolution-driven identification of Arabidopsis rhizosphere competence genes in *Pseudomonas* protegens. *bioRxiv* [Preprint]. doi: 10.1101/2020.12.01.407551

Little, A. E., Robinson, C. J., Peterson, S. B., Raffa, K. F., and Handelsman, J. (2008). Rules of engagement: interspecies interactions that regulate microbial communities. *Annu. Rev. Microbiol.* 62, 375–401. doi: 10.1146/annurev.micro.030608.101423

Lu, T., Ke, M., Lavoie, M., Jin, Y., Fan, X., and Zhang, Z. (2018). Rhizosphere microorganisms can influence the timing of plant flowering. *Microbiome* 6:231. doi: 10.1186/s40168-018-0615-0

Maharjan, R., Seeto, S., Notley-McRobb, L., and Ferenci, T. (2006). Clonal adaptive radiation in a constant environment. *Science* 313, 514–517. doi: 10.1126/science.1129865

Marchetti, M., Capela, D., Glew, M., Cruveiller, S., Chane-Woon-Ming, B., Gris, C., et al. (2010). Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* 8:e1000280. doi: 10.1371/journal.pbio.1000280

Marchetti, M., Clerissi, C., Yousfi, Y., Gris, C., Bouchez, O., Rocha, E., et al. (2017). Experimental evolution of rhizobia may lead to either extra- or intracellular symbiotic adaptation depending on the selection regime. *Mol. Ecol.* 26, 1818–1831. doi: 10.1111/mec.13895

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

McDonald, M. J. (2019). Microbial experimental evolution - a proving ground for evolutionary theory and a tool for discovery. *EMBO Rep.* 20:e46992. doi: 10.15252/embr.201846992

Meaden, S., and Koskella, B. (2017). Adaptation of the pathogen, *Pseudomonas syringae*, during experimental evolution on a native vs. alternative host plant. *Mol. Ecol.* 26, 1790–1801. doi: 10.1111/mec.14060

Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J. H. M., et al. (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science* 332, 1097–1100. doi: 10.1126/science.1203980

Morella, N. M., Weng, F. C., 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. U.S.A.* 117, 1148–1159. doi: 10.1073/pnas.1908600116

Nair, R. R., Vasse, M., Wielgoss, S., Sun, L., Yu, Y. N., and Velicer, G. J. (2019). Bacterial predator-prey coevolution accelerates genome evolution and selects on virulence-associated prey defences. *Nat. Commun.* 10:4301. doi: 10.1038/s41467-019-12140-6

Oliveira, N. M., Niehus, R., and Foster, K. R. (2014). Evolutionary limits to cooperation in microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17941–17946. doi: 10.1073/pnas.1412673111

Osmond, M. M., and de Mazancourt, C. (2013). How competition affects evolutionary rescue. *Philos. Trans. R. Soc. Lond. B* 368:20120085. doi: 10.1098/rstb.2012.0085

Panke-Buisse, K., Poole, A. C., Goodrich, J. K., Ley, R. E., and Kao-Kniffin, J. (2015). Selection on soil microbiomes reveals reproducible impacts on plant function. *ISME J.* 9, 980–989. doi: 10.1038/ismej.2014.196

Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A. J., Thomson, N. R., et al. (2010). Antagonistic coevolution accelerates molecular evolution. *Nature* 464, 275–278. doi: 10.1038/nature08798

Perrier, A., Peyraud, R., Rengel, D., Barlet, X., Lucasson, E., Gouzy, J., et al. (2016). Enhanced in planta fitness through adaptive mutations in EfpR, a dual regulator of virulence and metabolic functions in the plant pathogen *Ralstonia solanacearum*. *PLoS Pathog.* 12:e1006044. doi: 10.1371/journal.ppat.1006044

Piccardi, P., Vessman, B., and Mitri, S. (2019). Toxicity drives facilitation between 4 bacterial species. *Proc. Natl. Acad. Sci. U.S.A.* 116, 15979–15984. doi: 10.1073/pnas.1906172116

Pickersgill, B. (2018). Parallel vs. convergent evolution in domestication and diversification of crops in the Americas. *Front. Ecol. Evol.* 6:56. doi: 10.3389/fevo.2018.00056

Rainey, P. B., and Quistad, S. D. (2020). Toward a dynamical understanding of microbial communities. *Philos. Trans. R. Soc. B* 375:20190248. doi: 10.1098/rstb.2019.0248

Ratcliff, W. C., Denison, R. F., Borrello, M., and Travisano, M. (2012). Experimental evolution of multicellularity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1595–1600. doi: 10.1073/pnas.1115323109

Ratzke, C., and Gore, J. (2018). Modifying and reacting to the environmental pH can drive bacterial interactions. *PLoS Biol.* 16:e2004248. doi: 10.1371/journal.pbio.2004248

Rendueles, O., and Velicer, G. J. (2017). Evolution by flight and fight: diverse mechanisms of adaptation by actively motile microbes. *ISME J.* 11, 555–568. doi: 10.1038/ismej.2016.115

Rendueles, O., and Velicer, G. J. (2020). Hidden paths to endless forms most wonderful: complexity of bacterial motility shapes diversification of latent phenotypes. *BMC Evol. Biol.* 20:145. doi: 10.1186/s12862-020-01707-3

Renoud, S., Bouffaud, M. L., Dubost, A., Prigent-Combaret, C., Legendre, L., Moënne-Loccoz, Y., et al. (2020). Co-occurrence of rhizobacteria with nitrogen fixation and/or 1-aminocyclopropane-1-carboxylate deamination abilities in the maize rhizosphere. *FEMS Microbiol. Ecol.* 96:fiaa062. doi: 10.1093/femsec/fiaa062

Rodríguez-Verdugo, A., and Ackermann, M. (2021). Rapid evolution destabilizes species interactions in a fluctuating environment. *ISME J.* 15, 450–460. doi: 10.1038/s41396-020-00787-9

Sachs, J. L., and Wilcox, T. P. (2006). A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proc. R. Soc. B Biol. Sci.* 273, 425–429. doi: 10.1098/rspb.2005.3346

Scanlan, P. D., Hall, A. R., Lopez-Pascua, L. D. C., and Buckling, A. (2011). Genetic basis of infectivity evolution in a bacteriophage. *Mol. Ecol.* 20, 981–989. doi: 10.1111/j.1365-294X.2010.04903.x

Scheuerl, T., Cairns, J., Becks, L., and Hiltunen, T. (2019). Predator coevolution and prey trait variability determine species coexistence. *Proc. R. Soc. B* 286:20190245. doi: 10.1098/rspb.2019.0245

Scheuerl, T., Hopkins, M., Nowell, R. W., Rivett, D. W., Barraclough, T. G., and Bell, T. (2020). Bacterial adaptation is constrained in complex communities. *Nat. Commun.* 11:754. doi: 10.1038/s41467-020-14570-z

Schluter, D., Clifford, E., Nemethy, M., and Mckinnon, J. (2004). Parallel evolution and inheritance of quantitative traits. *Am. Nat.* 163, 809–822. doi: 10.1086/383621

Segrè, A. V., Murray, A. W., and Leu, J.-Y. (2006). High-resolution mutation mapping reveals parallel experimental evolution in yeast. *PLoS Biol.* 4:e256. doi: 10.1371/journal.pbio.0040256

Seth, E. C., and Taga, M. E. (2014). Nutrient cross-feeding in the microbial world. *Front. Microbiol.* 5:350. doi: 10.3389/fmicb.2014.00350

Smith, P., and Schuster, M. (2019). Public goods and cheating in microbes. *Curr. Biol.* 29, R442–R447. doi: 10.1016/j.cub.2019.03.001

Strotz, L. C., Simões, M., Girard, M. G., Breitkreuz, L., Kimmig, J., and Lieberman, B. S. (2018). Getting somewhere with the Red Queen: chasing a biologically modern definition of the hypothesis. *Biol. Lett.* 14, 20170734. doi: 10.1098/rsbl.2017.0734

Swenson, W., Wilson, D. S., and Elias, R. (2000). Artificial ecosystem selection. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9110–9114. doi: 10.1073/pnas.150237597

Tannières, M., Lang, J., Barnier, C., Shykoff, J. A., and Faure, D. (2017). Quorum-quenching limits quorum-sensing exploitation by signal-negative invaders. *Sci. Rep.* 7:40126. doi: 10.1038/srep40126

Tenaillon, O. (2014). The utility of fisher's geometric model in evolutionary genetics. *Annu. Rev. Ecol. Evol. Syst.* 45, 179–201. doi: 10.1146/annurev-ecolsys-120213-091846

Toft, C., and Andersson, S. G. E. (2010). Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat. Rev. Genet.* 11, 465–475. doi: 10.1038/nrg2798

Travisano, M., Mongold, J., Bennett, A., and Lenski, R. (1995). Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* 267, 87–90. doi: 10.1126/science.7809610

Treves, D. S., Manning, S., and Adams, J. (1998). Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Mol. Biol. Evol.* 15, 789–797. doi: 10.1093/oxfordjournals.molbev.a025984

Tso, G. H. W., Reales-Calderon, J. A., Tan, A. S. M., Sem, X., Le, G. T. T., Tan, T. G., et al. (2018). Experimental evolution of a fungal pathogen into a gut symbiont. *Science* 362, 589–595. doi: 10.1126/science.aat0537

Vacheron, J., Desbrosses, G., Bouffaud, M. L., Touraine, B., Moënne-Loccoz, Y., Muller, D., et al. (2013). Plant growth-promoting rhizobacteria and root system functioning. *Front. Plant Sci.* 4:356. doi: 10.3389/fpls.2013.00356

Vacheron, J., Desbrosses, G., Renoud, S., Padilla, R., Walker, V., Muller, D., et al. (2018). Differential contribution of plant-beneficial functions from *Pseudomonas kilonensis* F113 to root system architecture alterations in *Arabidopsis thaliana* and *Zea mays*. *Mol. Plant Microbe Interact.* 31, 212–223. doi: 10.1094/MPMI-07-17-0185-R

Van den Bergh, B., Swings, T., Fauvart, M., and Jan, M. (2018). Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiol. Mol. Biol. Rev.* 82:e00008-18. doi: 10.1128/MMBR.00008-18

van Elsas, J. D., Turner, S., and Bailey, M. J. (2003). Horizontal gene transfers in the phytosphere. *New Phytol.* 157, 525–537. doi: 10.1046/j.1469-8137.2003.00697.x

Van Valen, L. (1977). The red queen. *Am. Nat.* 111, 809–810. doi: 10.1086/283213

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.

Vidal, M. C., and Segraves, K. A. (2020). Coevolved mutualists experience fluctuating costs and benefits over time. *Evolution* 75, 219–230. doi: 10.1111/evo.14155

Vorholt, J. A., Vogel, C., Carlström, C. I., and Müller, D. B. (2017). Establishing causality: opportunities of synthetic communities for plant microbiome research. *Cell Host Microbe* 22, 142–155. doi: 10.1016/j.chom.2017.07.004

Westphal, L. L., Lau, J., Negro, Z., Moreno, I. J., Ismail Mohammed, W., Lee, H., et al. (2018). Adaptation of *Escherichia coli* to long-term batch culture in various rich media. *Res. Microbiol.* 169, 145–156. doi: 10.1016/j.resmic.2018.01.003

Winder, M., and Schindler, D. E. (2004). Climate change uncouples trophic interactions in an aquatic ecosystem. *Ecology* 85, 2100–2106. doi: 10.1890/04-0151

Wiser, M. J., Ribeck, N., and Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science* 342, 1364–1367. doi: 10.1126/science.1243357

Wisniewski-Dyé, F., Borziak, K., Khalsa-Moyers, G., Alexandre, G., Sukharnikov, L. O., Wuichet, K., et al. (2011). Azospirillum genomes reveal transition of bacteria from aquatic to terrestrial environments. *PLoS Genet.* 7:e1002430. doi: 10.1371/journal.pgen.1002430

Woods, L. C., Gorrell, R. J., Taylor, F., Connallon, T., Kwok, T., and McDonald, M. J. (2020). Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. *Proc. Natl. Acad. Sci. U.S.A.* 117, 26868–26875. doi: 10.1073/pnas.2005331117

Yan, Q., Lopes, L. D., Shaffer, B. T., Kidarsa, T. A., Vining, O., Philmus, B., et al. (2018). Secondary metabolism and interspecific competition affect accumulation of spontaneous mutants in the GacS-GacA regulatory system in *Pseudomonas protegens*. *mBio* 9:e01845-17. doi: 10.1128/mBio.01845-17

Yang, L., Jelsbak, L., Marvig, R. L., Damkiaer, S., Workman, C. T., Rau, M. H., et al. (2011). Evolutionary dynamics of bacteria in a human host environment. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7481–7486. doi: 10.1073/pnas.1018249108

Yeh, P., Tschumi, A. I., and Kishony, R. (2006). Functional classification of drugs by properties of their pairwise interactions. *Nat. Genet.* 38, 489–494. doi: 10.1038/ng1755

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

Zhang, Z., Gosset, G., Barabote, R., Gonzalez, C. S., Cuevas, W. A., and Saier, M. H. (2005). Functional interactions between the carbon and iron utilization regulators, Crp and Fur, in *Escherichia coli*. *J. Bacteriol.* 187, 980–990. doi: 10.1128/JB.187.3.980-990.2005

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 28 May 2021
doi: 10.3389/fmicb.2021.655903

Lead Drives Complex Dynamics of a Conjugative Plasmid in a Bacterial Community

Valentine Cyriaque[1,2], Jonas Stenløkke Madsen[2], Laurence Fievez[3], Baptiste Leroy[1], Lars H. Hansen[4], Fabrice Bureau[3], Søren J. Sørensen[2] and Ruddy Wattiez[1]*

[1] Laboratory of Proteomics and Microbiology, Research Institute for Biosciences, University of Mons, Mons, Belgium, [2] Section of Microbiology, Department of Biology, University of Copenhagen, Copenhagen, Denmark, [3] Cellular and Molecular Immunology Service, GIGA Research, University of Liège (ULG), Liège, Belgium, [4] Section for Microbial Ecology and Biotechnology, Department of Plant and Environmental Sciences, University of Copenhagen, Frederiksberg, Denmark

OPEN ACCESS

Edited by:
Alejandro Couce, Polytechnic University of Madrid, Spain

Reviewed by:
Tal Dagan, University of Kiel, Germany
Francisco Dionisio, University of Lisbon, Portugal
Anne Farewell, University of Gothenburg, Sweden

***Correspondence:**
Valentine Cyriaque
valentinecyriaque@gmail.com

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *19 January 2021*
Accepted: *06 April 2021*
Published: *28 May 2021*

Citation:
Cyriaque V, Madsen JS, Fievez L, Leroy B, Hansen LH, Bureau F, Sørensen SJ and Wattiez R (2021) Lead Drives Complex Dynamics of a Conjugative Plasmid in a Bacterial Community. Front. Microbiol. 12:655903. doi: 10.3389/fmicb.2021.655903

Plasmids carrying metal resistance genes (MRGs) have been suggested to be key ecological players in the adaptation of metal-impacted microbial communities, making them promising drivers of bio-remediation processes. However, the impact of metals on plasmid-mediated spread of MRGs through selection, plasmid loss, and transfer is far from being fully understood. In the present study, we used two-member bacterial communities to test the impact of lead on the dispersal of the IncP plasmid pKJK5 from a *Pseudomonas putida* KT2440 plasmid donor and two distinct recipients, *Variovorax paradoxus* B4 or *Delftia acidovorans* SPH-1 after 4 and 10 days of mating. Two versions of the plasmid were used, carrying or not carrying the lead resistance *pbr*TRABCD operon, to assess the importance of fitness benefit and conjugative potential for the dispersal of the plasmid. The spread dynamics of metal resistance conveyed by the conjugative plasmid were dependent on the recipient and the lead concentration: For *V. paradoxus*, the *pbr* operon did not facilitate neither lead resistance nor variation in plasmid spread. The growth gain brought by the *pbr* operon to *D. acidovorans* SPH-1 and *P. putida* KT2440 at 1 mM Pb enhanced the spread of the plasmid. At 1.5 mM Pb after 4 days, the proteomics results revealed an oxidative stress response and an increased abundance of pKJK5-encoded conjugation and partitioning proteins, which most likely increased the transfer of the control plasmid to *D. acidovorans* SPH-1 and ensured plasmid maintenance. As a consequence, we observed an increased spread of pKJK5-*gfp*. Conversely, the *pbr* operon reduced the oxidative stress response and impeded the rise of conjugation- and partitioning-associated proteins, which slowed down the spread of the *pbr* carrying plasmid. Ultimately, when a fitness gain was recorded in the recipient strain, the spread of MRG-carrying plasmids was facilitated through positive selection at an intermediate metal concentration, while a high lead concentration induced oxidative stress with positive impacts on proteins encoding plasmid conjugation and partitioning.

Keywords: plasmid spread, metal, lead, conjugation, plasmid-mediated resistance, proteomics

INTRODUCTION

Metals constitute a serious risk for ecosystems because of their biotoxicity and bioaccumulation. Previous studies have revealed that metals impact microbial communities by modifying their composition (Gillan et al., 2005; Guo et al., 2019), decreasing their diversity (Sun et al., 2013; Kwon et al., 2015; Guo et al., 2019), and/or reducing the activity rates of their members (Jurburg et al., 2017). However, after a long-term metal contamination, microbial communities proved to be resilient. In several cases, long-term metal exposure had no visible impacts on alpha diversity or activity (Gillan et al., 2015; Ni et al., 2016; Jacquiod et al., 2018; Cyriaque et al., 2020a). Horizontal gene transfer, especially when mediated by plasmids, was proposed as a mechanism involved in the resilience of microbial communities in metal-impacted ecosystems (Jacquiod et al., 2018; Cyriaque et al., 2020b). Mobile genetic elements thus seem to bridge clinical and environmental ecosystems, where metal resistance genes (MRGs) are associated (*via* cross- and co-resistance systems) with antibiotic resistance genes (ARGs), a major global threat to public health (Perry and Wright, 2013; Pal et al., 2015; Nishida and Oshima, 2019). Plasmids are considered key players in bacterial adaptation, as they mobilize genes contributing to genome innovation (Norman et al., 2009). Their persistence in a community is dependent on (i) their acquisition rate (conjugation and transformation), (ii) their fitness (i.e., ability to survive in a competitive environment) cost/benefit on their host, and (iii) loss rate (stability) (Bahl et al., 2009; Lopatkin et al., 2017). Furthermore, the fitness effect of the plasmid may also depend on specific genetic interactions between the plasmid and the rest of the host genome as demonstrated across *Pseudomonas* species in a mercury-selective environment (Kottara et al., 2018). Therefore, drawing a general scheme on the role of a metal as a selection factor for the maintenance and spread of plasmids encoding MRGs in a microbial community is not trivial. In *Cupriavidus metallidurans* CH34, metals increased the abundance of conjugative transfer proteins (Monchy et al., 2007), and cadmium was shown to increase plasmid dispersal in subsurface-derived sediment microcosms (Smets et al., 2003; Pu et al., 2021). Copper can either decrease conjugation frequency (Parra et al., 2019) or promote plasmid-mediated gene transfer (Zhang et al., 2019). Furthermore, metals were shown to decrease plasmid dispersal in a soil microbial community without impacting the diversity of transconjugants (Klümper et al., 2017). Metals either negatively or positively modulate plasmid uptake depending on the metal, metal concentration, and recipient, with a large variability between metals and metal concentrations (Klümper et al., 2017; Wang et al., 2020). Bio-remediation processes through plasmid-mediated MRG bio-augmentation rely on the spread of the plasmid in a metal-contaminated environment (Garbisu et al., 2017). Additional knowledge is needed to understand the impact of metals on selection, plasmid maintenance, and conjugation processes. The present work focuses on the effects of lead on the spread of a broad-host-range plasmid encoding a lead resistance system into a recipient population. Our initial hypothesis was that the plasmid-encoded lead resistance system would facilitate the spread of the plasmid in a metal-impacted environment.

Lead is a non-essential metal and a major metal pollutant in the environment with strong adverse effects. We investigated the impact of lead [Pb(II)] on the spread of a conjugative IncP-1ε plasmid in a recipient population carrying, or not, the lead resistance operon *pbrTRABCD* from *C. metallidurans* CH34. The spreads of pKJK5-*gfp* and pKJK5-*gfp-pbr* were investigated between a plasmid donor *Pseudomonas putida* KT2440 and a recipient strain. Several species identified in both non-contaminated and metal-contaminated (including lead) sediment (Gillan et al., 2015) were tested as recipient strains. *Variovorax paradoxus* B4 and *Delftia acidovorans* SPH-1 were chosen because a significant amount of transconjugants was obtained, because the pKJK5 plasmid could be easily introduced by electroporation for burden assays, and because these strains have a growth rate similar to the donor *P. putida* KT2440, limiting the impact of native growth rate on plasmid spread. The spread of the plasmid in the recipient pool was quantified, and the molecular adaptative response of plasmid donors and recipients was studied at increasing lead concentrations (Pb-0, Pb-0.5, Pb-1, and Pb-1.5 mM – **Figure 1**). To take the effect of the *pbr* operon on the fitness of the host into account, we monitored plasmid-free and plasmid-carrying cells after 4 and 10 days of co-culture. The long-term conjugation assay did not report a direct transfer from the donor cells to the recipients but instead the stability of the plasmid in the recipient community which is the sum of plasmid loss and gain as well as host fitness loss or gain. The involvement of these factors was deciphered by measuring the fitness effect of the plasmids on each strain and profiling the proteome of the mating co-cultures.

MATERIALS AND METHODS

Strains, Plasmids, and Growth Conditions

For the construction of plasmid pKJK5-*gfpmut3-pbr*TRABCD-*kan*R-*tet*R, *C. metallidurans* CH34 and Electrocomp™ GeneHogs® *Escherichia coli* (Invitrogen) were cultivated in Luria–Bertani (LB) medium (see "**Supplementary Method**" for details on how the *pbr*TRABCD operon was inserted into pKJK5). The strain *P. putida* KT2440/Plpp*mCherry-kan*R harboring the plasmid pKJK5-*gfpmut3-kan*R-*tet*R (pKJK5-*gfp*) or pKJK5-*gfpmut3-pbr*TRABCD-*kan*R-*tet*R (pKJK5-*gfp-pbr*) was used as the plasmid donor. Precultures of those strains were grown in threefold diluted LB broth buffered with MOPS (2.1 g/L) (LB3D; peptone 3.3 g/L, NaCl 3.3 g/L, YE 1.7 g/L, MOPS 2.1 g/L) and supplemented with 50 μg/ml tetracycline (30°C, overnight). *V. paradoxus* B4 (DSMZ, Germany) and *D. acidovorans* SPH-1 (DSMZ, Germany) were used as plasmid recipients. An overview of these strains' characteristics is displayed in **Supplementary Table 1**. Precultures of these strains were grown overnight in 457 medium (DSMZ) supplemented with 2 g/L mercaptosuccinate and LB3D (30°C). To assess the impact of the plasmid on

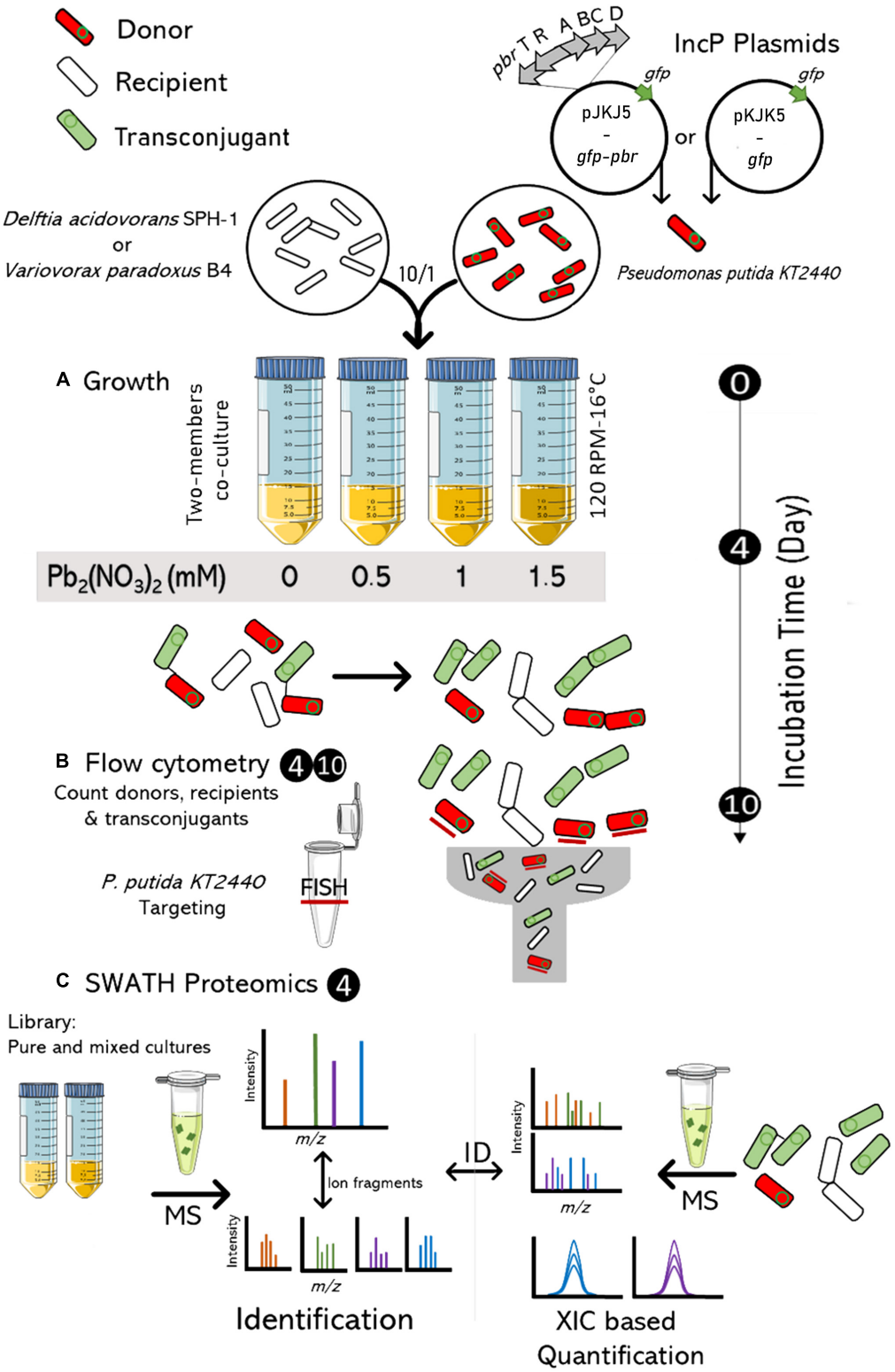


FIGURE 1 | Experimental design combining flow cytometry and proteomics to assess lead impact on the functional profile of a mating coculture and the spread of the plasmid. Recipient and donor cells were grown (10/1 ratio) in LB3D (120 RMP, 16°C) over 10 days with an increasing concentration of $Pb_2(NO_3)_2$ (0, 0.5, 1, 1.5 mM) **(A)**. After 4 and 10 days, flow cytometry was used to quantify recipient, donor (FISH labelled) and transconjugant cells **(B)**. After 4 days, they were also analyzed by SWATH-MS proteomics **(C)**. MS: Mass spectrometry; *m/z*: mass-to-charge ratio; XIC: extracted ion chromatogram.

the growth of its host, both plasmid donors and the clones *V. paradoxus* B4/pKJK5-*gfp*, *V. paradoxus* B4/pKJK5-*gfp-pbr*, *D. acidovorans* SPH-1/pKJK5-*gfp*, and *D. acidovorans* SPH-1/pKJK5-*gfp-pbr* were obtained by conjugation between those recipient cells and GeneHogs® *E. coli*-bearing plasmids ("**Supplementary Method**").

Assessing Plasmid Fitness Effect

The fitness effect of plasmids pKJK5-*gfp* and pKJK5-*gfp-pbr* when carried by *P. putida* KT2440/PlppmCherry-*kan*R, *V. paradoxus* B4, or *D. acidovorans* SPH-1 was measured relatively to the corresponding plasmid-free strain as the ratio between the growth rate of plasmid-free and plasmid-carrying cells ($\mu 0/\mu 1$) (Lopatkin et al., 2017) considering that pKJK5 displays a low rate of plasmid loss (Bahl et al., 2004). Plasmid-carrying clones were obtained by conjugation as described in the "**Supplementary Method**" and selected on tetracycline. Second precultures in 100 ml LB3D (30°C) (supplemented with tetracycline 50 μg/ml for plasmid-carrying strains) were grown until an optical density (OD, 600 nm) between 0.4 and 0.6, measured with a Helios Zeta UV–vis spectrophotometer (Thermo Fisher Scientific). Then, the precultures were washed twice in LB3D (2 min, 7,000 × *g*), and cells were counted using a Bright-Line™ Hemacytometer (Merck) following the manufacturer's instructions. A stated number of cells (**Supplementary Table 2**) were sampled from precultures and diluted in a final volume of 15 ml LB3D supplemented with 0, 0.5, 1, or 1.5 mM $Pb(NO_3)_2$ in 50 ml falcon (n = 3) and incubated (16°C, 120 RPM) until the stationary phase. The OD (595 nm) was measured in 96-well plates multiplied with FLUO star OPTIMA (BMG LABTECH). Growth rate (μ) was quantified by log-transforming the growth curves and fitting the linear portion to obtain the slope of the regression (**Supplementary Table 3**). The significance of the fitness effect for each strain at different lead concentrations was assessed by comparing their growth rate with the plasmid-free corresponding strain.

Plasmid Spread Assay in Liquid Community and Cytometry Analysis

The spreads of plasmids pKJK5-*gfp* and pKJK5-*gfp-pbr* carried by *P. putida* KT2440/Plpp*mCherry-kan*R as plasmid donor were assessed in co-culture with *V. paradoxus* B4 or *D. acidovorans* SPH-1 (**Figure 1**) with the equation T / (DR), where D is the count of donors (plasmid-carrying *P. putida* KT2440/Plpp*mCherry-kan*R), R is the count of recipients (plasmid-free *V. paradoxus* B4 or *D. acidovorans* SPH-1 cells), and T is the count of transconjugants (recipient cells that acquired the plasmid) (Jacquiod et al., 2017). In this long-term mating experiment, the recipient × donor normalization did not portray transfer frequency but instead a spread index of long-term plasmid transfer, plasmid loss, and a selection dynamics. Donor normalization was done as donor cells act as a plasmid source–sink (Hall et al., 2016) whose numbers, determined by *P. putida* KT440 growth, impact the plasmid invasion rate.

To start co-cultures, the strains were first grown separately in 100 ml LB3D (30°C; supplemented with 50 μg/ml tetracycline for donor strains) until an OD (600 nm) between 0.4 and 0.6 was measured using a Helios Zeta UV–vis spectrophotometer (Thermo Fisher Scientific). Then, bacteria were washed twice in LB3D (2 min, 7,000 × *g*), and cells were counted using a Bright-Line™ Hemacytometer (Merck) following the manufacturer's instructions. A stated number of cells (**Supplementary Table 2**) were sampled from the preculture and diluted in a final volume of 15 ml LB3D supplemented with 0, 0.5, 1, or 1.5 mM $Pb(NO_3)_2$ in 50-ml falcon (n = 3). Co-cultures were grown at 16°C and 120 RPM for 10 days and sampled (i) after 4 days for cell count and proteomics analyses (section "MRM Identification and Quantification of the PbrA Protein" and "Proteomic Analysis by SWATH Mass Spectrometry") and (ii) after 10 days for cell count (**Figure 1**). Donor, recipient, and transconjugant counts were assessed by flow cytometry for the detection of green fluorescent protein (GFP) fluorescence expressed by the *gfpmut3* gene carried by the pKJK5 plasmid carried by donor and transconjugant cells as described previously (Klümper et al., 2015; Cyriaque et al., 2020b). Plasmid donor cells were differentiated from transconjugants by flow cytometry (FC)-associated fluorescence *in situ* hybridization (FISH) targeting the 16S rRNA of *P. putida* as previously used (Gougoulias and Shaw, 2012). To that end, 1 ml of co-culture was centrifuged (6 min, 6,000 × *g*) and fixed by resuspending it in 1 ml PFA (4%), pH 7, for 15 min. The fixed cells were washed twice in PBS (6 min, 6,000 × *g*), resuspended in 247 μl of prewarmed (48°C) hybridization buffer supplemented with 3 μl of the probe PSE1284 (Yamaguchi et al., 2006) associated with an Alexa Fluor 647 fluorochrome (Eurogentec, Liège, Belgium) (250 ng/μl), and incubated for 4 h. The samples were centrifuged (5 min, 16,000 × *g*) and resuspended in 500 μl of hybridization buffer for 20 min. They were centrifuged (5 min, 16,000 × *g*) again and resuspended in 500 μl of wash buffer. Finally, the samples were centrifuged, resuspended in 1 ml PBS, and stored at 4°C until cytometry analyses. The hybridization [urea (1 M), NaCl (0.9 M), Tris HCl (pH 7.4, 20 μM)] and wash [urea (4 M), NaCl (0.9 M), Tris HCl (pH 7.4, 20 μM)] buffers contained urea as a denaturation agent to avoid GFP denaturation as described previously (Lawson et al., 2012; Kommerein et al., 2017). Cytometry analyses were carried out using a BD Science FACS Fortessa in a 96-well microplate with the following parameters: forward scatter, 390 V; side scatter, 176 V; detectors for green fluorescence associated with GFPmut3 (bandpass filter 530/30 nm, 501 V) and for red fluorescence associated with Alexa Fluor 647 (bandpass filter 670/14 nm, 550 V). FlowJo V10 was used to analyze the results to count donors, empty donors, recipients, and transconjugants (**Supplementary Table 4**). To confirm the accuracy of FC-FISH and GFP integrity, pure cultures of each donor and recipient were tested in upstream mating experiments (**Supplementary Figure 1**).

Proteomic Analysis by SWATH Mass Spectrometry

A quantitative proteomic approach was used to assess the metal's impact on plasmid transfer machinery and bacterial proteomes after 4 days of mating, concurrently with plasmid

dispersion measurement. For that, 600 µl of 0 and 0.5 mM lead-cultured samples and 1,200 µl of 1 and 1.5 mM lead-cultured samples were harvested, centrifuged, and washed twice (6,000 × *g*, 6 min, 4°C) with PBS. The proteins were extracted, reduced, alkylated, precipitated, and trypsinized from the pellet using the PreOmics Kit (PreOmics GmbH, Germany), following the manufacturer's instructions. The obtained peptides were quantified using the Pierce™ Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific). For post-acquisition retention time calibration, a PepCalMix solution (Protein Extract Digest) (AB SCIEX, Framingham, MA, United States) was added to 4 µg of peptides (50 fmol on column) following the manufacturer's instructions. The peptides (2 µg on column) were analyzed on a UHPLC-HRMS/MS instrument (AB SCIEX LC420 and TripleTOF™ 6600) using SWATH data-independent acquisition as described in the "**Supplementary Method**." SWATH wiff files were processed using AB SCIEX PeakView 2.2 software and SWATH™ Acquisition MicroApp. Up to six peptides with at least 95% confidence were selected, with six transitions per peptide. The XIC extraction window was set to 15 min, and the XIC width was set to 70 ppm. The peptide area corresponds to the sum of the fragment ion area, and the protein area corresponds to the sum of the peptide area. The protein areas were extracted and exported in AB SCIEX MarkerView™ 1.2 software for normalization and statistical analysis. The protein extraction characteristics, the number of proteins identified in both libraries at 1% FDR, and the proteome coverage of the SWATH proteomic analyses are displayed in **Supplementary Tables 5**, **6**. To be able to compare the different samples, protein relative abundances were obtained by dividing each protein abundance by the cumulated protein area of the corresponding sample. The obtained dataset underwent a second normalization: (i) chromosomally encoded proteins were normalized by the total protein abundance associated with the specific strain. We used the protein biomass (and not cell count) to normalize protein counts because the expression level of proteins per cell depends on the strain (Cortes et al., 2019). Therefore, using cell count as a normative factor unbalanced the protein counts (**Supplementary Figures 2–5**); and (ii) for plasmid-encoded proteins, it was impossible to discriminate what bacterium expressed the proteins. In this case, protein counts were instead normalized by the proportion of plasmid-carrying cells in the two-member community obtained by flow cytometry. The results were deposited on the Peptide Atlas public repository[1] under accession number PASS01468[2]. Significant differences between norm2-protein abundances (log-2-transformed) were determined using a two-tailed Student's *t*-test across the different lead concentrations for each filter mating association and for the comparison of the same lead condition impacting the mating pair with or without *pbr*TRABCD in the exchanged plasmid. Proteins displaying a minimum *p*-value below 0.05 were taken into consideration, and strain-specific proteins with a *p*-value below 0.01 were plotted in heat maps using hierarchical cluster dendrograms (Euclidean distance and average clustering) in RGui software [vegan (Oksanen et al., 2019), rioja (Juggins, 2019), and gplots (Warnes et al., 2016) R-packages]. The strain-specific proteins are displayed in the "Heat Map **Supplementary Material**."

[1]http://www.peptideatlas.org/

[2]http://www.peptideatlas.org/PASS/PASS01468

MRM Identification and Quantification of the PbrA Protein

Quantification of PbrA was performed using multiple reaction monitoring (MRM)-based relative quantification. The spectral signature of PbrA was obtained through the analysis of *C. metallidurans* CH34 samples using a regular liquid chromatography–tandem mass spectrometry (LC–MS/MS) procedure (Leroy et al., 2015). This sample was selected because *C. metallidurans* is known to express, in the presence of a high lead load, a high level of PbrA. Bacterial culture was obtained as described previously (30), proteins were extracted and digested using the PreOmics Kit (PreOmics GmbH, Planegg/Martinsried, Germany), and peptides were quantified using a Pierce™ Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific) following the manufacturer's instructions. A total of 13 peptides were detected in the LC–MS/MS analysis of *C. metallidurans* and evaluated for quantification in MRM mode on a Q-Trap 6500+ coupled to LC420 chromatography (SCIEX) operated in microflow mode. After transition optimization and interference removal, four peptides with four to five transitions each were selected for quantification (**Supplementary Table 5**). Quantification was assessed on (i) the pKJK5-*gfp*-carrying sample with 1 mM lead as a control and (ii) all pKJK5-*gfp-pbr*-carrying samples. For that, 3 µg of trypsin-digested proteins was separated on a C18 YMC-Triart 0.3 × 150-mm column operated at a flow rate of 5 µl/min with an ACN gradient from 2 to 35% in 17 min. Skyline was used for visual inspection of the data and area under the curve integration. Peak picking for each peptide was manually refined using the transition intensity ratio and retention time as leading parameters. These parameters were obtained from a *C. metallidurans* sample that contained a higher level of PbrA. The intensity of all four to five transitions was summed up for each peptide. Protein abundance was obtained as the average of the Ln-transformed area under the curve of each of the four peptides.

RESULTS

Fitness Effect of Plasmids on Their Host

The fitness effect of the pKJK5-*gfp* and pKJK5-*gfp-pbr* plasmids impeding the growth of plasmid donors and recipients was calculated as the ratio of the growth rate of plasmid-free and plasmid-carrying cells. Without lead exposure, none of the plasmids had any significant fitness effect on *P. putida* KT2440, *D. acidovorans*, or *V. paradoxus* B4 (**Figure 2**).

In the presence of lead, neither pKJK5-*gfp* nor pKJK5-*gfp-pbr* had a cost for *V. paradoxus* B4. We, therefore, expected no lead-mediated positive selection of the plasmids into *V. paradoxus* B4.

However, the growth rate of *D. acidovorans* SPH-1/pKJK5-*gfp* was significantly decreased at high lead concentrations (Pb-1 and Pb-1.5). This fitness effect was compensated by the

pbrTRABCD operon, dedicated to lead resistance. Despite being non-significant, the growth rate of *D. acidovorans* SPH-1 carrying pKJK5-*gfp-pbr* showed a decreasing trend in Pb-0.5. At low lead concentrations, the cost imposed by the *pbr* operon would then be larger than the benefit it procured to the host. Lead had a negative impact on the growth rate of plasmid carrier *P. putida* KT2440/pKJK5-*gfp*. In Pb-0.5 and Pb-1, the *pbr* operon compensated for the negative fitness effect of the plasmid. In Pb-1.5, decreased growth was significant in both plasmid carriers, with an attenuated effect on pKJK5-*gfp-pbr*-carrying cells (**Figure 2**).

Two-Member Community Dynamics Assessed by Flow Cytometry

The proportion of donors (*P. putida* KT2440), recipients, and transconjugants in a two-member mating assay was measured using flow cytometry (**Figures 1**, **3** and **Supplementary Table 4**). We used either *D. acidovorans* SPH-1 or *V. paradoxus* B4 as recipient strains. Plasmid spread was detected in both recipients, with a higher frequency in *V. paradoxus* B4 (**Figures 3**, **4**). For this strain, lead only had a significant negative impact on the spread of pKJK5-*gfp* in the recipient community in Pb-1 after both 4 and 10 days.

When *D. acidovorans* SPH-1 was the recipient strain, after 4 days, Pb had a negative impact on pKJK5-*gfp* dispersion until Pb-1. Interestingly, the spread of pKJK5-*gfp* in Pb-1.5 was not different from that in the control (Pb-0) after 4 days of mating, but the transconjugants relapsed after 10 days. After 4 days, in Pb-1, the spread of pKJK5-*gfp-pbr* was not impacted as opposed to pKJK5-*gfp*.

When *V. paradoxus* B4 was used as the recipient, plasmid loss in the donor cell fraction (plasmid-free *P.p.*; **Figure 3** and **Supplementary Table 4**) progressively increased with lead concentration and time. When *D. acidovorans* SPH-1 was used as the recipient, plasmid loss in the donor *P. putida* KT2440 was high in Pb-1.5 where the fitness effect of both plasmids was the highest (the percentage of cured cells was calculated as the ratio between "empty *P.p.*" and total "*P.p.*": 84.3 ± 4.2% of cured cells from pKJK5-*gfp* and 74.1 ± 1.4% from pKJK5-*gfp-pbr*) after 4 days of mating, but plasmid-free *P. putida* KT2440 largely decreased after 10 days (18.5 ± 12.6% of cured cells from pKJK5-*gfp* and 11.4 ± 3.1% from pKJK5-*gfp-pbr*) (**Figure 3** and **Supplementary Table 4**).

Meta-Proteomic Profiling of Two-Member Communities

To obtain a better understanding of the functional response to lead in the co-cultures, a quantitative meta-proteomic analysis was performed using a SWATH approach. A spectral library built with a data-dependent acquisition (DDA) workflow was generated with monocultures and co-cultures of donors and recipients (**Supplementary Table 2**) at 0 and 1 mM $Pb(NO_3)_2$. The label-free DDA analyses of pure cultures revealed that lead was associated with metal resistance proteins (**Supplementary Table 6**). In *P. putida* KT2440, we observed the induction of phosphate metabolism-associated proteins (e.g., phosphatase) and pyoverdine-associated proteins. The ATPase PbrA was exclusively observed in *P. putida* KT2440/pKJK5-*gfp-pbr* in the presence of lead. In *V. paradoxus* B4, a high level of TonB siderophore-related proteins and a putative ABC transporter iron-binding protein and the iron-sulfur assembly scaffold protein IscU were upregulated. In *D. acidovorans* SPH-1, a large diversity of metal resistance-associated proteins was observed, including (i) phosphate metabolism-associated proteins, (ii) TonB siderophore-related proteins, (iii) efflux pumps, (iv) glutathione S-transferase, (v) the iron–sulfur assembly protein IscA, (vi) thioesterase, (vii) iron permease, and (viii) bacterioferritin (**Supplementary Table 6**). Using SWATH-MS results, the relative biomass of the donor and recipient bacteria in co-cultures was evaluated as the sum of the area of all proteins attributed to the specific strain (**Supplementary Table 4**). The coverage of the proteome of each bacterial member is displayed in **Supplementary Table 7**.

Differential Impacts of Lead on Metabolic Response and Resistance Mechanisms

When the plasmid donor *P. putida* KT2440 was grown in the presence of *V. paradoxus* B4, a total of 632 proteins assigned to *P. putida* KT2440 were significantly impacted by lead (p-value < 0.05; 233, p-value < 0.01). A total of 734 proteins assigned to *V. paradoxus* B4 were significantly impacted by lead (p-value < 0.05; 296, p-value < 0.01). When the *P. putida* KT2440 plasmid donor was grown in the presence of *D. acidovorans* SPH-1, a total of 739 proteins were significantly impacted by lead (p-value < 0.05; 352, p-value < 0.01). The abundance of proteins identified as belonging to *D. acidovorans* SPH-1 that were significantly impacted by lead (585, p-value < 0.05; 254, p-value < 0.01) was increased the most at a lead concentration of 1.5 mM.

The detailed metabolic modulation in the SWATH proteomes of the three strains in both mating pair conditions with increasing lead concentrations is displayed in the "Heat Map **Supplementary Material**." In Pb-1.5, in co-cultures with *D. acidovorans* SPH-1, *P. putida* KT2440's DNA repair proteins [transcription repair coupling factor (mfd), MutS, SbcC, L COG; "Heat Map **Supplementary Material**"] and stress proteins (CysQ; P COG, "Heat Map **Supplementary Material**") were overabundant. Among lead-impacted proteins, enzymes requiring divalent cation cofactors [Fe(II), Mg(II), Zn(II), Co(II), and Mn(II)] or the [4Fe–4S] cluster were highly represented in Pb-1.5. Although lead cannot directly inactivate the [4Fe–4S]-dependent class I enzyme (Xu and Imlay, 2012), the oxidative stress that Pb most likely induced (Wang et al., 2012; Jarosławiecka and Piotrowska-Seget, 2014) decreased the abundance of these enzymes as previously shown with Al–Ga toxicity (Chenier et al., 2008). Among these enzymes, [4Fe–4S]-dependent fumarate hydratase class I was replaced by iron-independent fumarate hydratase class II (**Figure 5C**). When grown with *D. acidovorans* SPH-1, the class I enzyme of *P. putida* KT2440 was much less decreased in abundance at Pb1.5 when grown with the *pbr* operon. Similarly, the decreased abundance of the succinate dehydrogenase iron–sulfur subunit of *P. putida* KT2440 was amplified in the presence of *D. acidovorans* SPH-1

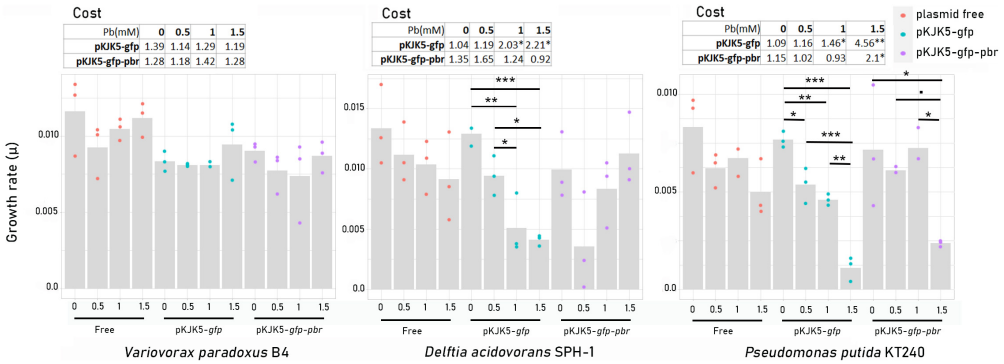


FIGURE 2 | Growth rates (μ) and plasmid fitness effect (cost, μ_0 / μ_1) ± SEM, where μ_0 = growth rate of the plasmid-free cells and μ_1 = growth rate of plasmid-carrying cells in pure cultures. The stars show significant differences among growth rates and the significance of the fitness effect (comparison between growth rates of the plasmid-carrying cells and plasmid-free cells at the same lead concentration). **p*-value < 0.05; ***p*-value < 0.01; ****p*-value < 0.001 (*N* = 3) (*n* = 3; Tukey test; $p \leq 0.05$).

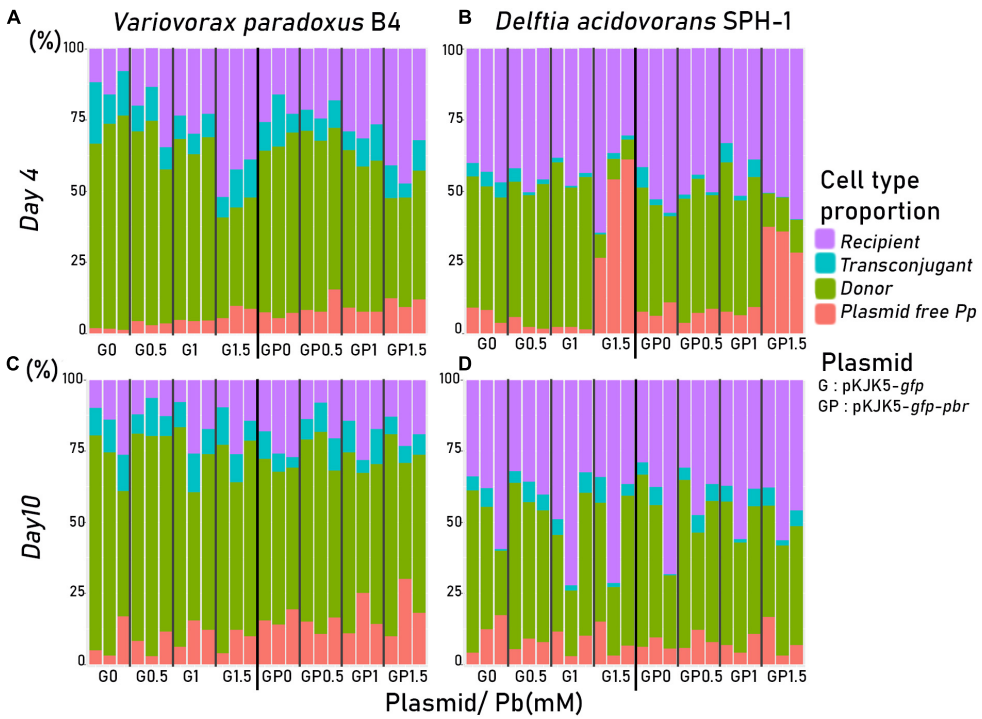


FIGURE 3 | Cell proportions measured among 30,000 cells after 4 days **(A,B)** or 10 days **(C,D)** of co-culture using a BD Science FACS Fortessa flow cytometer. Plasmid-carrying cells were detected by green fluorescence associated with GFPmut3. *Pseudomonas putida* KT2440 was identified by flow-FISH using an Alexa Fluor 647 fluorescent specific probe. The associated spread index is displayed in **Figure 4**.

(**Figures 5A,B**) and reduced with the benefit of the *pbrTRABCD* operon. In *V. paradoxus* B4, an increased amount of pilus-related proteins ("Heat Map **Supplementary Material**") may explain the large mobility of pKJK5 plasmids into this recipient strain compared with *D. acidovorans* SPH-1 (**Figures 3**, **4**).

At high lead concentrations, all strains upregulated metal resistance-associated proteins involving phosphatases (production of phosphate salt for the precipitation of metal cations; Jarosławiecka and Piotrowska-Seget, 2014), metabolism of glutathione (binding metal cations; Taghavi et al., 2009), and efflux transporters (RND efflux systems and heavy-metal transporters; P COG) ("Heat Map **Supplementary Material**"). In the *V. paradoxus* B4 proteome, additional siderophore-related proteins were also overabundant, such as TRAP dicarboxylate transporters, TonB receptors (P COG), a putative non-ribosomal peptide synthase (NRPS), and polyketide synthase (Q COG), for the biosynthesis of variochelin lipopeptide siderophores (Kurth et al., 2016). In *D. acidovorans* SPH-1,

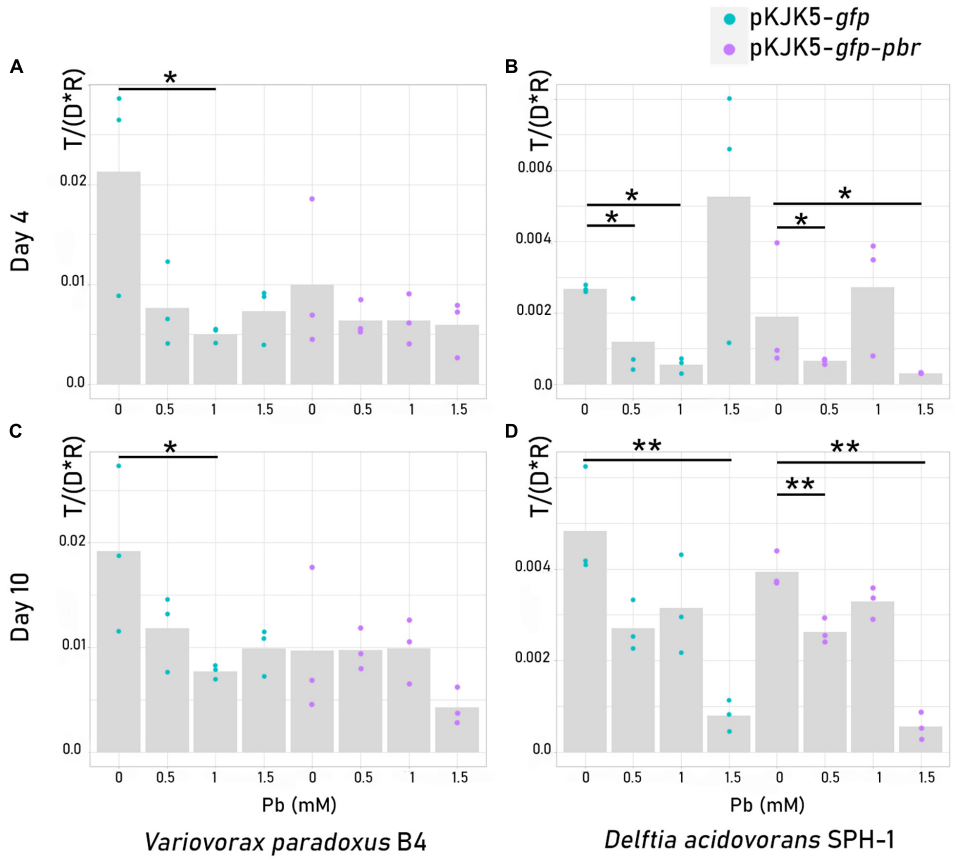


FIGURE 4 | Plasmid spread index [*T* / (*D* * *R*)] ± SEM in recipient cells, *Variovorax paradoxus* B4 **(A,C)** or *Delftia acidovorans* SPH-1 **(B,D)**, after 4 days **(A,B)** or 10 days **(C,D)** of mating. The *p*-values were calculated from Kruskal–Wallis tests (4-day mating) or log-2-transformed abundances using *t*-tests (10-day mating) depending on homoscedasticity. *p-value < 0.05; **p-value < 0.01; ***p-value < 0.001 (N = 3). The dotted lines are for pKJK5-*gfp* comparisons, and the continuous lines are for pKJK5-*gfp-pbr* comparisons.

a large number of metal resistance-induced proteins were overabundant, including TonB siderophores, phosphatases, and efflux pumps, explaining its high fitness (P COG) ("Heat Map **Supplementary Material**"). *P. putida* KT2440 also upregulated pyoverdine-associated proteins (V COG) (Taghavi et al., 2009; Naik and Dubey, 2011; Jarosławiecka and Piotrowska-Seget, 2014). When *P. putida* KT2440 was co-cultured with *D. acidovorans* SPH-1 in Pb-1.5, (i) penicillin-binding protein 1B was overly increased, especially when carrying pKJK5-*gfp* (**Figure 5B**), and (ii) toluene efflux periplasmic linker protein TtgA and outer membrane protein H1 were increased exclusively when carrying pKJK5-*gfp-pbr* (**Figure 5B**). The outer membrane protein H1 replaces divalent cations at binding sites on lipopolysaccharide and would then prevent the uptake of Pb(II) ions (Bell and Hancock, 1989). In Pb-1, phosphatase, phosphate-binding and transport-associated proteins as well as putative cation transporter (P COG) and PhoB and PhoR proteins (T COG) were more abundant when the *pbrTRABCD* operon was present (**Figure 6**). Displayed resistance systems were then dependent on the strain, its mating partner, and the presence of the *pbr* operon.

Lead Impacts on pKJK5 Backbone Genes

To decipher lead effects on the conjugation machinery, we also inspected conjugation-associated proteins encoded by the pKJK5 plasmids. We normalized the abundance of these proteins by the proportion of plasmid-carrying cells assessed by flow cytometry. When *V. paradoxus* B4 was the mating partner, 16 proteins encoded by the pKJK5 plasmid were significantly impacted. Among them, TraD, TrbA, TrbB, TrbF, TrbG, and TrbH were decreased at 1.5 mM (**Supplementary Table 8** and **Figure 7**). When *D. acidovorans* SPH-1 was the mating partner, in Pb-1, TraC and TraG of the pKJK5-*gfp-pbr* plasmid displayed a positive fold-change. In Pb-1.5, the proteins KorB, IncC1, TraC, TraE, TraG, TrbE, and TrbI were increased, especially in pKJK5-*gfp* (**Figure 7** and **Supplementary Table 8**).

Plasmid-Encoded PbrA Expression

Pbr proteins were not detected in the SWATH analysis, revealing their low abundance level. Therefore, PbrA relative abundance was measured for the samples carrying the pKJK5-*gfp-pbr* plasmid using a targeted MRM-based relative quantification approach (**Supplementary Figure 6**). This sensitive method revealed the presence of PbrA in the samples. No significant

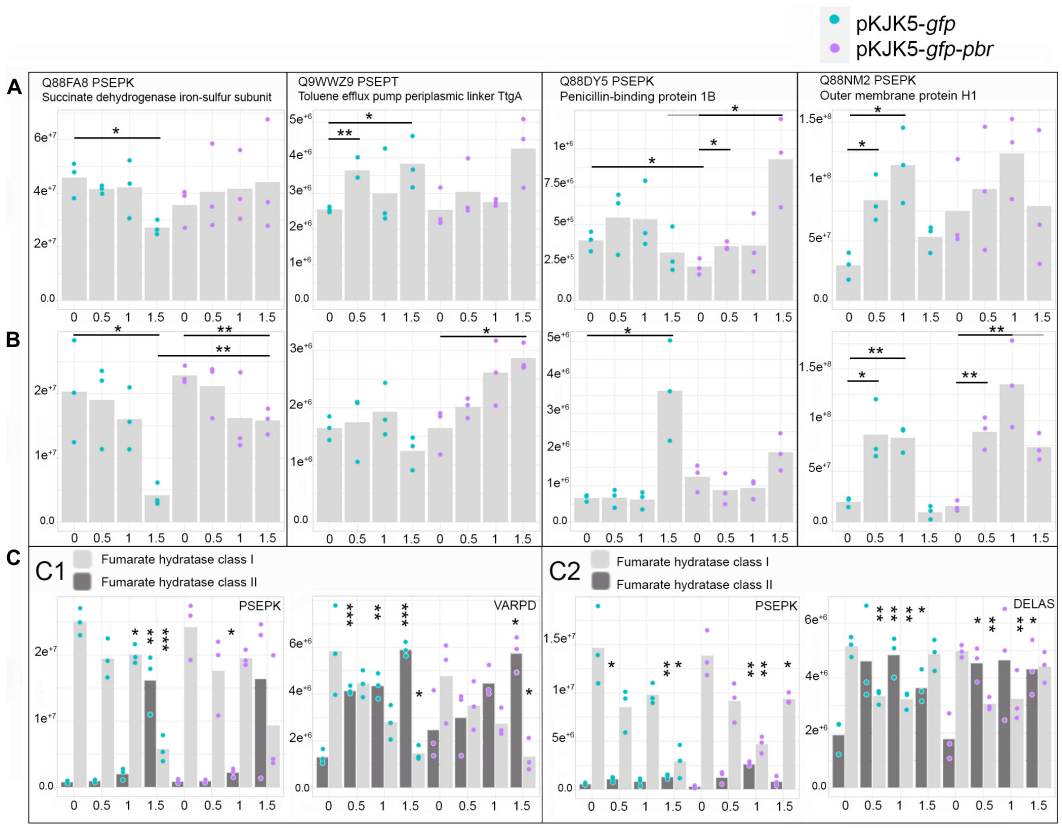


FIGURE 5 | Abundances (± SEM) of proteins of interest in the two-member community without the *pbr*TRABCD operon (blue) or including the *pbr*TRABCD operon (purple) depending on Pb(II) concentration (mM). Different panels represent the abundance of proteins identified as belonging to *Pseudomonas putida* KT2440 when cultivated with *Variovorax paradoxus* B4 **(A)** or with *Delftia acidovorans* SPH-1 **(B)**; the abundance of fumarate hydratase enzymes in *P. putida* KT2440 (PSEPK) and *V. paradoxus* B4 (VARPD) when cultivated together (C1) and in *P. putida* KT2440 (PSEPK) and *D. acidovorans* SPH-1 (DELAS) when cultivated together (C2). The *p*-values were calculated from log-2-transformed abundances using a *t*-test ($n = 3$). **p*-value < 0.05; ***p*-value < 0.01; ****p*-value < 0.001. In **(C)**, the *p*-values were calculated between the different lead concentrations (Pb0.5, Pb1, or Pb1.5) and the control (Pb0) of the same co-culture.

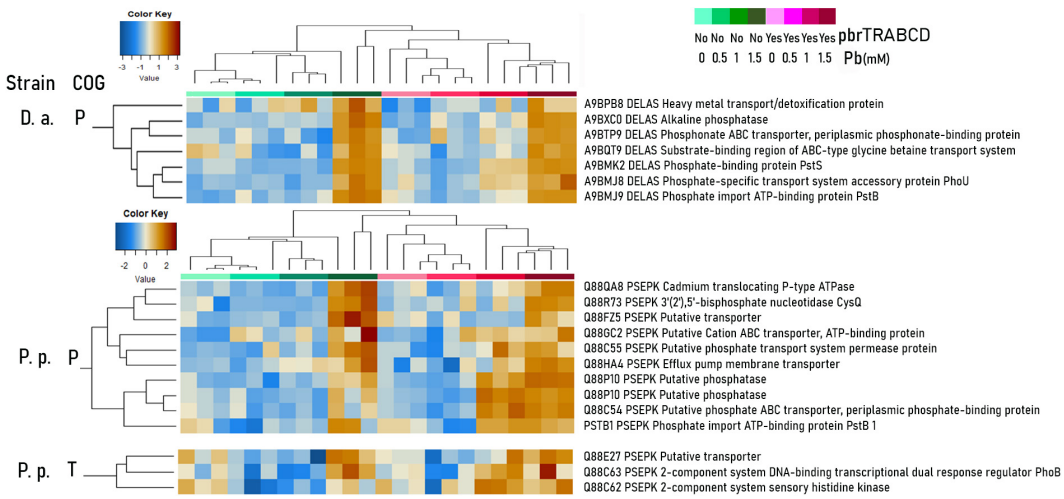


FIGURE 6 | Centered-scaled log-2-transformed abundances of *Pseudomonas putida* KT2440 (P.p.) and *Delftia acidovorans* SPH-1 (D.a.) proteins (two or more identified peptides) of interest belonging to P (inorganic ion transport and metabolism) and T (signal transduction mechanisms) COGs are displayed in heat maps (Euclidean distance, average clustering; entire heat maps: see "Heat Map **Supplementary Material**").

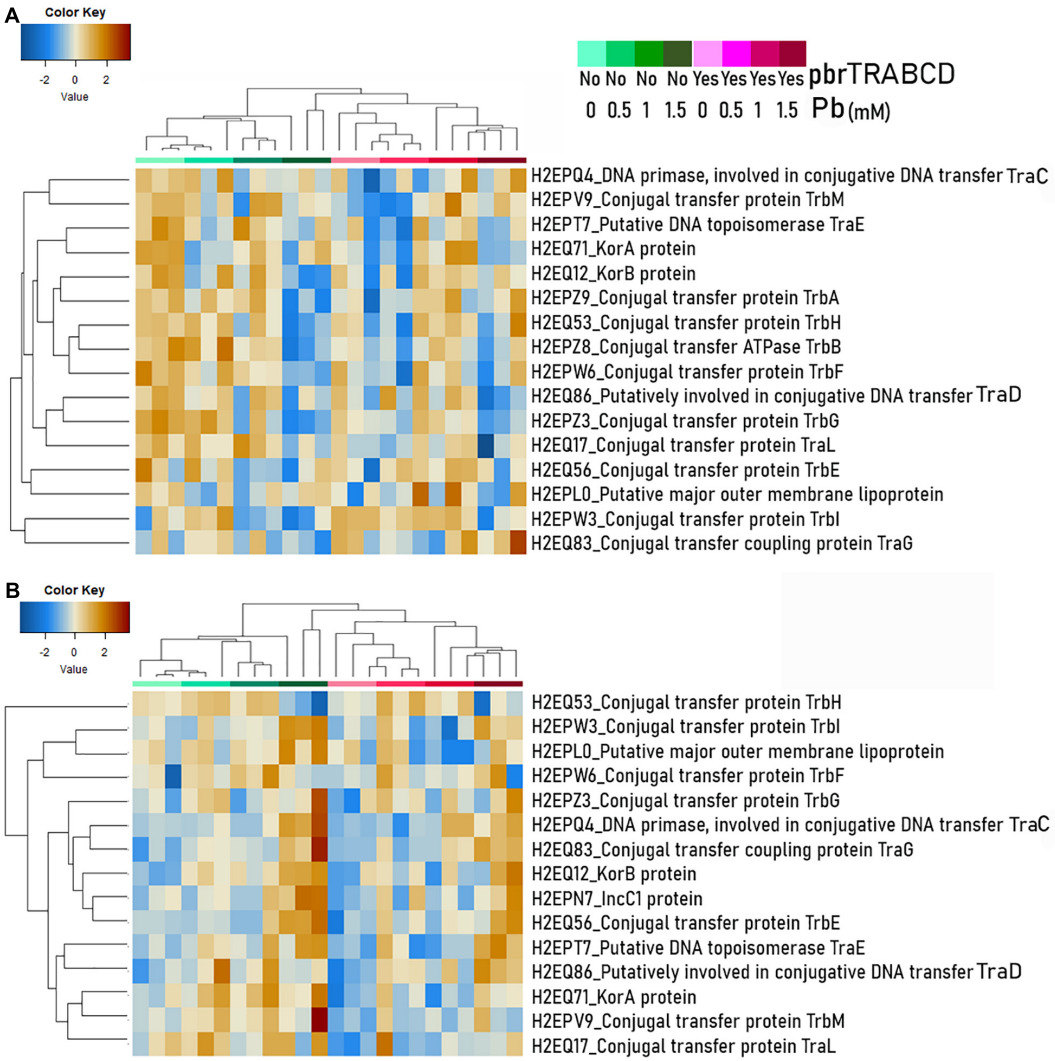


FIGURE 7 | Abundance of conjugation-associated proteins (two or more identified peptides) encoded by the pKJK5 plasmid [normalization on the summed area of all peptides for each sample and by cell proportion of plasmid carrier (donor + transconjugant) (see **Supplementary Table 8**)] displayed on a heat map (Euclidean distance and average clustering) as extracted from two-member communities using either *Variovorax paradoxus* B4 **(A)** or *Delftia acidovorans* SPH-1 **(B)** as the plasmid recipient.

differences in the *V. paradoxus* B4 mating pair and an increased relative abundance in the *D. acidovorans* SPH-1 mating pair (at Pb 1 and 1.5 mM) were observed.

DISCUSSION

Previous works revealed that the presence of metals either favor (Smets et al., 2003; Monchy et al., 2007) or inhibit the spread of plasmids (Klümper et al., 2017; Parra et al., 2019) with different impacts on the bacterial members of a soil community (Klümper et al., 2017). It was suggested that the cost and advantage trade-offs imposed by plasmids make their dispersion a dynamic process (Pinilla-Redondo et al., 2018; Cyriaque et al., 2020b). The pKJK5 plasmid has been shown to ensure its stability in a bacterial population using a high conjugation efficiency (Bahl et al., 2007) and an analogous partitioning system to the RK2 plasmid (Rosche et al., 2000). Lead may either modulate the maintenance machinery or topple the cost/benefit ratio.

Lead, Mating Partner, and MRGs Modulated the Fitness of the Hosts

Pseudomonas putida KT2440, *V. paradoxus* B4, and *D. acidovorans* SPH-1 separately displayed similar lead minimal inhibitory concentrations (MICs, 2 mM in LB 3D). Fitness effect was calculated as the ratio between the separate growth rate of plasmid-free and plasmid-carrying cells. Although the broad-host-range pKJK5 plasmid was shown to be stable in *Escherichia coli* (Bahl et al., 2004), a plasmid loss in a part of the population cannot be dismissed, especially since plasmid-free *P. putida* arose in the mating experiment, mostly after 4 days, at high lead

concentration. If so, the fitness effect of the plasmid might have been underestimated. Nevertheless, some observations can be addressed: First, lead had no significant effect on the growth of any of the plasmid-free strains. Moreover, while lead had a significant impact on plasmid-bearing *D. acidovorans* SPH-1 and *P. putida* KT2440, no significant effect was recorded on the growth rate of plasmid-bearing *V. paradoxus* B4. The presence of the *pbr* operon had no effect on the growth of *V. paradoxus* B4 when exposed to lead, suggesting that the *pbr* operon is not beneficial for *V. paradoxus* B4. This might be due to the high upregulation of siderophore-related proteins (e.g., NRPS, polyketide synthase and outer-membrane TonB receptor). As a consequence, the *pbr* operon would then be obsolete in the resistance against lead. PbrA is an efflux pump transporting zinc, cadmium, and lead ions from the cytoplasm to the periplasm, where phosphates released in the periplasm by the undecaprenyl pyrophosphate phosphatase PbrB form metallophosphates with lead cations (Hynninen et al., 2009). This system would then compete with lead-binding siderophores. Consequently, plasmid spread in the *V. paradoxus* B4 population depends on conjugation and stability systems regardless of the MRGs it carries. In *D. acidovorans* SPH1 and *P. putida* KT2440, the *pbr* defense system decreased the fitness effect of the plasmid, impacting the spread of the plasmid in *D. acidovorans* SPH-1. After 4 days of mating, at Pb1, the spread of pKJK5-*gfp-pbr* was higher than pKJK5-*gfp*. Similar selection dynamics where plasmids provide a fitness benefit have been recorded many times in the presence of antibiotics (Lopatkin et al., 2016). Such selection processes and the colocation of ARGs and MRGs on plasmids may amplify the risk of spread of ARGs in metal-contaminated environments. After 10 days, the plasmid continued to spread and the pKJK5-*gfp* plasmid caught up with the *pbr* variant.

It should be noted that, in these co-cultures, the *P. putida* KT2440 proteome reveals that the presence of the *pbr* operon was associated with an overabundance of alternative systems. For instance, we see an overabundance of the outer membrane protein H1, the toluene efflux pump TtgA, phosphatases, and proteins associated with phosphate import, especially in Pb-1, where the dispersion of pKJK5-*gfp-pbr* was maintained, in contrast to pKJK5-*gfp*. Phosphatases and phosphate import-related proteins may be required to ensure the turnover of undecaprenyl pyrophosphate in the membrane associated with the *pbr* resistance system. Beneficial *pbr* genes then most likely sustained plasmid dispersion in Pb-1 after 4 days of mating by increasing the fitness of their host.

Metal-Mediated Oxidative Stress at High Lead Concentrations Seemed to Have Consequences for the Stability of the Plasmids

At Pb 1.5, an increased number of plasmid-free *P. putida* KT2440 was recorded in co-cultures with *D. acidovorans* SPH-1, concomitantly to an increased cost of both plasmids for *P. putida* KT2440. This large fitness decrease most probably combined the negative effects on the growth rate imposed by the plasmid and the high lead concentration, while the *pbr* system was not efficient enough to completely alleviate the negative fitness effect of the plasmid. Surprisingly, when co-cultured with *V. paradoxus* B4, plasmid-free *P. putida* KT2440 cells did not increase drastically. As mentioned above, an up-regulation of siderophore-associated proteins of *V. paradoxus* B4 was observed. Siderophores are public goods – secreted molecules that benefit neighboring individuals – in this case by decreasing the bio-availability of the metal (Hesse et al., 2018). Siderophores might then have efficiently decreased the cumulated fitness effects of the plasmid and lead for both *V. paradoxus* B4 and *P. putida* KT2440.

The *pbr* operon alleviated the cost of the pKJK5 plasmid on *P. putida* KT2440 and the loss of pKJK5 in *P. putida* KT2440 when associated with *D. acidovorans* SPH-1. However, the number of plasmid-free cells was still very high. Interestingly, despite a reduction in potential plasmid donors, transconjugants were still detected among *D. acidovorans* SPH-1. The large spread of the pKJK5-*gfp* plasmid in the *D. acidovorans* SPH-1 recipient population at 1.5 mM lead after 4 days of co-culture concurred with the upregulation of partitioning- and conjugation-involved proteins (TraG, TraC, TrbI, TrbE, IncC1, and KorB), which most likely indicate an increased transfer and a decreased loss of the plasmid. IncC1 and KorB are part of the partitioning system (Rosche et al., 2000), T4CP and TraG are conjugative coupling proteins (T4CP, TraG) and the DNA primase TraC allows plasmid replication (Lanka and Barth, 1981; Bates et al., 1998; Schröder et al., 2002), ensuring the fertility of conjugation (Bates et al., 1998). The high increase in abundance of pKJK5-*gfp* conjugation proteins at Pb-1.5 concurs with signs of oxidative stress demonstrated by the negative impacts of the metal on iron–sulfur cluster-dependent proteins and the overabundant stress proteins (CysQ, MutS, SbcC, mdf) as previously shown (Zhang et al., 2019; Wang et al., 2020; Pu et al., 2021), whereas *pbr* genes most likely decreased this oxidative stress, subsequently reducing the promotion of partitioning- and conjugation-associated genes. In the long-term co-cultures (10 days), plasmid settlement was most likely determined by its fitness effect as plasmid occurrence in the recipient community decreased. Considering the potential effects of oxidative stress, the low fluctuation in the spread index of both plasmids into *V. paradoxus* might have been due to an inexistant stress, thanks to the previously mentioned involvement of siderophore-related proteins. As *pbr* brings no advantage to *V. paradoxus* B4, no fitness gain nor oxidative stress reduction would impact the spread of the plasmid.

These results unraveled factors impacting the plasmid-mediated spread of a MRG at high lead concentration in a two-member community *in vitro*, showing that associated metal differently impacted the spread of the pKJK5 plasmids. The metal-mediated oxidative stress modulated the spread of the plasmid by promoting the expression of conjugation and partitioning proteins. MRGs, if bringing any advantage to the host, would reduce the oxidative stress and (i) increase the fitness benefit of the plasmid but (ii) impede stress-activated conjugation/partitioning protein promotion, yet this assertion should be investigated on a shorter term to measure the actual effect of oxidative stress on plasmid transfer and its subsequent

persistence in a strain population. Moreover, it was previously shown that host plasmid co-adaptation leads to a decreased plasmid cost, which might have influenced plasmid stability in the recipient pool over the 10 days of co-culture.

CONCLUSION

In the present study, when a fitness gain was recorded in the recipient strain, the spread of MRG-carrying plasmids was facilitated through positive selection at an intermediate metal concentration. At high metal concentrations, while metal-mediated oxidative stress increased the abundance of proteins involved in conjugation/partitioning and facilitated the spread of the plasmid, the MRG curbed this oxidative stress and subsequently slowed down plasmid spread. Future studies will test a larger range of strains with different lead resistance potentials to increase the advantage brought by the plasmid-carried MRG, and the fine regulation of the conjugation machinery must be deciphered. Nonetheless, in light of these results, metals most likely influence the journey of plasmids in a diversified recipient community. These data thus represent valuable insights when considering plasmids for metal pollution bioremediation.

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: http://www.peptideatlas.org/, PASS01468.

AUTHOR CONTRIBUTIONS

VC and RW conceived the study. VC and LF performed the experiments and data analysis. JM and BL helped with experiments and assisted with computational analysis. LH, SS, FB, and RW contributed to reagents, materials, and analysis tools. VC, LF, JM, SS, and RW wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was sponsored by the "Belgian Fund for Scientific Research" (Grant equipment-F.R.S-FNRS), by the Pôles d'Attraction Inter-Universitaire (PAI) no. P7/25, and by the Fund for Scientific Research (F.R.S-FNRS) FRFC 7050357, PF014F828/A09016F, and T.0127. The bioprofiling platform used for the proteomic analysis was supported by the European Regional Development Fund and the Walloon Region, Belgium.

ACKNOWLEDGMENTS

We thank Dr. Sandra Ormenese and Raafat Stephan from the Cell Imaging and Flow Cytometry GIGA Platform. We acknowledge Servier Medical Art for providing the images in **Figure 1**.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.655903/full#supplementary-material

REFERENCES

Bahl, M. I., Hansen, L. H., and Sørensen, S. J. (2007). Impact of conjugal transfer on the stability of IncP-1 plasmid pKJK5 in bacterial populations. *FEMS Microbiol. Lett.* 266, 250–256. doi: 10.1111/j.1574-6968.2006.00536.x

Bahl, M. I., Hansen, L. H., and Sørensen, S. J. (2009). "Persistence Mechanisms of Conjugative Plasmids," in *Horizontal Gene Transfer: Genomes in Flux*, eds M. B. Gogarten, J. P. Gogarten, and L. C. Olendzenski (Totowa, NJ: Humana Press), doi: 10.1007/978-1-60327-853-9

Bahl, M. I., Sørensen, S. J., and Hansen, L. H. (2004). Quantification of plasmid loss in *Escherichia coli* cells by use of flow cytometry. *FEMS Microbiol. Lett.* 232, 45–49. doi: 10.1016/S0378-1097(04)00015-1

Bates, S., Cashmore, A. M., and Wilkins, B. M. (1998). IncP plasmids are unusually effective in mediating conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*: involvement of the Tra2 mating system. *J. Bacteriol.* 180, 6538–6543.

Bell, A., and Hancock, R. E. W. (1989). Outer membrane protein H1 of *Pseudomonas aeruginosa*: purification of the protein and cloning and nucleotide sequence of the gene. *J. Bacteriol.* 171, 3211–3217. doi: 10.1128/jb.171.6.3211-3217.1989

Chenier, D., Beriault, R., Mailloux, R., Baquie, M., Abramia, G., Lemire, J., et al. (2008). Involvement of fumarase C and NADH oxidase in metabolic adaptation of *Pseudomonas* fluorescens cells evoked by aluminum and gallium toxicity. *Appl. Environ. Microbiol.* 74, 3977–3984. doi: 10.1128/AEM.02702-07

Cortes, L., Wopereis, H., Tartiere, A., Piquenot, J., Gouw, J. W., Tims, S., et al. (2019). Metaproteomic and 16S rRNA gene sequencing analysis of the infant fecal microbiome. *Int. J. Mol. Sci.* 20, 9–12. doi: 10.3390/ijms20061430

Cyriaque, V., Géron, A., Billon, G., Nesme, J., Werner, J., Gillan, D. C., et al. (2020a). Metal-induced bacterial interactions promote diversity in river-sediment microbiomes. *FEMS Microbiol. Ecol.* 96:5826176. doi: 10.1093/femsec/fiaa076

Cyriaque, V., Jacquiod, S., Riber, L., Abu Al-soud, W., Gillan, D. C., Sørensen, S. J., et al. (2020b). Selection and propagation of IncP conjugative plasmids following long-term anthropogenic metal pollution in river sediments. *J. Hazard. Mater.* 382:121173. doi: 10.1016/j.jhazmat.2019.121173

Garbisu, C., Garaiyurrebaso, O., Epelde, L., Grohmann, E., and Alkorta, I. (2017). Plasmid-mediated bioaugmentation for the bioremediation of contaminated soils. *Front. Microbiol.* 8:1966. doi: 10.3389/fmicb.2017.01966

Gillan, D. C., Danis, B., Pernet, P., Joly, G., and Dubois, P. (2005). Structure of sediment-associated microbial communities along a heavy-metal contamination gradient in the marine environment. *Appl. Environ. Microbiol.* 71, 679–690. doi: 10.1128/AEM.71.2.679

Gillan, D. C., Roosa, S., Kunath, B., Billon, G., and Wattiez, R. (2015). The long-term adaptation of bacterial communities in metal-contaminated sediments: a metaproteogenomic study. *Environ. Microbiol.* 17, 1991–2005. doi: 10.1111/1462-2920.12627

Gougoulias, C., and Shaw, L. J. (2012). Evaluation of the environmental specificity of fluorescence in situ hybridization (FISH) using fluorescence-activated cell sorting (FACS) of probe (PSE1284)-positive cells extracted from rhizosphere soil. *Syst. Appl. Microbiol.* 35, 533–540. doi: 10.1016/j.syapm.2011.11.009

Guo, Q., Li, N., and Xie, S. (2019). Heavy metal spill influences bacterial communities in freshwater sediments. *Arch. Microbiol.* 201, 847–854. doi: 10.1007/s00203-019-01650-y

Hall, J. P. J., Wood, A. J., Harrison, E., and Brockhurst, M. A. (2016). Source–sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities. *Proc. Natl. Acad. Sci. U.S.A.* 113, 8260–8265. doi: 10.1073/pnas.1600974113

Hesse, E., O'Brien, S., Tromas, N., Bayer, F., Luján, A. M., van Veen, E. M., et al. (2018). Ecological selection of siderophore-producing microbial taxa in response to heavy metal contamination. *Ecol. Lett.* 21, 117–127. doi: 10.1111/ele.12878

Hynninen, A., Touzé, T., Pitkänen, L., Mengin-Lecreulx, D., and Virta, M. (2009). An efflux transporter PbrA and a phosphatase PbrB cooperate in a lead-resistance mechanism in bacteria. *Mol. Microbiol.* 74, 384–394. doi: 10.1111/j.1365-2958.2009.06868.x

Jacquiod, S., Brejnrod, A., Morberg, S. M., Abu Al-Soud, W., Sørensen, S. J., and Riber, L. (2017). Deciphering conjugative plasmid permissiveness in wastewater microbiomes. *Mol. Ecol.* 26, 3556–3571. doi: 10.1111/mec.14138

Jacquiod, S., Cyriaque, V., Riber, L., Al-soud, W. A., Gillan, D. C., Wattiez, R., et al. (2018). Long-term industrial metal contamination unexpectedly shaped diversity and activity response of sediment microbiome. *J. Hazard. Mater.* 344, 299–307. doi: 10.1016/j.jhazmat.2017.09.046

Jarosławiecka, A., and Piotrowska-Seget, Z. (2014). Lead resistance in micro-organisms. *Microbiology (United Kingdom)* 160, 12–25. doi: 10.1099/mic.0.070284-0

Juggins, S. (2019). *rioja: Analysis of Quaternary Science Data*.

Jurburg, S. D., Nunes, I., Brejnrod, A., Jacquiod, S., Priemé, A., Sørensen, S. J., et al. (2017). Legacy effects on the recovery of soil bacterial communities from extreme temperature perturbation. *Front. Microbiol.* 8:1832. doi: 10.3389/fmicb.2017.01832

Klümper, U., Dechesne, A., Riber, L., Brandt, K. K., Gülay, A., Sørensen, S. J., et al. (2017). Metal stressors consistently modulate bacterial conjugal plasmid uptake potential in a phylogenetically conserved manner. *ISME J.* 11, 152–165. doi: 10.1038/ismej.2016.98

Klümper, U., Riber, L., Dechesne, A., Sannazzarro, A., Hansen, L. H., Sørensen, S. J., et al. (2015). Broad host range plasmids can invade an unexpectedly diverse fraction of a soil bacterial community. *ISME J.* 9, 934–945. doi: 10.1038/ismej.2014.191

Kommerein, N., Stumpp, S. N., Musken, M., Ehlert, N., Winkel, A., Haussler, S., et al. (2017). An oral multispecies biofilm model for high content screening applications. *PLoS One* 12:e0173973. doi: 10.1371/journal.pone.0173973

Kottara, A., Hall, J. P. J., Harrison, E., and Brockhurst, M. A. (2018). Variable plasmid fitness effects and mobile genetic element dynamics across *Pseudomonas* species. *FEMS Microbiol. Ecol.* 94:fix172. doi: 10.1093/femsec/fix172

Kurth, C., Schieferdecker, S., Athanasopoulou, K., Seccareccia, I., and Nett, M. (2016). Variochelins, lipopeptide siderophores from variovorax boronicumulans discovered by genome mining. *J. Nat. Prod.* 79, 865–872. doi: 10.1021/acs.jnatprod.5b00932

Kwon, M. J., Yang, J. S., Lee, S., Lee, G., Ham, B., Boyanov, M. I., et al. (2015). Geochemical characteristics and microbial community composition in toxic metal-rich sediments contaminated with Au-Ag mine tailings. *J. Hazard. Mater.* 296, 147–157. doi: 10.1016/j.jhazmat.2015.04.035

Lanka, E., and Barth, P. T. (1981). Plasmid RP4 specifies a deoxyribonucleic acid primase involved in its conjugal transfer and maintenance. *J. Bacteriol.* 148, 769–781.

Lawson, T. S., Connally, R. E., Vemulpad, S., and Piper, J. A. (2012). Dimethyl formamide-free, urea-NaCl fluorescence in situ hybridization assay for *Staphylococcus aureus*. *Lett. Appl. Microbiol.* 54, 263–266. doi: 10.1111/j.1472-765X.2011.03197.x

Leroy, B., De Meur, Q., Moulin, C., Wegria, G., and Wattiez, R. (2015). New insight into the photoheterotrophic growth of the isocytrate lyase-lacking purple bacterium rhodospirillum rubrum on acetate. *Microbiology (United Kingdom)* 161, 1061–1072. doi: 10.1099/mic.0.000067

Lopatkin, A. J., Meredith, H. R., Srimani, J. K., Pfeiffer, C., Durrett, R., and You, L. (2017). Persistence and reversal of plasmid-mediated antibiotic resistance. *Nat. Commun.* 8:1689. doi: 10.1038/s41467-017-01532-1

Lopatkin, A. J., Sysoeva, T. A., and You, L. (2016). Dissecting the effects of antibiotics on horizontal gene transfer: analysis suggests a critical role of selection dynamics. *Bioassays* 38, 1283–1292. doi: 10.1002/bies.201600133. Dissecting

Monchy, S., Benotmane, M. A., Janssen, P., Vallaeys, T., Taghavi, S., Van Der Lelie, D., et al. (2007). Plasmids pMOL28 and pMOL30 of *Cupriavidus metallidurans* are specialized in the maximal viable response to heavy metals. *J. Bacteriol.* 189, 7417–7425. doi: 10.1128/JB.00375-07

Naik, M. M., and Dubey, S. K. (2011). Lead-enhanced siderophore production and alteration in cell morphology in a Pb-resistant *Pseudomonas aeruginosa* strain 4EA. *Curr. Microbiol.* 62, 409–414. doi: 10.1007/s00284-010-9722-2

Ni, C., Horton, D. J., Rui, J., Henson, M. W., Jiang, Y., Huang, X., et al. (2016). High concentrations of bioavailable heavy metals impact freshwater sediment microbial communities. *Ann. Microbiol.* 66, 1003–1012. doi: 10.1007/s13213-015-1189-8

Nishida, H., and Oshima, T. (2019). *DNA Traffic in the Environment*. Singapore: Springer Singapore.

Norman, A., Hansen, L. H., and Sørensen, S. J. (2009). Conjugative plasmids: vessels of the communal gene pool. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2275–2289. doi: 10.1098/rstb.2009.0037

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *Vegan: Community Ecology Package. R Package Version 2.5-6*.

Pal, C., Bengtsson-Palme, J., Kristiansson, E., and Larsson, D. G. J. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics* 16:964. doi: 10.1186/s12864-015-2153-5

Parra, B., Tortella, G. R., Cuozzo, S., and Martínez, M. (2019). Negative effect of copper nanoparticles on the conjugation frequency of conjugative catabolic plasmids. *Ecotoxicol. Environ. Saf.* 169, 662–668. doi: 10.1016/j.ecoenv.2018.11.057

Perry, J. A., and Wright, G. D. (2013). The antibiotic resistance "mobilome": searching for the link between environment and clinic. *Front. Microbiol.* 4:138. doi: 10.3389/fmicb.2013.00138

Pinilla-Redondo, R., Cyriaque, V., Jacquiod, S., Sørensen, S. J., and Riber, L. (2018). Monitoring plasmid-mediated horizontal gene transfer in microbiomes: recent advances and future perspectives. *Plasmid* 99, 56–67. doi: 10.1016/j.plasmid.2018.08.002

Pu, Q., Fan, X., Li, H., An, X., Lassen, S. B., and Su, J. (2021). Cadmium enhances conjugative plasmid transfer to a fresh water microbial community. *Environ. Pollut.* 268:115903. doi: 10.1016/j.envpol.2020.115903

Rosche, T. M., Siddique, A., Larsen, M. H., and Figurski, D. H. (2000). Incompatibility protein IncC and global regulator KorB interact in active partition of promiscuous plasmid RK2. *J. Bacteriol.* 182, 6014–6026. doi: 10.1128/JB.182.21.6014-6026.2000

Schröder, G., Krause, S., Zechner, E. L., Traxler, B., Yeo, H. J., Lurz, R., et al. (2002). TraG-like proteins of DNA transfer systems and of the *Helicobacter* pylori type IV secretion system: inner membrane gate for exported substrates? *J. Bacteriol.* 184, 2767–2779. doi: 10.1128/JB.184.10.2767-2779.2002

Smets, B. F., Morrow, J. B., and Pinedo, C. A. (2003). Plasmid introduction in metal-stressed, subsurface-derived microcosms: plasmid fate and community response. *Appl. Environ. Microbiol.* 69, 4087–4097. doi: 10.1128/AEM.69.7.4087

Sun, M. Y., Dafforn, K. A., Johnston, E. L., and Brown, M. V. (2013). Core sediment bacteria drive community response to anthropogenic contamination over multiple environmental gradients. *Environ. Microbiol.* 15, 2517–2531. doi: 10.1111/1462-2920.12133

Taghavi, S., Lesaulnier, C., Monchy, S., Wattiez, R., Mergeay, M., and Lelie, D. (2009). Lead(II) resistance in Cupriavidus metallidurans CH34: Interplay between plasmid and chromosomally-located functions. *Antonie van Leeuwenhoek Int. J. Gen. Mol. Microbiol.* 96, 171–182. doi: 10.1007/s10482-008-9289-0

Wang, P., Zhang, S., Wang, C., and Lu, J. (2012). Effects of Pb on the oxidative stress and antioxidant response in a Pb bioaccumulator plant *Vallisneria natans*. *Ecotoxicol. Environ. Saf.* 78, 28–34. doi: 10.1016/j.ecoenv.2011.11.008

Wang, Q., Liu, L., Hou, Z., Wang, L., Ma, D., Yang, G., et al. (2020). Heavy metal copper accelerates the conjugative transfer of antibiotic resistance genes in freshwater microcosms. *Sci. Total Environ.* 717:137055. doi: 10.1016/j.scitotenv.2020.137055

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W. H. A., Lumley, T., et al. (2016). *Package "gplots": Various R programming Tools for Plotting Data. R Packag. Version 2.17.0*. 1–68. doi: 10.1111/j.0022-3646.1997.00569.x

Xu, F. F., and Imlay, J. A. (2012). Silver(I), mercury(II), cadmium(II), and zinc(II) target exposed enzymic iron-sulfur clusters when they toxify *Escherichia coli*. *Appl. Environ. Microbiol.* 78, 3614–3621. doi: 10.1128/AEM.07368-11

Yamaguchi, N., Ohba, H., and Nasu, M. (2006). Simple detection of small amounts of *Pseudomonas* cells in milk by using a microfluidic device. *Lett. Appl. Microbiol.* 43, 631–636. doi: 10.1111/j.1472-765X.2006.02013.x

Zhang, S., Wang, Y., Song, H., Lu, J., Yuan, Z., and Guo, J. (2019). Copper nanoparticles and copper ions promote horizontal transfer of plasmid-mediated multi-antibiotic resistance genes across bacterial genera. *Environ. Int.* 129, 478–487. doi: 10.1016/j.envint.2019.05.054

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 21 July 2021
doi: 10.3389/fmicb.2021.709259

The Selective Advantage of the *lac* Operon for *Escherichia coli* Is Conditional on Diet and Microbiota Composition

Catarina Pinto[1†], Rita Melo-Miranda[2†], Isabel Gordo[3] and Ana Sousa[2]*

[1] CEDOC, Chronic Diseases Research Centre, NOVA Medical School, Faculdade de Ciências Médicas, Universidade NOVA de Lisboa, Lisbon, Portugal, [2] Department of Medical Sciences, Institute of Biomedicine (iBiMED), University of Aveiro, Aveiro, Portugal, [3] Instituto Gulbenkian de Ciência, Oeiras, Portugal

The *lac* operon is one of the best known gene regulatory circuits and constitutes a landmark example of how bacteria tune their metabolism to nutritional conditions. It is nearly ubiquitous in *Escherichia coli* strains justifying the use of its phenotype, the ability to consume lactose, for species identification. Lactose is the primary sugar found in milk, which is abundant in mammals during the first weeks of life. However, lactose is virtually non-existent after the weaning period, with humans being an exception as many consume dairy products throughout their lives. The absence of lactose during adulthood in most mammals and the rarity of lactose in the environment, means that the selective pressure for maintaining the *lac* operon could be weak for long periods of time. Despite the ability to metabolize lactose being a hallmark of *E. coli*'s success when colonizing its primary habitat, the mammalian intestine, the selective value of this trait remains unknown in this ecosystem during adulthood. Here we determine the competitive advantage conferred by the *lac* operon to a commensal strain of *E. coli* when colonizing the mouse gut. We find that its benefit, which can be as high as 11%, is contingent on the presence of lactose in the diet and on the presence of other microbiota members in the gut, but the operon is never deleterious. These results help explaining the pervasiveness of the *lac* operon in *E. coli*, but also its polymorphism, as *lac*-negative *E. coli* strains albeit rare can naturally occur in the gut.

Keywords: *lac* operon, fitness effect, *Escherichia coli*, *Lactobacillus murinus*, *Bacteroides thetaiotaomicron*, lactose, gut microbiota, microbe–microbe interactions

OPEN ACCESS

Edited by:
Maria Elena Martino, University of Padua, Italy

Reviewed by:
Åsa Sjöling, Karolinska Institutet (KI), Sweden
Scott Miller, University of Montana, United States

***Correspondence:**
Ana Sousa
amsousa@ua.pt

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *13 May 2021*
Accepted: *28 June 2021*
Published: *21 July 2021*

Citation:
Pinto C, Melo-Miranda R, Gordo I and Sousa A (2021) The Selective Advantage of the lac Operon for Escherichia coli Is Conditional on Diet and Microbiota Composition. Front. Microbiol. 12:709259. doi: 10.3389/fmicb.2021.709259

INTRODUCTION

The *lac* operon, first described by Jacob and Monod (1961), codes for the cellular machinery to transport and metabolize lactose. It is constituted by three structural genes: *lacZ* codes for the lactose-degrading enzyme β-galactosidase, that breaks down lactose into glucose and galactose; *lacY* encodes the β-galactoside permease, that facilitates lactose transport into the cell; and *lacA* encodes galactoside transacetylase, whose function in lactose metabolism remains unknown (Müller-Hill, 1996b; Roderick, 2005). The *lac* operon is expressed under the presence of lactose and low levels of

glucose (Müller-Hill, 1996a), being the textbook example of gene regulation at the transcriptional level in prokaryotes. As it is present in the vast majority of *Escherichia coli* strains (Stoebel, 2005), the phenotype of lactose consumption was used to identify *E. coli* among environmental samples (Hartl and Dykhuizen, 1984).

Lactose is the primary sugar in milk (Holsinger, 1988) and abundant in the mammalian gut during breastfeeding. However, lactose is virtually absent after that period. In fact, mammals become lactose-intolerant in adulthood, except for about a third of the human population, which exhibits lactase persistence into adulthood resulting from selection in dairy farming cultures (Storhaug et al., 2017). Thus, given the absence of lactose during adulthood in most mammals and also its scarcity in the environment, the selective pressure for preserving the *lac* operon could be weak. Yet, β-galactosidase can also degrade other compounds, including galactosylglycerols, which result from the pancreatic degradation of galactolipids (Egel, 1979, 1988). As galactolipids are main components of the chloroplast membranes, the frequent ingestion of green leaves could act as a selective pressure to maintain the *lac* operon in *E. coli* beyond the breastfeeding period. Studies on the origin of the operon were not able to trace its occurrence to a horizontal gene transfer event (Stoebel, 2005), further supporting a continuous pressure to keep the operon and the hypothesis of it being an important component of the species identity.

E. coli's primary habitat is the mucus layer of the mammalian intestinal tract, where it grows in a multi-species biofilm and accounts for around 0.1% of the human gut microbiota population (Raman et al., 2005). *E. coli*, together with streptococci, is one of the first bacteria to colonize the intestinal track, being established as soon as 48 h after birth (Mackie et al., 1999). The large amounts of lactose present in the gut at this stage might confer *E. coli* an advantage for colonizing this environment (Ochman et al., 2000; Blount, 2015).

Still, the actual selective value of the *lac* operon for *E. coli* in this niche, whether it increases in the presence of lactose, and how it changes according to microbiota composition, remains undetermined. To tackle these questions, we determined the competitive advantage given by the *lac* operon to a commensal strain of *E. coli* when colonizing the mouse intestine both in the presence and absence of lactose. We found that the ability to consume lactose is mainly beneficial in the presence of lactose and that this benefit can have an individualized effect, possibly reflecting differences in microbiota composition.

To disentangle the effect of lactose from the effect of the microbiota we next performed similar competitions in animals devoid of microbiota [germ free (GF) mice]. To our surprise, the *lac* operon was neutral in this scenario, irrespectively of the presence of lactose. We hypothesized that the overabundance of nutritional sources and the lack of interspecies competition could be responsible for this result. We corroborate this hypothesis via a series of competitions varying the level of inter and intraspecies competition and show that the selective benefit of the *lac* operon depends on the presence of lactose and also gradually increases with the number of different species present in the gut.

MATERIALS AND METHODS

Mice

All experiments were conducted using female 6–8 weeks old C57BL6/6J mice under germ free (GF) or specific pathogen-free (SPF) conditions at the Instituto Gulbenkian de Ciência (IGC) animal facilities. Mice were individually caged and had access to water and food *ad libitum*. GF mice were bred and raised at the IGC gnotobiology facility in dedicated axenic isolators (La Calhene/ORM). Young adults were transferred into sterile ISOcages (Tecniplast) before the competition experiments. This research project was approved by both the Ethics Committee and the IGC Animal Welfare Body (license reference: A009.2018), and by the Portuguese National Entity that regulates the use of laboratory animals (DGAV—Direção Geral de Alimentacão e Veterinária; license reference: 009676). All experiments were performed following the Portuguese (Decreto-Lei n° 113/2013) and European (Directive 2010/63/EU) legislations concerning animal welfare.

Bacterial Strains

The two *E. coli* strains used in this work were derived from *E. coli* K-12 MG1655 and were resistant to streptomycin due to the point mutation *rpsL* K43T (Str^R). The two strains are isogenic except for the deletion of the *lacZ* (Δ*lacZ, lac*−), differing only in the ability to metabolize lactose. *L. murinus* was isolated from feces of SPF mice and species identity was confirmed by sequencing the whole 16S rRNA gene. *Bacteroides thetaiotaomicron* strain (VPI-5482) was acquired from the German Collection of Microorganisms and Cell Cultures (DSMZ 2079).

Mouse Gut Colonizations

To colonize the mouse intestine, we used a streptomycin treated colonization model (Conway et al., 2004) under two antibiotic regimes, i.e., treatment lasted for 1 or 7 days. GF animals were treated continuously with streptomycin. SPF animals were colonized with *E. coli,* while GF mice were either mono-colonized (*E. coli*), double-colonized (*E. coli* + *L. murinus*), or triple-colonized (*E. coli* + *L. murinus* + *B. thetaiotaomicron*). In all colonization conditions, mice were given autoclaved drinking water supplemented with streptomycin (5 g/L) for 24 h.

After 4 h of starvation for water and food, animals were gavaged with 100 μL of a suspension of 10^8 colony forming units (CFUs) of a mixture (1:1) of *lac*− and *lac* + (SPF and mono-colonization). Double and triple colonizations started with the gavage of 100 μL of a suspension of 10^8 CFU containing equal amounts of *E. coli* and each of the other species.

Bacteria for gavage were prepared as follows: *E. coli* was grown in LB broth overnight with aeration followed by a 1:100 dilution and grown in BHI broth until reaching an optical density of 2. *L. murinus* was inoculated in De Man, Rogosa and Sharpe (MRS) broth and incubated in anaerobiosis. *B. thetaiotaomicron* was inoculated in Chopped Meat Medium with Carbohydrates and grown in anaerobiosis. All bacteria were incubated overnight at 37°C and 1 mL of each culture was centrifuged and resuspended in PBS 1x + 0.1% cysteine.

Water with streptomycin was maintained throughout the 6 days of experiment and replaced after 3 days, except for the colonization where streptomycin was only administered for 24 h before gavage. In half of the animals, the water was also supplemented with 2% lactose. This percentage was chosen to mimic to the concentration found in mice milk (Görs et al., 2009). Fecal samples were collected and weighed daily. Samples were then suspended in PBS 1x and appropriate dilutions were then inoculated in MacConkey agar supplemented with 100 μg/mL of streptomycin (for *E. coli*) and MRS agar (for *L. murinus*) or in LB agar with X-Gal and 100 μg/mL of streptomycin. *E. coli* and *L. murinus* were grown at 37°C for 18 and 24 h, respectively in aerobiosis. When grown in MacConkey, *E. coli lac* + and *lac*− strains were discriminated by colony color (red colonies are *lac* +, white colonies are *lac*−).

Competitive Fitness Assays *in vivo* to Measure the Selective Coefficient of the *lac* Operon

To assess if the *lac* operon confers an advantage to *E. coli* in the tested conditions, total numbers and relative frequencies of the bacteria were determined by counting the number of colonies of each strain. The selection coefficient (fitness gain) per generation of the *lac* + strain *in vivo* was calculated as the slope of the linear regression of ln ($freq_{lac+}/freq_{lac-}$) divided by the number of generations for *E. coli* per day. The number of days considered to calculate the selection coefficient in the colonization of SPF mice (1 day of streptomycin) was 2 ($n = 3$), 3 ($n = 1$), 4 ($n = 3$), or 5 ($n = 1$), to use a time interval with a linear tendency of selection. For the remaining colonizations, the first 3 days were considered. As for the number of generations per day, we considered this value to be 18 based on the previous estimation of 80-min generation time for *E. coli* in streptomycin-treated mice (Rang et al., 1999).

Microbiota Analysis

To characterize the effect of different streptomycin regimes in the microbiota, we analyzed the V4 region of 16S rRNA gene sequenced from fecal samples from young animals before and after antibiotic treatment (days 7/8 of the evolution experiment). 16S sequence data from animals treated with streptomycin for only 1 day is available from Frazão et al. (2019) and 16S sequence data from animals treated with continuous streptomycin is available from Barreto et al. (2020).

Raw reads were processed using QIIME2 version 2020.8 with default parameters (Bolyen et al., 2019). Deblur was used for quality filtering and denoising (Amir et al., 2017). The generated table of ASVs was then used in the R package *phyloseq* (McMurdie and Holmes, 2013) to determine the Bray-Curtis dissimilarity index, using rarefaction based on the sample with the lowest sequencing depth. Within-group distance based on Bray-Curtis was calculated in QIIME2.

The observed ASVs were also calculated in QIIME2 with *diversity core-metrics-phylogenetic* and *diversity alpha-group-significance,* also using rarefaction based on the sample with the lowest sequencing depth.

Statistical Analysis

All graphical representations were performed in GraphPad Prism 8.2.1 for Mac[1] (GraphPad Software, San Diego, California, United States) and statistical analysis in Prism or in R version 3.6.3 (R Core Team, 2020). Normal data distribution was assessed using the Shapiro-Wilk test and QQ plots observation, and the homogeneity of variances using Levene's test. Statistical significance was defined for $p < 0.05$. Significance levels were defined as: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$. The detailed statistics and sample size for each experiment are described in the figure legends.

RESULTS

The Selective Advantage of the *lac* Operon Is Stronger on the Presence of Lactose and Reflects the Amount of Competition

The selective advantage conferred by the ability to metabolize a given nutrient is expected to change with its relative abundance, its energetic potential, and the amount of competition for that nutrient.

Here we evaluated the selective advantage conferred by the *lac* operon to an *E. coli* strain (*lac* +), by allowing for its competition with a *lac*− strain in the intestine of mice after a short treatment (1 day) with streptomycin. This procedure maintains a complex microbiota while allowing for *E. coli* colonization (Frazão et al., 2019). We started by colonizing 8 animals with a mixture of 1:1 (*lac* + and *lac*−) isogenic strains, except for the deletion of the *lacZ* in the *lac*- strain, by oral gavage. We further supplemented the diet of 4 animals with 2% lactose whereas the remaining 4 animals were fed regular chow. The frequencies of *lac* +/*lac*− strains were assessed by daily collecting fecal samples and plating appropriate dilutions for 6 days.

During this period, in the absence of lactose, the frequency of *lac* + remained close to the initial 50% in two animals and increased slightly in the other two (**Figures 1A,B**), showing a mean advantage of 1 ± 0.6% (SEM) (**Figure 2**).

The presence of lactose had a significant effect in the *lac* + frequency over time [**Figures 1A,B**, $F(6, 36) = 5.63$, $p = 0.0003$] and a more variable effect was observed for the selective advantage of the *lac* + strain, ranging from 1.4 to 11%, with a mean of 5 ± 4% (SEM) (**Figure 2**).

Besides the presence of lactose, differences between measurements likely reflect the common variation in microbiota composition between individuals which is known to further increase with the antibiotic treatment (Leónidas Cardoso et al., 2020). The effect of the chosen antibiotic regime was assessed by estimating the Bray-Curtis beta diversity, a measure of group heterogeneity, before and after antibiotic treatment in an independent set of 5 animals. A mean increase of ∼30% in Bray-Curtis diversity was observed after antibiotic treatment (**Supplementary Figures 1A,B**) possibly contributing to the

[1] www.graphpad.com

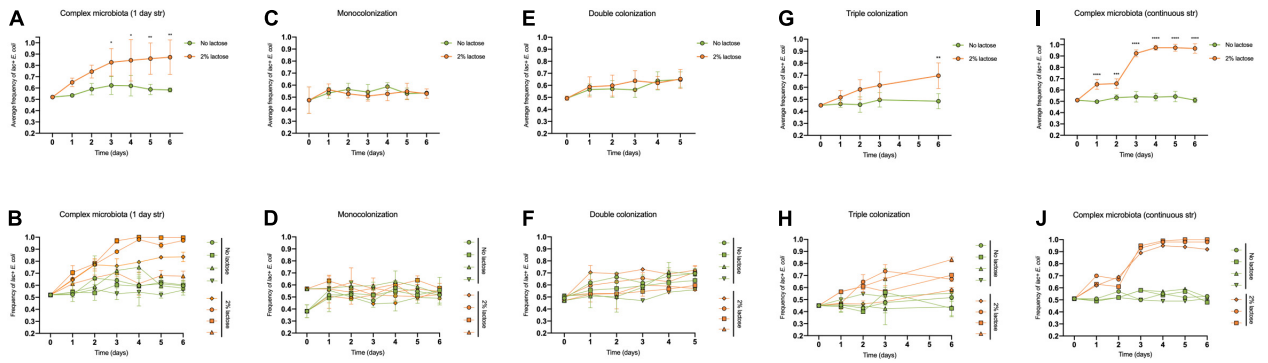


FIGURE 1 | The advantage of the *lac* operon is dependent on diet and the microbiota composition. Frequency of *lac* + *E. coli* when the diet is supplemented or not with lactose, during the first 6 days of colonization of the mouse gut, with microbiotas of different complexity: **(A,B)** complex microbiota (SPF mice) after 1 day of streptomycin ($n = 4$ per group); **(C,D)** monocolonization with *E. coli* ($n = 4$ per group); **(E,F)** double colonization with *E. coli* and *L. murinus* ($n = 4$ per group); **(G,H)** triple colonization with *E. coli*, *L. murinus* and *B. thetataiomicron* ($n = 4$ per group); **(I,J)** complex microbiota (SPF mice) with continuous streptomycin treatment ($n = 3$–4 per group). In the top panels **(A,C,E,G,I)**, values are represented as mean ± SD. In the bottom panels, each animal is individually represented, and the error bars correspond to experimental replicates ($n = 3$). The effects of the microbiota composition and time on the *lac* + *E. coli* frequency were tested by a linear mixed model followed by a type III ANOVA and *post hoc* contrasts with Bonferroni's correction.

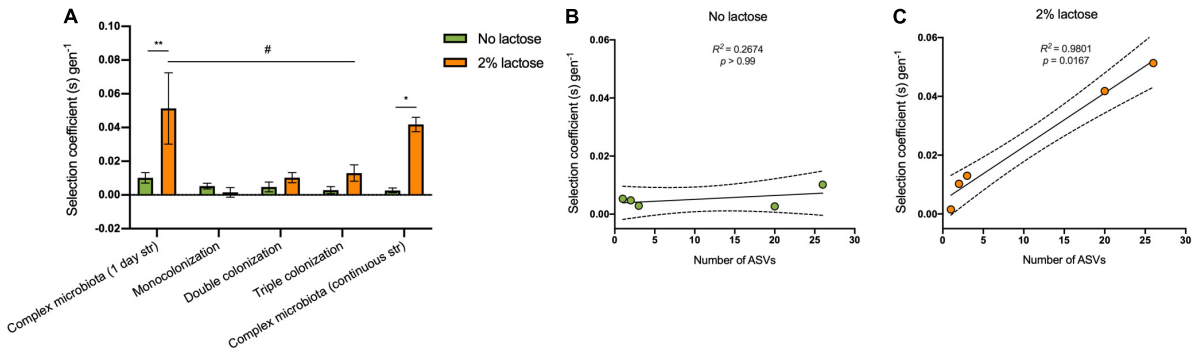


FIGURE 2 | The selective advantage provided by the *lac* operon in the presence of lactose increases with microbiota complexity. **(A)** Selection coefficient per generation of the *lac* + *E. coli* strain when the diet is supplemented or not with lactose, with microbiotas of different complexity: Complex microbiota (SPF mice) after 1 day of streptomycin ($n = 4$ per group); Monocolonization with *E. coli* ($n = 4$ per group); double colonization with *E. coli* and *L. murinus* ($n = 4$ per group); triple colonization with *E. coli*, *L. murinus* and *B. thetataiomicron* ($n = 4$ per group); complex microbiota (SPF mice) with continuous streptomycin treatment ($n = 3$–4 per group). Values are represented as mean ± SEM. The effects of the microbiota composition and lactose on the selection coefficient were tested by a Two-Way ANOVA followed by *post hoc* selected contrasts with Bonferroni's correction. **(B,C)** Linear regression of the number of ASVs vs. the selection coefficient per generation in the **(B)** absence and **(C)** presence of lactose. *P*-values refer to the significance of the Spearman's correlation. The dashed lines represent the 95% confidence intervals. $^{*}p < 0.05$, $^{**}p < 0.01$ refers to comparisons between the presence or absence of lactose within each microbiota composition. $^{\#}p < 0.05$ refers to comparisons between the different microbiota compositions in the presence of lactose.

variation in the fitness effect of the *lac* operon in the gut of different mice.

These results led us to conclude that the *lac* operon can afford a significant advantage to an *E. coli* strain while colonizing the intestine, but its benefit is influenced by lactose and possibly microbiota composition.

To isolate the effect of lactose and test for the influence of a microbiota diversity to the selective coefficient of the *lac* operon, we engaged in a series of colonizations within a range of microbiota diversities. These comprised a mono, double and a triple colonization. A group of animals continuously treated with antibiotic (7 days) was further included, representing a complex yet simpler microbiota than the group treated only for a short period (1 day).

Although not statistically significant, after antibiotic treatment the two regimes differed in microbiota richness, with an average of 20 (continuous regime) and 26 (1 day regime) ASVs (**Supplementary Figure 1C** and **Supplementary Table 1**).

These microbiota compositions likely represent different levels of competition to *E. coli*, which could influence the abundance of lactose and other nutritional sources.

To reduce competition to a minimum we monoassociated GF animals with a mixture of 1:1 (*lac* +/*lac*−) *E. coli* and followed its frequency for a period of 6 days. As before, a group of animals received a lactose supplemented diet ($n = 4$) whereas the remaining animals were fed regular chow. Interestingly we observed that when *E. coli* is the only inhabitant of the gut, the ability to metabolize lactose, irrespectively

of the presence of its cognate substrate, is a neutral trait (**Figures 1C,D**, **2**).

Though it could be hypothesized that in this simple environment the available glucose could preclude the expression of the *lac* operon by catabolite repression this seems unlikely since glucose is less abundant in the cecum of monocolonized than in SPF mice (Barroso-Batista et al., 2020), where the *lac* operon is presumably expressed.

We proceeded by co-colonizing GF animals with *E. coli* and similar numbers of a representative species of Firmicutes, *Lactobacilus murinus,* which belongs to lactic acid bacteria, thus can compete for lactose (de Vos and Vaughan, 1994). As before, the *lac* operon conferred no particular advantage both in the presence or absence of lactose (**Figures 1E,F**, **2**).

Therefore, to further increase the competition level, we performed a triple colonization of GF animals with equal amounts of *E. coli*, *L. murinus,* and *Bacteroides thetaiotaomicron*, which is a representative species of the phylum Bacteroidetes, usually present in large numbers in the gut microbiota and that can also compete for lactose (Chia et al., 2020).

The addition of a third species was enough to restore the advantage of the *lac* operon in the presence of lactose [**Figures 1G,H**, $F(4, 24) = 4.41$, $p = 0.0082$], providing a selective advantage around 1.3 ± 1% (SEM) per generation (**Figure 2**). Nevertheless, we cannot rule out the possibility that *B. thetaiotaomicron per se*, and not just the number of species, was responsible for the increase in the effect of the *lac* operon since we did not perform a colonization including only *E. coli* and *B. thetaiotaomicron*.

Coherently, in the context of continuously antibiotic treated mice and in the presence of lactose, the effect of the *lac* operon was further increased showing a mean advantage of 4.2 ± 1% per generation (**Figure 2**) which allowed it to reach a frequency close to fixation by day 4 (**Figures 1I,J**). Interestingly, the variation in *lac* + selective coefficient between animals was also reduced in comparison to the 1 day antibiotic regime (**Figure 2**), in agreement with the smaller effect of the antibiotic in group heterogeneity in the continuous regime [**Supplementary Figures 1A,B**, $p = 0.0007$ (Kruskal-Wallis test, *post hoc* comparisons with Dunn's correction)].

Globally, these experiments demonstrate that the benefit of the *lac* operon is potentiated by the presence of lactose and increases with the number of co-colonizers (**Figures 2B,C**).

The Absolute Success of *E. coli* in the Gut Depends on the Net Result Between Synergism and Competition

Increasing microbiota diversity likely promotes competition but it also raises the probability of synergism, e.g., via cross-feeding.

Here we observe that *E. coli*'s abundance (measured by its loads—CFU/g of feces) increases with the number of co-colonizers in ex-GF mice, but then it drops when the gut environment reaches the complexity of a regular microbiota (**Figure 3**). Specifically, when comparing the double with the mono colonization we observed a non-significant but sustained increase in the average loads of *E. coli* in the double colonization

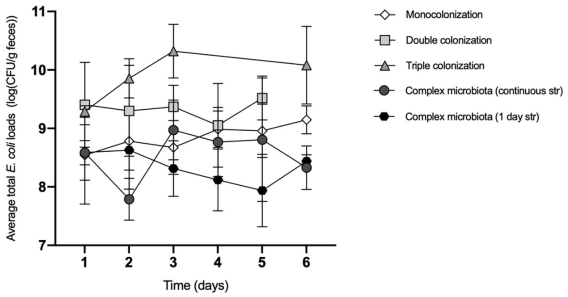


FIGURE 3 | The total *E. coli* loads vary with the composition of the microbiota. Average *E. coli* total loads ($\log_{10}$ transformed) during the first 6 days of colonization of the mouse gut, with increasing microbiota complexity: Monocolonization with *E. coli*, $n = 8$; Double colonization with *E. coli* and *L. murinus*, $n = 6$; Triple colonization with *E. coli*, *L. murinus* and *B. thetataiomicron*, $n = 8$; Complex microbiota with continuous streptomycin treatment, $n = 7$; Complex microbiota after 1 day of streptomycin, $n = 8$. Values are represented as mean ± SD. The effects of the microbiota composition and time on the total *E. coli* loads were tested by a linear mixed model followed by type III ANOVA. *Post hoc* comparisons were performed between all the levels of microbiota complexity for day 5 and 6 (Bonferroni's correction). In day 5, *E. coli* loads in the complex microbiota (continuous streptomycin) are significantly lower than: the monocolonization ($p = 0.0036$); double colonization ($p < 0.0001$); and the complex microbiota (continuous str) ($p = 0.038$). In day 6, *E. coli* loads in the triple colonization are higher than the monocolonization ($p = 0.0117$) and the complex microbiota (both 1 day and continuous str) ($p < 0.0001$).

in 4 out of the 5 days of the experiment (**Figure 3**). Upon 5 days of colonization *E. coli* was on average 3 times more abundant in the double ($\sim 4.4 \times 10^9$ CFU/g) than in the mono colonization ($\sim 1.3 \times 10^9$ CFU/g).

The sharpest increase was observed when comparing the mono with the triple colonization where a significant difference ($\sim$10 times) in the average loads of *E. coli* was measured (2×10^{10} and 1.6×10^9 CFU/g at the 6th day of the experiment, in the triple and mono colonization, respectively ($p = 0.012$, *post hoc* comparison with Bonferroni's correction).

These data suggest that *E. coli*'s niche is expanded by the presence of *L. murinus* and *B. thetaiotaomicron*, hinting at a synergistic relationship between these bacteria and *E. coli*.

In contrast, the expansion of *E. coli* in an environment with multiple species of bacteria, such as the microbiota of streptomycin-treated mice, regardless of the antibiotic regime, was limited to $\sim 10^8$ CFU/g upon 6 days of colonization, likely due to the predominance of interspecies competition.

These differences are similar for the *lac* + and *lac*− strains separately (**Supplementary Figure 2**) and are independent of the presence of lactose in the mouse diet (**Supplementary Figure 3**).

DISCUSSION

The nearly ubiquitous presence of the *lac* operon in *E. coli* and the restricted timeframe of lactose availability in its natural environment, the mammalian gut, has for long raised questions about the adaptive value of the *lac* operon in *E. coli.* Breastfeeding

represents only a brief period in mammalians' life, but it is particularly influential in determining the first stages of microbiota establishment (Milani et al., 2017), possibly with long-term effects. *E. coli* represents a small fraction of the adult microbiota (Raman et al., 2005), but is among the first colonizers of the gut (Milani et al., 2017) and therefore, consuming lactose, could in principle, contributes to its success. In contrast, it has it has been shown that lactose consumption is not essential for *E. coli* to colonize the intestine of streptomycin-treated SPF CD-1 mice, while essential sugars for either the initiation or maintenance of gut colonization are arabinose, fucose, gluconate, N-acetyl-glucosamine, and N-acetyl-neuraminic acid (Fabich et al., 2008).

In our work, by allowing for direct competition between *lac* + and *lac*− *E. coli* strains in the adult mammalian gut we found that, in the presence of lactose, the *lac* operon confers a selective advantage that can be as high as 11%, while being close to neutral in its absence. Therefore, our results suggest that the strong interspecies competition for nutritional resources that occurs in these settings (Coyte et al., 2015) considerably increases the advantage for lactose consumers, particularly in the presence of lactose.

The emergence of constitutive mutants for lactose consumption during the first days of life was observed in a controlled experiment (Ghalayini et al., 2019), further supporting the hypothesis that the *lac* operon is under strong selection during this period, which in the long term could help *E. coli* to secure a place in the complex adult microbiota.

Besides the presence of lactose, the fitness advantage of the *lac* operon for *E. coli* was also shown to depend on microbiota composition, particularly on the number of species present. One way to interpret this result is to consider that an increasing number of species increases the opportunity for more competitive interactions. Thus, the ability to consume lactose would become a bigger advantage as other nutritional sources get depleted by the growing number of competitors. On the other hand, as the number of species increases, so does the opportunity for synergism, with the absolute success of *E. coli* reflecting the net result between these two forms of ecological interactions.

The observation that the population size of *E. coli* substantially increased in the triple-colonization, in comparison to the situation where it is alone (mono-colonization), showed us that, from the perspective of *E. coli*, synergism dominated in this scenario. This is consistent with the notion that *E. coli* consumes the byproducts of complex nutrients' degraded by *B. thetaiotaomicron*, such as maltose and fucose which result from the breakdown of dietary glycans (Chang et al., 2004; Li et al., 2015; Baümler and Sperandio, 2016). This seems to occur regardless of the ability to metabolize lactose, and it is independent of the availability of that additional carbohydrate.

In contrast to our findings, in a scenario of DSS-induced colitis *E. coli* was antagonized by *B. thetaiotaomicron* and another species of *Lactobacillus* (*L. johnsonii*). Here the oral administration of these species was enough to reduce the overgrowth of *E. coli* (Charlet et al., 2020). The higher microbiota complexity of the animals in the referred study when compared to ours, as well as potentially different strains, may account for this difference.

Conversely, in the presence of a complex microbiota (several tens of species), its absolute success decreased and the most likely scenario was one dominated by competition (Coyte et al., 2015). A scenario where the population size of *E. coli* in the gut is inversely correlated with microbiota diversity has been previously observed in more than one occasion (Sousa et al., 2017).

Surprisingly, when comparing the mono with the triple colonization, lactose showed to be more advantageous in the situation where *E. coli*'s absolute success was higher, suggesting that the fitness effect of the *lac* operon may be density dependent. Future work should address this issue, possibly through investigating the geographic distribution of bacteria abundances and possible differences in the *lac* operon selective advantage along the intestinal tract. Since lactose is mainly absorbed in the small intestine, this could be the place where *lac* + is positively selected with the cecum being the intestinal compartment where *E. coli* expands irrespectively of its ability to consume lactose. Such a hypothetical situation would simultaneously explain the larger advantage of a *lac* + strain and the higher absolute success of the species.

Another interesting observation was the correlation between the variance of the *lac* operon selective effect and microbiota heterogeneity. This was unraveled by the two antibiotic regimes but probably best resemble the natural diversity of the microbiota of animals which are not genetically homogenous or eat the same controlled diet as in our experimental setup.

Overall, the results indicate that the specificity of interactions established by *E. coli*, and, therefore, the balance between beneficial and antagonistic interactions, determines its niche size and location. The adaptive value of the *lac* operon for *E. coli* is dependent on the competitors and the availability of nutritional resources and might be especially useful in highly competitive environments and for exploring new niches.

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.

ETHICS STATEMENT

The animal study was reviewed and approved by Direção Geral de Alimentacão e Veterinária (license reference: 009676) and Instituto Gulbenkian de Ciência Animal Welfare Body (license reference: A009.2018).

AUTHOR CONTRIBUTIONS

CP, IG, and AS conceived and designed the experiments. CP performed the experiments. CP, RM-M, and AS analyzed

the data. IG and AS contributed reagents, materials, and analysis tools. RM-M and AS wrote the manuscript. CP and IG contributed to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Fundação para a Ciência e Tecnologia (PTDC/BIA-EVF/3247/2012 and PTDC/BIA-EVL/30212/2017) and iBiMED (UID/BIM/04501/2019, project POCI-01-0145-FEDER-007628). CP was the recipient of a fellowship (138/BI/13) from Fundação para a Ciência e Tecnologia. RM-M was the recipient of a doctoral fellowship (DFA/BD/5130/2020) from Fundação para a Ciência e Tecnologia. AS was funded from National Funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., under the Scientific Employment Stimulus - Institutional Call - reference CEECINST/00026/2018.

ACKNOWLEDGMENTS

We would like to acknowledge Karina Bivar Xavier for providing the strain *Bacteroides thetaiotaomicron*.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.709259/full#supplementary-material

REFERENCES

Amir, A., Daniel, M., Navas-Molina, J., Kopylova, E., Morton, J., Xu, Z. Z., et al. (2017). Deblur rapidly resolves single-nucleotide community sequence patterns. *Am. Soc. Microbiol.* 2, 1–7.

Barreto, H. C., Sousa, A., and Gordo, I. (2020). The landscape of adaptive evolution of a gut commensal bacteria in aging Mice. *Curr. Biol.* 30, 1102.e5–1109.e5. doi: 10.1016/j.cub.2020.01.037

Barroso-Batista, J., Pedro, M. F., Sales-Dias, J., Pinto, C. J. G., Thompson, J. A., Pereira, H., et al. (2020). Specific eco-evolutionary contexts in the mouse gut reveal *Escherichia coli* metabolic versatility. *Curr. Biol.* 30, 1049.e7–1062.e7. doi: 10.1016/j.cub.2020.01.050

Baümler, A. J., and Sperandio, V. (2016). Interactions between the microbiota and pathogenic bacteria in the gut. *Nature* 535, 85–93. doi: 10.1038/nature18849

Blount, Z. D. (2015). The unexhausted potential of *E. coli*. *elife* 4, 1–12. doi: 10.7554/eLife.05826

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Chang, D. E., Smalley, D. J., Tucker, D. L., Leatham, M. P., Norris, W. E., Stevenson, S. J., et al. (2004). Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7427–7432. doi: 10.1073/pnas.0307888101

Charlet, R., Bortolus, C., Sendid, B., and Jawhara, S. (2020). *Bacteroides* thetaiotaomicron and Lactobacillus johnsonii modulate intestinal inflammation and eliminate fungi via enzymatic hydrolysis of the fungal cell wall. *Sci. Rep.* 10, 1–13. doi: 10.1038/s41598-020-68214-9

Chia, L. W., Mank, M., Blijenberg, B., Aalvink, S., Bongers, R. S., Stahl, B., et al. (2020). *Bacteroides* thetaiotaomicron fosters the growth of butyrate-producing *Anaerostipes caccae* in the presence of lactose and total human milk carbohydrates. *Microorganisms* 8:1513.

Conway, T., Krogfelt, K. A., and Cohen, P. S. (2004). The life of commensal *Escherichia coli* in the mammalian intestine. *EcoSal Plus* 1, 1–16. doi: 10.1128/ecosalplus.8.3.1.2

Coyte, K. Z., Schluter, J., and Foster, K. R. (2015). The ecology of the microbiome: networks, competition, and stability. *Science* 350, 663–666. doi: 10.1126/science.aad2602

de Vos, W. M., and Vaughan, E. E. (1994). Genetics of lactose utilization in lactic acid bacteria. *FEMS Microbiol. Rev.* 15, 217–237. doi: 10.1016/0168-6445(94)90114-7

Egel, R. (1979). The lac-operon for lactose degradation, or rather for the utilization of galactosylglycerols from galactolipids? *J. Theor. Biol.* 79, 117–119. doi: 10.1016/0022-5193(79)90260-1

Egel, R. (1988). The lac operon - an irrelevant paradox? *Trends Genet.* 4:9525.

Fabich, A. J., Jones, S. A., Chowdhury, F. Z., Cernosek, A., Anderson, A., Smalley, D., et al. (2008). Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. *Infect. Immun.* 76, 1143–1152. doi: 10.1128/IAI.01386-07

Frazão, N., Sousa, A., Lässig, M., and Gordo, I. (2019). Horizontal gene transfer overrides mutation in *Escherichia coli* colonizing the mammalian gut. *Proc. Natl. Acad. Sci. U.S.A.* 116, 17906–17915. doi: 10.1073/pnas.1906958116

Ghalayini, M., Magnan, M., Dion, S., Zatout, O., Bourguignon, L., Tenaillon, O., et al. (2019). Long-term evolution of the natural isolate of *Escherichia coli* 536 in the mouse gut colonized after maternal transmission reveals convergence in the constitutive expression of the lactose operon. *Mol. Ecol.* 28, 4470–4485. doi: 10.1111/mec.15232

Görs, S., Kucia, M., Langhammer, M., Junghans, P., and Metges, C. C. (2009). Technical note: milk composition in mice - Methodological aspects and effects of mouse strain and lactation day. *J. Dairy Sci.* 92, 632–637. doi: 10.3168/jds.2008-1563

Hartl, D. L., and Dykhuizen, D. E. (1984). The population genetics of *Escherichia coli*. *Annu. Rev. Genet.* 18, 31–68. doi: 10.1146/annurev.ge.18.120184.000335

Holsinger, V. H. (1988). "Lactose," in *Fundamentals of Dairy Chemistry*, eds N. P. Wong, R. Jenness, M. Keeney, and E. H. Marth (New York, NY: Van Nostrand Reinhold Company Inc), 583–607. doi: 10.1007/978-1-4615-7050-9_11

Jacob, F., and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318–356. doi: 10.1016/S0022-2836(61)80072-7

Leónidas Cardoso, L., Durão, P., Amicone, M., and Gordo, I. (2020). Dysbiosis individualizes the fitness effect of antibiotic resistance in the mammalian gut. *Nat. Ecol. Evol.* 4, 1268–1278. doi: 10.1038/s41559-020-1235-1

Li, H., Limenitakis, J. P., Fuhrer, T., Geuking, M. B., Lawson, M. A., Wyss, M., et al. (2015). The outer mucus layer hosts a distinct intestinal microbial niche. *Nat. Commun.* 6:8292. doi: 10.1038/ncomms9292

Mackie, R. I., Sghir, A., and Gaskins, H. R. (1999). Developmental microbial ecology of the neonatal gastrointestinal tract. *Am. J. Clin. Nutr.* 69, 1035S–1045S. doi: 10.1093/ajcn/69.5.1035s

McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e0061217. doi: 10.1371/journal.pone.0061217

Milani, C., Duranti, S., Bottacini, F., Casey, E., Turroni, F., Mahony, J., et al. (2017). The first microbial colonizers of the human gut: composition, activities, and health implications of the infant gut microbiota. *Microbiol. Mol. Biol. Rev.* 81, 1–67. doi: 10.1128/mmbr.00036-17

Müller-Hill, B. (1996a). "Activation of the lac promoter by CAP protein," in *The Lac Operon: A Short History of a Genetic Paradigm*, ed. Berlin (New York, NY: Walter de Gruyter), 137–141.

Müller-Hill, B. (1996b). "ß-Galactosidase, lac permease and transacetylase," in *The Lac Operon: A Short History of a Genetic Paradigm*, ed. Berlin (New York, NY: Walter de Gruyter), 181.

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. doi: 10.1038/35012500

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Core Team.

Raman, R., Thomas, R. G., Weiner, M. W., Jack, C. R., Ernstrom, K., Aisen, P. S., et al. (2005). Diversity of the human intestinal microbial Flora. *Science* 308, 1635–1638.

Rang, C. U., Licht, T. R., Midtvedt, T., Conway, P. L., Chao, L., Krogfelt, K. A., et al. (1999). Estimation of growth rates of *Escherichia coli* BJ4 in streptomycin-treated and previously germfree mice by in situ rRNA hybridization. *Clin. Diagn. Lab. Immunol.* 6, 434–436. doi: 10.1128/cdli.6.3.434-436.1999

Roderick, S. L. (2005). The lac operon galactoside acetyltransferase. *Compt. Ren. Biol.* 328, 568–575. doi: 10.1016/j.crvi.2005.03.005

Sousa, A., Ramiro, R. S., Barroso-Batista, J., Güleresi, D., Lourenčo, M., and Gordo, I. (2017). Recurrent reverse evolution maintains polymorphism after strong bottlenecks in commensal gut bacteria. *Mol. Biol. Evol.* 34, 2879–2892. doi: 10.1093/molbev/msx221

Stoebel, D. M. (2005). Lack of evidence for horizontal transfer of the lac operon into *Escherichia coli*. *Mol. Biol. Evol.* 22, 683–690. doi: 10.1093/molbev/msi056

Storhaug, C. L., Fosse, S. K., and Fadnes, L. T. (2017). Country, regional, and global estimates for lactose malabsorption in adults: a systematic review and meta-analysis. *Lancet Gastroenterol. Hepatol.* 2, 738–746. doi: 10.1016/S2468-1253(17)30154-1

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 30 July 2021
doi: 10.3389/fmicb.2021.699190

Interspecific Niche Competition Increases Morphological Diversity in Multi-Species Microbial Communities

Xiao-Lin Chu[1,2], Quan-Guo Zhang[2], Angus Buckling[1] and Meaghan Castledine[1]**

[1]College of Life and Environmental Sciences, Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, United Kingdom, [2]State Key Laboratory of Earth Surface Processes and Resource Ecology and MOE Key Laboratory for Biodiversity Science and Ecological Engineering, Beijing Normal University, Beijing, China

Intraspecific competition for limited niches has been recognized as a driving force for adaptive radiation, but results for the role of interspecific competition have been mixed. Here, we report the adaptive diversification of the model bacteria *Pseudomonas fluorescens* in the presence of different numbers and combinations of four competing bacterial species. Increasing the diversity of competitive community increased the morphological diversity of focal species, which is caused by impeding the domination of a single morphotype. Specifically, this pattern was driven by more diverse communities being more likely to contain key species that occupy the same niche as otherwise competitively superior morphotype, and thus preventing competitive exclusion within the focal species. Our results suggest that sympatric adaptive radiation is driven by the presence or absence of niche-specific competitors.

Keywords: competition, niche, community ecology, adaptive radiation, microbial ecology, evolutionary ecology

OPEN ACCESS

Edited by:
Alejandro Couce, Polytechnic University of Madrid, Spain

Reviewed by:
Michael Travisano, University of Minnesota Twin Cities, United States
Despoina Mavridou, University of Texas at Austin, United States
David S. Guttman, University of Toronto, Canada

****Correspondence:***
Xiao-Lin Chu
chuxl@mail.bnu.edu.cn
Meaghan Castledine
mcastledine96@gmail.com

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *23 April 2021*
Accepted: *02 July 2021*
Published: *30 July 2021*

Citation:
Chu X-L, Zhang Q-G, Buckling A and Castledine M (2021) Interspecific Niche Competition Increases Morphological Diversity in Multi-Species Microbial Communities. Front. Microbiol. 12:699190. doi: 10.3389/fmicb.2021.699190

INTRODUCTION

Adaptive radiation is a key component of biodiversity generation (Dieckmann and Doebeli, 1999; Schluter, 2000; Losos, 2010; Kassen, 2014). Population diversification – the precursor to adaptive radiation – not only depends on genetic characteristics of a population, such as cryptic genetic variation (Zheng et al., 2019) and mutation/recombination rate (Gavrilets and Vose, 2005; Lanfear et al., 2010), but also on abiotic and biotic environmental conditions (Grant, 1986; Rainey and Travisano, 1998; Buckling and Rainey, 2002; Hall and Colegrave, 2007; Meyer and Kassen, 2007; Betts et al., 2018). Intraspecific competition, in particular, is suggested to be a key driver of diversifying selection in a wide range of taxa (Schluter, 2000; MacLean et al., 2005; Svanback and Bolnick, 2007). However, the role of interspecific competition is less clear. Interspecific competitors might contribute to increased diversification by creating new ecological niches or could select for different resource usage to that of coexisting species (Emerson and Kolm, 2005; Svanback and Bolnick, 2007; Erwin, 2008; Calcagno et al., 2017). On the other hand, interspecific competitors may inhibit diversification due to fewer vacant niches or reduced population sizes (Gómez and Buckling, 2013; Ghoul and Mitri, 2016; Schluter and Pennell, 2017; Pontarp and Petchey, 2018; Harvey et al., 2020). Moreover, the neutral theory of community assembly proposes that all species are functionally equivalent for community

assembly and maintenance (Hubbell, 2001; Harris et al., 2017), implying a neutral role of interspecific competition in diversification.

Direct experimental tests of the role of interspecific competitors on diversification have typically used the bacterium *Pseudomonas fluorescens* as the focal species because of its propensity to morphologically diversify over 10s of generations. These studies have generated contrasting results. Gómez and Buckling (2013) demonstrated lower evolved diversity in resource use of an initially isogenic population of *P. fluorescens* in the presence vs. absence of the natural microbial community in soil; Jousset et al. (2016) reported that increasing the number of competing *P. fluorescens* isolates resulted in an increased diversifying selection of the focal strain; while Zhang et al. (2012) reported little effect of a competitor (*Pseudomonas putida*) on *P. fluorescens* diversification.

We argue that the inconsistent results could be reconciled based on the theoretical framework emerging from biodiversity-ecosystem functioning studies (Schulze and Mooney, 1993; Tilman, 1997; Loreau et al., 2001). Specifically, more diverse competitors may generally reduce the diversification of focal species due to the reduced niche availability and total abundance of a focal species, which is akin to a complementarity effect. Meanwhile, a sampling effect may also take place, as a species-richer competitor community is more likely to contain particular species that can strongly affect the focal species. However, its effect on the diversification within the latter may depend on the specific competitive interactions, either positive (if it competes with a dominant genotype and thus prevent competitive exclusion within the focal species) or negative (if it competes with all genotypes and strongly reduces the total abundance of the focal species). While the Jousset study created a diversity gradient, all the competitors were *P. fluorescens* isolates, limiting the potential for resident niche occupation (Jousset et al., 2016). A recent study of natural microbial communities from the Earth Microbiome Project found a unimodal relationship between diversity and diversification: community diversity beget diversity in low-diversity biomes and reach plateaus when niches are increasingly filled in high-diversity biomes (Madi et al., 2020).

To further investigate the role of interspecific diversity on diversification, we evolved *P. fluorescens* SBW25 across a diversity gradient of a synthetic microbial community isolated from potting compost, which consists of four species that can stably coexist in a relatively oligotrophic medium when incubated in the lab (Castledine et al., 2020). *Pseudomonas fluorescens* can rapidly diversify into three main colony morphotypes when propagated in spatially heterogeneous microcosms (Rainey and Travisano, 1998). These morphotypes can be typically identified as: ancestor-like smooth (SM) that occupies the liquid phase, wrinkly spreader (WS), which arises by spontaneous mutations and forms a self-supporting mat on the air-liquid interface, and fuzzy spreader (FS), which occupies the anaerobic niche (Rainey and Travisano, 1998). The driver of diversification has been shown to be competition for oxygen and other nutrients (Koza et al., 2011; Kassen, 2014). The present study measured the morphological diversification of *P. fluorescens* in the presence of microbial communities, in which a diversity gradient was set-up using four species, to determine the diversity of competing microbial communities on adaptive diversification.

Abbreviations: A, *Achromobacter* sp.; O, *Ochrobactrum* sp.; S, *Stenotrophomonas* sp.; V, *Variovorax* sp.; SM, Smooth morph; WS, Wrinkly spreaders; FS, Fuzzy spreader.

MATERIALS AND METHODS

Experimental Evolution

Communities were set up with a SM clone of *P. fluorescens* SBW25 with isogenic clones of *Achromobacter* sp. (A), *Ochrobactrum* sp. (O), *Stenotrophomonas* sp. (S), and *Variovorax* sp. (V), which can be distinguished by their unique colony morphologies (**Supplementary Figure 1A**; Castledine et al., 2020; Padfield et al., 2020 unpublished). Every community combination of the four species with *P. fluorescens* was set up across all the levels of diversity with six replicates for each community combination. Therefore, a total of 16 different communities were established with total species richness ranging from 1 (only *P. fluorescens*) to 5 (all five species). Specifically, there were one 1-species community treatment, four 2-species treatments, six 3-species treatments, four 4-species treatments, and one 5-species treatment (**Figure 1**). Communities were grown in 25 ml glass vials with loosened lids with 6 ml of M9KB media (glycerol 10 g L^{-1}, proteose peptone no.3 20 g L^{-1}, KH_2PO_4 3 g L^{-1}, NaCl 0.5 g L^{-1}, and NH_4Cl 1 g L^{-1}). Prior to experimental set-up, each species was grown for 2 days in M9KB media at 28°C to achieve high-cell densities. Cell densities of each species were diluted into M9 buffer to approximately 10^4 CFUs uL^{-1}. Each microcosm was inoculated with 20 uL per species (~10^5 cells) and incubated at 28°C in static, after which 60uL was transferred every 7 days for a total of three transfers. Culture samples were cryogenically frozen at −80°C in 50% glycerol (final concentration: 25%) at each transfer. Species densities within each microcosm at each transfer were determined by plating culture dilutions onto KB agar and counting the number of colony-forming units (CFUs) after 2 days of incubation at 28°C.

Measurement of Diversity

Community diversity was estimated by species richness (the number of species). The ancestral smooth *P. fluorescens* diversified into three morphotypes (SM, WS, and FS; **Supplementary Figure 1B**) at the end point and the sympatric diversity of *P. fluorescens* populations was obtained by measuring the morphologies of 100 randomly chosen colonies (Buckling et al., 2000) and calculated as Simpson's index: $\left(\frac{N}{N-1}\right)(1-\lambda)$, where N is the total number of colonies sampled from the focal population and $\lambda = 1 - \sum {p_i}^2$ (p_i is the frequency of the ith morphotype; Simpson, 1949).

Statistical Analysis

All analysis was conducted using R (version 3.5.2; R Core Team, 2018), and all plots were made using the R package "*ggplot2*" (Wickham, 2016). Simpson's index was calculated using the package "*vegan*" (Oksanen et al., 2019). A linear

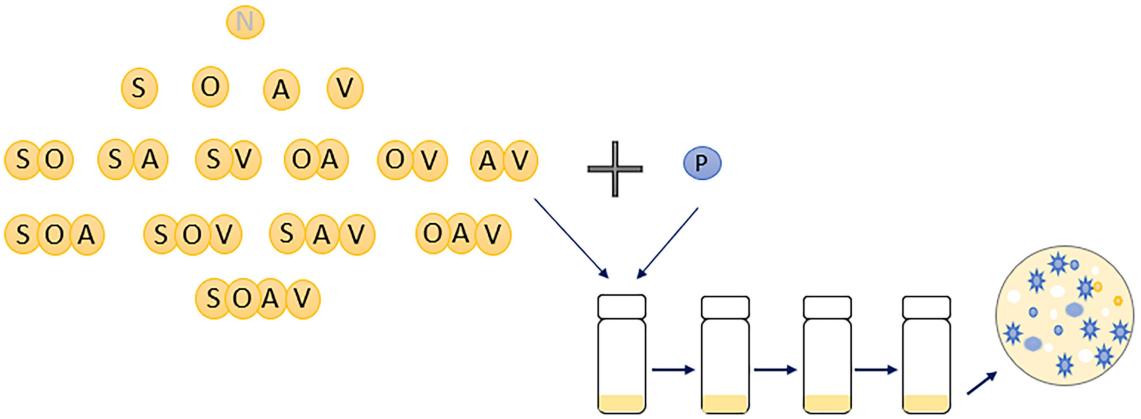


FIGURE 1 | An illustration of experimental design of the present study. Sixteen different microbial communities were set-up with species diversity ranging from 1 (monoculture of *Pseudomonas fluorescens*) to 5 (with all five species). Six replicates were conducted for each community combination. N represents no other species were inoculated into the microcosms; P, A, O, S, and V represents *Pseudomonas fluorescens*, *Achromobacter*, *Ochrobactrum*, *Stenotrophomonas* and *Variovorax*, respectively. Diversity of the focal species, *Pseudomonas fluorescens*, at the end of the evolution experiment was estimated based on its morphological diversity of 100 randomly chosen colonies (blue ones).

model was used to analyze the relationship between *P. fluorescens* diversity and the proportion of WS morphotype; the relationship between inoculated density of the competing community and proportion of WS or *P. fluorescens* diversity; and the relationship between the initial or final density of *P. fluorescens* and its sympatric diversity. ANOVA was performed to analyze whether the presence or absence of competitors affected *P. fluorescens* diversity and the proportion of WS. Density data were log-transformed [$\log_{10}$ (1 + CFUs ml^{-1})].

The effect of community richness on the diversity of *P. fluorescens* were analyzed by an analysis of variance with sequential sum of squares (type I) using linear models (Schmid et al., 2002; Buzhdygan et al., 2020). To assess whether diversity effects were only due to the presence of single species, we fitted each species (presence/absence) before species richness in separate sequential analyses. If fitting single genotypes before species richness removed the effect of richness indices, the observed diversity effects may have been mainly caused by a sampling effect (the inclusion of a particular genotype in the community; Huston, 1997). To further explain whether species composition matters, we used ANOVA to test the effect of the presence of each species and their interactions on *P. fluorescens* diversity. The simplest model was obtained by AICc ranking using the R package "*MuMIn*" (Barton, 2020). The effects of the presence of single species and initial community diversity on the proportion of WS, and on the density of WS were tested with sequential linear models as described above.

RESULTS

Dominant WS Morphotype Inhibits *Pseudomonas fluorescens* Diversity

Across all evolution lines, the WS morphotype was found to dominate the diversified *P. fluorescens* populations (mean ± SD of the proportion of WS in *P. fluorescens* populations: 0.847 ± 0.128). A negative relationship was found between *P. fluorescens* diversity and proportion of WS ($F_{1,94}$ = 369.82, p < 0.001; **Figure 2**), indicating that *P. fluorescens* diversity is caused by reduced proportions of WS morphotypes. This implies that factors limiting the dominant WS morphotype would increase the sympatric diversity of *P. fluorescens*. Although the presence of competing communities was not found to affect *P. fluorescens* diversity ($F_{1,94}$ = 3.606, p = 0.061) or the proportion of WS ($F_{1,94}$ = 2.521, p = 0.116), the density of inoculated competing communities reduced the proportion of WS ($F_{1,88}$ = 5.346, p = 0.023; **Supplementary Figure 2**). In addition, a positive relationship was observed between the density of inoculated competing communities and the diversity of focal species ($F_{1,88}$ = 10.169, p = 0.002; **Supplementary Figure 2**), suggesting a competing effect of the microbial communities on *P. fluorescens*.

Species-Specific Effects on *Pseudomonas fluorescens* Morphotype Diversity

Pseudomonas fluorescens diversity was positively affected by increasing species richness ($F_{1,94}$ = 10.204, p = 0.002). However, the significant effect of species richness was removed when fitted after the presence/absence of the *Ochrobactrum* species (**Figures 3A,B**; **Table 1**), indicating that the observed diversity effect is mainly due to a sampling effect. Further analysis revealed that the diversity of *P. fluorescens* was also affected by the interaction between *Ochrobactrum* and *Achromobacter* species ($F_{1,92}$ = 9.537, p = 0.003). Qualitatively similar results were obtained using alternative community diversity metric (Simpson's index; see **Supplementary Text**).

While species *Achromobacter* (A), *Ochrobactrum* (O), *Stenotrophomonas* (S), and *Variovorax* (V) stably coexisted in a previous study, the presence of *P. fluorescens* destabilized this community in many combinations (**Supplementary Figure 3**). V was driven extinct in at least 3/6 replicates in 7/8 species

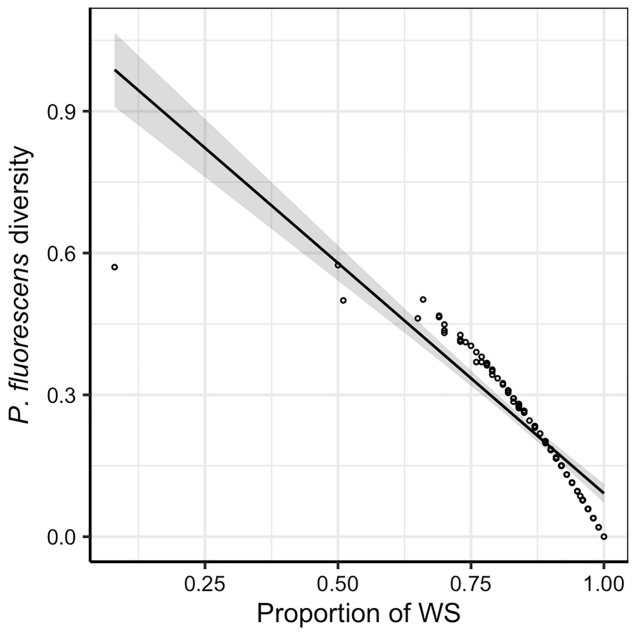


FIGURE 2 | *Pseudomonas fluorescens* diversity is negatively correlated with the proportion of wrinkly spreader (WS). The regression line represents a significant linear relationship and shaded areas around lines show the 95% confidence intervals: $y = 1.066 - 0.974x$, $F_{1,94} = 369.820$, $p < 0.001$, adjusted $R^2 = 0.795$.

combinations with V going extinct in 2/6 replicates in the PSAV evolution line. Similarly, S went extinct in 4/6 replicates in evolution lines PSO and PSOA, and in 1/6 replicates in evolution lines PSOV and PSV. These extinctions resulted in variable community diversities and the *P. fluorescens* final densities ranging from 4.40×10^8 to 2.78×10^9. These results indicate that O and A (to a less extent) are the main competitors of *P. fluorescens* and affect its diversity.

Potential Mechanisms

We first tested if the patterns of diversification could be affected by *P. fluorescens* density by affecting the supply of mutations and the strength of diversifying selection. Neither inoculated *P. fluorescens* density nor its final density affected diversification. However, the final population size of *P. fluorescens* decreased in the presence of competing species ($F_{1,94} = 13.410$, $p < 0.001$), suggesting the interspecific resource competition during the diversification of focus species. As A and O coexisted with *P. fluorescens*, while S and V went extinct in many combinations (**Supplementary Figure 3**), they are presumably the main competitors of *P. fluorescens*.

Given that *P. fluorescens* diversity is caused by the reduced proportion of WS, the increased diversity of focal species in the presence of O, and O and A together may be a result of competition with WS. Therefore, we tested whether the presence of particular species affected diversity of focal species by affecting the proportion of WS (the sampling effect). Although the proportion of WS was affected by species richness ($F_{1,94} = 8.061$, $p = 0.006$), its effect was removed when fitted after the presence of O (**Figures 3C,D; Table 2**). In addition, a marginal sampling effect of A was also detected (**Table 2**). These results supported the competition between WS and O (and to a lesser extent A) in driving the diversification of *P. fluorescens*. As WS forms mats to occupy the air-liquid interface, we hypothesized that O (and A to a lesser extent) may compete with WS in this niche, reducing invasion of WS. O, A, and WS were found to form mats in media when grown in monoculture while V formed very thin mats and mats were absent for S and SM. The presence of O and A were, indeed, found to affect the density of WS (**Supplementary Text; Supplementary Figure 4**; and **Supplementary Table 1**), suggesting the niche competition of O and A with the derived WS morphs of *P. fluorescens*.

DISCUSSION

Here, we demonstrated that the colony morphological diversity of evolving *P. fluorescens* populations increased with the number of bacterial taxa it was co-cultured with. This pattern was primarily driven by the presence of a specific species (O) in more diverse communities, although there were also some interactive effects with other species. *Pseudomonas fluorescens* diversity was negatively correlated with the proportion of evolved WS morphotypes, which had a mean frequency of ~0.8 across all replicates. Most of the remaining *P. fluorescens* populations were SM morphotypes. Diversity was therefore maximized when the evolved WS less successfully dominating populations.

The presence of O (and to a lesser extent A) impeded the dominance of WS. The other species (S and V) were frequently driven extinct by the end of experiment, and hence would have imposed relatively little competition. We also observed that O and A produced more detectable mats (or biofilms) than the other two species, and mat formation is characteristic of WS [we tried to quantify the biofilms with the resazurin assay, (Peeters et al., 2008) but failed as their mats are not as definable as the biofilms of *P. fluorescens*]. This suggests that O and A, which are the strongest competitors in the community, are competing more with WS than SM by occupying a similar ecological niche to WS (Kim et al., 2014). Similar constraints on the invasion of evolved WS have been observed when the resident *P. fluorescens* WS genotypes were present in the populations (Brockhurst et al., 2007; Flohr et al., 2013).

The formation of biofilms and inhibiting the biofilm formation of other microbial populations could be an evolved competitive strategy since these genotypes could gain preferential access to resources and the biofilms could protect cells from environmental hazards (An et al., 2006; Stacy et al., 2016). For example, it has been reported that *P. aeruginosa* could prevent *Agrobacterium* from forming its own biofilm by surface blanketing (An et al., 2006), and *E. coli* can reduce the biofilm formation of Gram-positive and Gram-negative bacteria by producing extracellular polysaccharides or surfactants (Valle et al., 2006; Rendueles et al., 2011).

WS density was highly variable in the presence of O and A, and this may be explained by the advantage of forming biofilms and historical contingences. The diversified WS is a

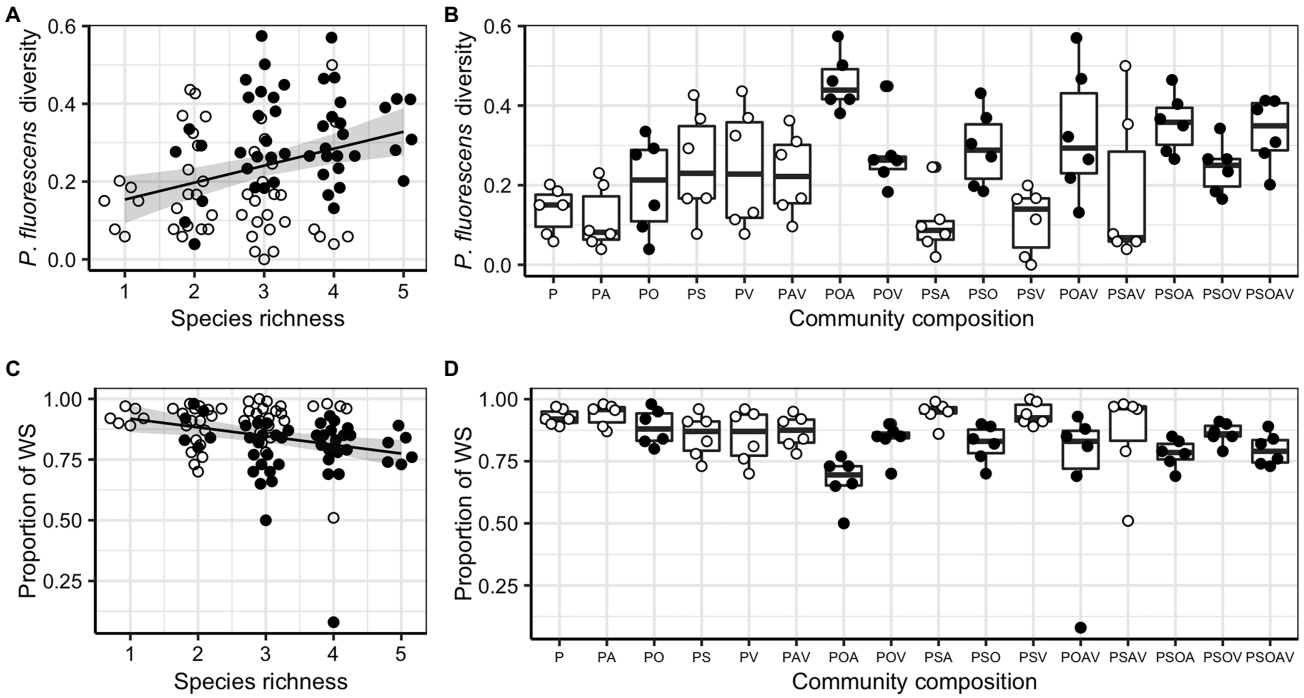


FIGURE 3 | The effect of species richness, the presence of *Ochrobactrum* sp. (O) and community composition on *Pseudomonas fluorescens* diversity **(A,B)** and the proportion of wrinkly spreader **(**WS; **C,D)**. For both responsive variables, regression lines showed significant linear relationships with species richness: $y = 0.110 + 0.044x$, $F_{1,94} = 369.820$, $p < 0.001$, adjusted $R^2 = 0.088$ **(A)**; and $y = 0.954 - 0.036x$, $F_{1,94} = 8.061$, $p = 0.006$, adjusted $R^2 = 0.069$ **(B)**. Shaded areas around lines show the 95% confidence intervals. Treatments containing O (filled dots) and not containing O (open dots) are highlighted to show the sampling effect of O. Tops and bottoms of the bars represent the 75th and 25th percentiles of the data, the middle lines are the medians, and the whiskers extend from their respective hinge to the smallest or largest value no further than 1.5 × interquartile range.

TABLE 1 | Analysis of variance table of *F*-values on the effects of presence of specific species, and species richness on the diversity of *Pseudomonas fluorescens* populations.

Factor	*df*	***Achromobacter* sp.**		***Ochrobactrum* sp.**		***Stenotrophomonas* sp.**		***Variovorax* sp.**	
		F	***P***	***F***	***P***	***F***	***P***	***F***	***P***
Species presence	1	2.558	0.113	31.898	**<0.001**	0.293	0.590	0.027	0.871
Species richness	1	7.538	**0.007**	0.620	0.433	16.849	**<0.001**	13.174	**<0.001**
Residuals	93								

The bold values indicate significant effects.

TABLE 2 | Analysis of variance table of *F*-values on the effects of presence of specific species, and species richness on the proportion of wrinkly spreader.

Factor	*df*	***Achromobacter* sp.**		***Ochrobactrum* sp.**		***Stenotrophomonas* sp.**		***Variovorax* sp.**	
		F	***P***	***F***	***P***	***F***	***P***	***F***	***P***
Species presence	1	3.868	0.052	18.257	**<0.001**	0.807	0.371	0.249	0.619
Species richness	1	4.548	**0.036**	0.942	0.334	15.272	**<0.001**	8.955	**0.004**
Residuals	93								

The bold values indicate significant effects.

cooperating group that can enable individuals to align with them by overproducing an adhesive polymer and promote the colonization of the air-liquid interface though the overproduction is costly to individuals (Spiers et al., 2002; Rainey and Rainey, 2003). Therefore, if mutations to form WS arose late in a population or *P. fluorescens* initially grew slowly, WS may lose the chance to spread because the available niche is already partially occupied by competitor species.

It has been previously shown that the *Pseudomonas* sp. can stably coexist with the other four competing species (Castledine et al., 2020), while S and V were at low densities or went extinct in many combinations by the end of our experiment.

Although a different *Pseudomonas* was involved here, it is not suggested to be a main cause of the instability as family-level bacteria are found to have similar function (Goldford et al., 2018; Louca et al., 2018). Therefore, the instability may be a result of different media used in the present studies that altered community composition (Goldberg and Miller, 1990; Goldford et al., 2018). Specifically, the rich media used here may have increased the likelihood of competitive exclusion by favoring the fast-growing *Pseudomonas*.

Limitations of the present study are noteworthy. The relative short-term evolution limited the study to further investigate more nuanced processes such as how within-species diversification feedback to affect community structure. Though the three morphotypes of *P. fluorescens* has been widely used for diversity estimation, less consideration has been given to its genetic diversity (Spiers et al., 2002; Fukami et al., 2007).

Diversity-dependent adaptive radiation theory predicts that more diverse communities are more likely to experience evolutionary diversification (Calcagno et al., 2017). This is not supported by our study despite a greater diversity of focal species in the presence of more diverse competitors was observed. Our study is instead consistent with previous work, suggesting that evolutionary processes are impeded by the presence of competitors (Hall et al., 2018; Scheuerl et al., 2020). In this case, competitors reduced the dominance of the evolved WS genotypes, most likely because of shared niche occupation in spatially structured (static) microcosms. Though the reduced population size of focal species by interspecific competition may affect the adaptive radiation process, *P. fluorescens* is less affected in our working system as the focal species showed an advantage and dominated the microbial communities after three transfers (Johansson, 2008; Osmond and de Mazancourt, 2012; Zhao et al., 2018). Our results may help to understand the inconsistent effects of interspecific competitors on diversification of resident species across different studies and suggest that the presence of niche-specific competitors is a possible explanation.

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://doi.org/10.6084/m9.figshare.14471229.

AUTHOR CONTRIBUTIONS

All authors contributed to the design of the study. X-LC and MC conducted the experiments. All authors contributed to the data analysis and the writing of the manuscript.

FUNDING

The work was funded by NERC (NE/P001130/1). X-LC was funded by the China Scholarship Council. MC was funded by a studentship from the MRC (111846). Q-GZ was supported by the National Natural Science Foundation of China (31725006 and 31670376), and the 111 project (B13008).

ACKNOWLEDGMENTS

We thank Jesica Soria-Pascual and Elze Hesse for help in the laboratory and Daniel Padfield for helpful advice on statistical analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articless/10.3389/fmicb.2021.699190/full#supplementary-material

REFERENCES

An, D., Danhorn, T., Fuqua, C., and Parsek, M. R. (2006). Quorum sensing and motility mediate interactions between *Pseudomonas aeruginosa* and *Agrobacterium tumefaciens* in biofilm cocultures. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3828–3833. doi: 10.1073/pnas.0511323103

Barton, K. (2020). MuMIn: multi-model inference. Available at: https://cran.r-project.org/web/packages/MuMIn/index.html (Accessed April 15, 2020).

Betts, A., Gray, C., Zelek, M., MacLean, R. C., and King, K. C. (2018). High parasite diversity accelerates host adaptation and diversification. *Science* 360, 907–911. doi: 10.1126/science.aam9974

Brockhurst, M. A., Colegrave, N., Hodgson, D. J., and Buckling, A. (2007). Niche occupation limits adaptive radiation in experimental microcosms. *PLoS One* 2:e193. doi: 10.1371/journal.pone.0000193

Buckling, A., Kassen, R., Bell, G., and Rainey, P. B. (2000). Disturbance and diversity in experimental microcosms. *Nature* 408, 961–964. doi: 10.1038/35050080

Buckling, A., and Rainey, P. B. (2002). The role of parasites in sympatric and allopatric host diversification. *Nature* 420, 496–499. doi: 10.1038/nature01164

Buzhdygan, O. Y., Meyer, S. T., Weisser, W. W., Eisenhauer, N., Ebeling, A., Borrett, S. R., et al. (2020). Biodiversity increases multitrophic energy use efficiency, flow and storage in grasslands. *Nat. Ecol. Evol.* 4, 393–405. doi: 10.1038/s41559-020-1123-8

Calcagno, V., Jarne, P., Loreau, M., Mouquet, N., and David, P. (2017). Diversity spurs diversification in ecological communities. *Nat. Commun.* 8:15810. doi: 10.1038/ncomms15810

Castledine, M., Padfield, D., and Buckling, A. (2020). Experimental (co)evolution in a multi-species microbial community results in local maladaptation. *Ecol. Lett.* 23, 1673–1681. doi: 10.1111/ele.13599

Dieckmann, U., and Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature* 400, 354–357. doi: 10.1038/22521

Emerson, B. C., and Kolm, N. (2005). Species diversity can drive speciation. *Nature* 434, 1015–1017. doi: 10.1038/nature03450

Erwin, D. H. (2008). Macroevolution of ecosystem engineering, niche construction and diversity. *Trends Ecol. Evol.* 23, 304–310. doi: 10.1016/j.tree.2008.01.013

Flohr, R. C. E., Blom, C. J., Rainey, P. B., and Beaumont, H. J. E. (2013). Founder niche constrains evolutionary adaptive radiation. *Proc. Natl. Acad. Sci. U. S. A.* 110, 20663–20668. doi: 10.1073/pnas.1310310110

Fukami, T., Beaumont, H. J. E., Zhang, X. X., and Rainey, P. B. (2007). Immigration history controls diversification in experimental adaptive radiation. *Nature* 446, 436–439. doi: 10.1038/nature05629

Gavrilets, S., and Vose, A. (2005). Dynamic patterns of adaptive radiation. *Proc. Natl. Acad. Sci. U. S. A.* 102, 18040–18045. doi: 10.1073/pnas.0506330102

Ghoul, M., and Mitri, S. (2016). The ecology and evolution of microbial competition. *Trends Microbiol.* 24, 833–845. doi: 10.1016/j.tim.2016.06.011

Goldberg, D. E., and Miller, T. E. (1990). Effects of different resource additions of species diversity in an annual plant community. *Ecology* 71, 213–225. doi: 10.2307/1940261

Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., et al. (2018). Emergent simplicity in microbial community assembly. *Science* 361, 469–474. doi: 10.1126/science.aat1168

Gómez, P., and Buckling, A. (2013). Real-time microbial adaptive diversification in soil. *Ecol. Lett.* 16, 650–655. doi: 10.1111/ele.12093

Grant, P. R. (1986). *Ecology and Evolution of Darwin's Finches*. Princeton, New Jersey: Princeton University Press.

Hall, A. R., and Colegrave, N. (2007). How does resource supply affect evolutionary diversification? *Proc. R. Soc. B* 274, 73–78. doi: 10.1098/rspb.2006.3703

Hall, J. P. J., Harrison, E., and Brockhurst, M. A. (2018). Competitive species interactions constrain abiotic adaptation in a bacterial soil community. *Evol. Lett.* 2, 580–589. doi: 10.1002/evl3.83

Harris, K., Parsons, T. L., Ijaz, U. Z., Lahti, L., Holmes, I., and Quince, C. (2017). Linking statistical and ecological theory: Hubbell's unified neutral theory of biodiversity as a hierarchical dirichlet process. *Proc. IEEE* 105, 516–529. doi: 10.1109/JPROC.2015.2428213

Harvey, M. G., Bravo, G. A., Claramunt, S., Cuervo, A. M., Derryberry, G. E., Battilana, J., et al. (2020). The evolution of a tropical biodiversity hotspot. *Science* 370, 1343–1348. doi: 10.1126/science.aaz6970

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, New Jersey: Princeton University Press.

Huston, M. A. (1997). Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia* 110, 449–460. doi: 10.1007/s004420050180

Johansson, J. (2008). Evolutionary responses to environmental changes: how does competition affect adaptation? *Evolution* 62, 421–435. doi: 10.1111/j.1558-5646.2007.00301.x

Jousset, A., Eisenhauer, N., Merker, M., Mouquet, N., and Scheu, S. (2016). High functional diversity stimulates diversification in experimental microbial communities. *Sci. Adv.* 2:e1600124. doi: 10.1126/sciadv.1600124

Kassen, R. (2014). *Experimental Evolution and the Nature of Biodiversity*. Colorado: Roberts and Company Publisher.

Kim, W., Racimo, F., Schluter, J., Levy, S. B., and Foster, K. R. (2014). Importance of positioning for microbial evolution. *Proc. Natl. Acad. Sci. U. S. A.* 111, E1639–E1647. doi: 10.1073/pnas.1323632111

Koza, A., Moshynets, O., Otten, W., and Spiers, A. J. (2011). Environmental modification and niche construction: developing O_2 gradients drive the evolution of the wrinkly spreader. *ISME J.* 5, 665–673. doi: 10.1038/ismej.2010.156

Lanfear, R., Ho, S. Y. W., Love, D., and Bromham, L. (2010). Mutation rate is linked to diversification in birds. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20423–20428. doi: 10.1073/pnas.1007888107

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J. P., Hector, A., et al. (2001). Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* 294, 804–808. doi: 10.1126/science.1064088

Losos, J. B. (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* 175, 623–639. doi: 10.1086/652433

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'Connor, M. I., et al. (2018). Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* 2, 936–943. doi: 10.1038/s41559-018-0519-1

MacLean, R. C., Dickson, A., and Bell, G. (2005). Resource competition and adaptive radiation in a microbial microcosm. *Ecol. Lett.* 8, 38–46. doi: 10.1111/j.1461-0248.2004.00689.x

Madi, N., Vos, M., Murall, C. L., Legendre, P., and Shapiro, B. J. (2020). Does diversity beget diversity in microbiomes? *elife* 9:e58999. doi: 10.7554/eLife.58999

Meyer, J. R., and Kassen, R. (2007). The effects of competition and predation on diversification in a model adaptive radiation. *Nature* 446, 432–435. doi: 10.1038/nature05599

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Minchin, P. R., et al. (2019). vagan: community ecology package. Available at: https://cran.r-project.org/package=vegan (Accessed November 28, 2020).

Osmond, M. M., and de Mazancourt, C. (2012). How competition affects evolutionary rescue. *Philos. Trans. R. Soc. B Biol. Sci.* 368:20120085. doi: 10.1098/rstb.2012.0085

Peeters, E., Nelis, H. J., and Coenye, T. (2008). Comparison of multiple methods for quantification of microbial biofilms grown in microtiter plates. *J. Microbiol. Methods* 72, 157–165. doi: 10.1016/j.mimet.2007.11.010

Pontarp, M., and Petchey, O. L. (2018). Ecological opportunity and predator-prey interactions: linking eco-evolutionary processes and diversification in adaptive radiations. *Proc. R. Soc. B Biol. Sci.* 285:20172550. doi: 10.1098/rspb.2017.2550

R Core Team (2018). R: A language and environment for statistical computing. Available at: https://www.r-project.org/ (Accessed December 20, 2020).

Rainey, P. B., and Rainey, K. (2003). Evolution of cooperation and conflict in experimental bacterial populations. *Nature* 425, 72–74. doi: 10.1038/nature01906

Rainey, P. B., and Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature* 394, 69–72. doi: 10.1038/27900

Rendueles, O., Travier, L., Latour-Lambert, P., Fontaine, T., Magnus, J., Denamur, E., et al. (2011). Screening of *Escherichia coli* species biodiversity reveals new biofilm-associated antiadhesion polysaccharides. *Mbio* 2, e00043–e00011. doi: 10.1128/mBio.00043-11

Scheuerl, T., Hopkins, M., Nowell, R. W., Rivett, D. W., Barraclough, T. G., and Bell, T. (2020). Bacterial adaptation is constrained in complex communities. *Nat. Commun.* 11:754. doi: 10.1038/s41467-020-14570-z

Schluter, D. (2000). *The Ecology of Adaptive Radiations*. New York: Oxford University Press.

Schluter, D., and Pennell, M. W. (2017). Speciation gradients and the distribution of biodiversity. *Nature* 546, 48–55. doi: 10.1038/nature22897

Schmid, B., Hector, A., Huston, M. A., Inchasusti, P., Nijs, I., and Leadley, P. W. (2002). "The design and analysis of biodiversity experiments," in *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*. eds. M. Loreau, S. Naeem and P. Inchausti (New York: Oxford University Press).

Schulze, E. D., and Mooney, H. A. (1993). *Biodiversity and Ecosystem Function*. Berlin: Springer-Verlag.

Simpson, E. H. (1949). Measurement of diversity. *Nature* 163, 688–688. doi: 10.1038/163688a0

Spiers, A. J., Kahn, S. G., Bohannon, J., Travisano, M., and Rainey, P. B. (2002). Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics* 161, 33–46. doi: 10.1093/genetics/161.1.33

Stacy, A., McNally, L., Darch, S. E., Brown, S. P., and Whiteley, M. (2016). The biogeography of polymicrobial infection. *Nat. Rev. Microbiol.* 14, 93–105. doi: 10.1038/nrmicro.2015.8

Svanback, R., and Bolnick, D. I. (2007). Intraspecific competition drives increased resource use diversity within a natural population. *Proc. R. Soc. B Biol. Sci.* 274, 839–844. doi: 10.1098/rspb.2006.0198

Tilman, D. (1997). Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* 78, 81–92. doi: 10.1890/0012-9658(1997)078[0081:CIRLAG]2.0.CO;2

Valle, J., Da Re, S., Henry, M., Fontaine, T., Balestrino, D., Latour-Lambert, P., et al. (2006). Broad-spectrum biofilm inhibition by a secreted bacterial polysaccharide. *Proc. Natl. Acad. Sci. U. S. A.* 103, 12558–12563. doi: 10.1073/pnas.0605399103

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer Press.

Zhang, Q. G., Ellis, R. J., and Godfray, H. C. J. (2012). The effect of a competitor on a model adaptive radiation. *Evolution* 66, 1985–1990. doi: 10.1111/j.1558-5646.2011.01559.x

Zhao, X. F., Buckling, A., Zhang, Q. G., and Hesse, E. (2018). Specific adaptation to strong competitors can offset the negative effects of population size reductions. *Proc. R. Soc. B Biol. Sci.* 285:20180007. doi: 10.1098/rspb.2018.0007

Zheng, J., Payne, J. L., and Wagner, A. (2019). Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. *Science* 365, 347–353. doi: 10.1126/science.aax1837

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.

frontiers in Microbiology

REVIEW
published: 27 August 2021
doi: 10.3389/fmicb.2021.719000

Ecology and Evolution of Marine Fungi With Their Adaptation to Climate Change

Vinit Kumar[1], V. Venkateswara Sarma[2], Kasun M. Thambugala[3], Jun-Jie Huang[1], Xiang-Yang Li[1] and Ge-Fei Hao[1]**

[1] State Key Laboratory Breeding Base of Green Pesticide and Agricultural Bioengineering, Key Laboratory of Green Pesticide and Agricultural Bioengineering, Ministry of Education, Center for Research and Development of Fine Chemicals, Guizhou University, Guiyang, China, [2] Department of Biotechnology, Pondicherry University, Puducherry, India, [3] Genetics and Molecular Biology Unit, Faculty of Applied Sciences, University of Sri Jayewardenepura, Nugegoda, Sri Lanka

Climate change agitates interactions between organisms and the environment and forces them to adapt, migrate, get replaced by others, or extinct. Marine environments are extremely sensitive to climate change that influences their ecological functions and microbial community including fungi. Fungi from marine habitats are engaged and adapted to perform diverse ecological functions in marine environments. Several studies focus on how complex interactions with the surrounding environment affect fungal evolution and their adaptation. However, a review addressing the adaptation of marine fungi to climate change is still lacking. Here we have discussed the adaptations of fungi in the marine environment with an example of *Hortaea werneckii* and *Aspergillus terreus* which may help to reduce the risk of climate change impacts on marine environments and organisms. We address the ecology and evolution of marine fungi and the effects of climate change on them to explain the adaptation mechanism. A review of marine fungal adaptations will show widespread effects on evolutionary biology and the mechanism responsible for it.

Keywords: *Aspergillus terreus*, ecology, eDNA, evolution, *Hortaea werneckii*, next-generation sequencing

OPEN ACCESS

Edited by:
Maria Elena Martino, University of Padua, Italy

Reviewed by:
Filomena De Leo, University of Messina, Italy
Sunil Kumar Deshmukh, The Energy and Resources Institute (TERI), India

***Correspondence:**
Xiang-Yang Li
xyli1@gzu.edu.cn
Ge-Fei Hao
gefei_hao@foxmail.com

Specialty section:
This article was submitted to Aquatic Microbiology, a section of the journal Frontiers in Microbiology

Received: *03 June 2021*
Accepted: *10 August 2021*
Published: *27 August 2021*

Citation:
Kumar V, Sarma VV, Thambugala KM, Huang J-J, Li X-Y and Hao G-F (2021) Ecology and Evolution of Marine Fungi With Their Adaptation to Climate Change. Front. Microbiol. 12:719000. doi: 10.3389/fmicb.2021.719000

INTRODUCTION

Global climate change affects the environment through shifts in mean temperatures and climate instability, along with other related changes such as ocean warming, stratification, acidification, eutrophication, and increased atmospheric carbon dioxide. Due to global warming, the average sea level is increasing by ∼3.2 mm per year (Stocker, 2014). Direct effects of climate change may alter the behavior, physiological functioning, and demography of organisms living in these environments (Hunter-Cevera et al., 2005). In marine environments, these alterations may affect species interactions and trophic pathways by propagating climate signals from primary producers to the degraders like marine fungi and affecting it in both bottom-up and top-down directions (Doney et al., 2012; Cavicchioli et al., 2019).

Marine fungi belong to different taxonomic groups and can be found colonizing and adapting to different substrates including driftwood, mangrove wood, roots, pneumatophores, seedlings, leaves of mangrove plants, soils and sediments in marine environments, seawater, and dead and decomposing animal substrata (Kohlmeyer and Kohlmeyer, 1979b; Hyde et al., 2000). Several species from marine habitats such as coastal oligotrophic and upwelling waters, deep-sea sediments, and sediments in anoxic zones have been reported (Damare and Raghukumar, 2008;

Raghukumar, 2008; Gutiérrez et al., 2016; Lear et al., 2018). While filamentous higher marine fungi predominantly occur in coastal habitats such as mangroves and driftwood, the yeasts occur in open seas and deep-sea habitats. Moreover, the recent addition of molecular inputs has shown an increased discovery of novel species of marine fungi (Dayarathne, 2020; Jones et al., 2020; Devadatha et al., 2021). New species of *Aureobasidium*, *Cryptococcus*, *Candida*, *Exophiala*, *Malassezia*, *Rhodosporidium*, and *Rhodotorula*, have also been discovered from marine habitats such as hydrothermal vents and sub-sea floors (Overy et al., 2014; Grossart et al., 2016; Amend et al., 2019). Therefore, novel habitats, substrates, and the environment will remarkably assist in identifying new species that are adapted to them.

Adaptations of marine fungi to climate change can be understood through exploring the ecology and evolution of marine fungi from such extreme habitats. Up to now, several studies have reported ecology and evolution by utilizing morphological and molecular techniques (Sarma and Jeewon, 2019). As a result, novel fungal lineages have been found, based on a minimum of 3% nucleotide difference (Richards et al., 2012). Major adaptation events by fungi have also been explained, eventually providing new insight into the "white origin of fungi" (Rédou et al., 2015; Naranjo-Ortiz and Gabaldón, 2019). These works are widely regarded as pioneering which has led the experimental demonstrations on the ecology and evolution of marine fungi to expand at an ever-increasing rate. However, a comprehensive review of the adaptation of marine fungi to climate change, through their ecology and evolution is relatively rare.

We attempted to provide an overview of the adaptation of marine fungi to climate change through the above-mentioned topics. The importance of using molecular techniques to study the ecology and evolution of marine fungi is also discussed here. It may help scientists to improve current research practices to understand the purpose of adaptation that is important for evolution, where different ecological agents are expected to have different purposes. It may also allow non-professionals to better understand how marine fungi widen marine microbiological horizons. Information on fungal halotolerant genes or genes involved in adaptation to climatic changes will help in developing transgenic plants which may tolerate conditions like high salinity and high temperature.

ECOLOGY OF MARINE FUNGI: CHALLENGES AND CONCERNS

Marine fungi play an important role in energy flow, exopolysaccharide complexes synthesis, and nutrient recycling. They intercede the cycling of dissolved organic matter and select appropriate decomposing techniques, such as comminution, non-enzymic chemical reactions, leaching, and volatilization (Sinsabaugh, 2005; Gessner et al., 2010). They perform denitrification in the hypoxic zones as reported from the Arabian sea (Raghukumar, 2008). Some marine fungi and fungi-like organisms degrade environmental pollutants in marine environments, e.g., *Thraustochytrids*, isolated from chronically polluted by oil spills in Goa, can degrade tar-balls (Raikar et al., 2001; Raghukumar, 2008).

Marine fungi produce various extracellular degradative enzymes, e.g., cellulases, ligninases, and xylanases (Raghukumar et al., 2004; Chi et al., 2007; Bonugli-Santos et al., 2015). Some enzymes are associated with nutrient-cycling in the deep-sea and they may be utilized as potential indicators of nutrient cycling processes, e.g., alkaline phosphatase in the deep sea plays a significant role in the recovery of inorganic phosphate by the catalysis of organic esters (Chróst, 1991). Fungi may be engaged in the production of humic aggregates in deep-sea sediments. The aggregate formation holds extracellular enzymes close to the secreting organisms and thus protectors and contributes to the overall sedimental nutrient cycling process (Damare and Raghukumar, 2008). However, it is difficult to comprehend the ecological functions of marine fungi (Gleason et al., 2012). Established roles include plant and algal waste degradation, chemical defense, pathogenicity, symbiosis, and contribution to various holobiont populations (Gleason et al., 2012; Balabanova et al., 2018). Following are some aspects of marine fungal ecology through different culture-dependent and culture-independent techniques.

Culture-Dependent Techniques

Conventional methods of culturing marine fungi include (1) direct detection of fungal reproductive structures on natural samples by observing using stereomicroscope followed by single spore isolation, (2) culturing after surface sterilization of plant leaves or soft animal tissues and particle plating, (3) Baiting followed by culturing, and (4) dilution plating/direct plating (Raghukumar et al., 2010). Many marine fungi have been detected and cultured *via* direct detection and isolation (Kohlmeyer and Kohlmeyer, 1979a; Hyde et al., 2000; Jones et al., 2009, 2015). Some marine fungi are recorded with a higher percentage of occurrence with those encountered with $\geq$10% frequency indicating the "core group" of fungi at that site (Sarma and Hyde, 2001; Sridhar and Maria, 2006). Frequently occurring fungi are primarily studied by utilizing culture-dependent approaches. One such example can be seen from mangroves, where most of the fungi documented to date have been obtained based on culture-dependent methods (Sarma and Hyde, 2001). Using this type of method guarantees identification according to morphological, biochemical, or genetic characteristics (Jany and Barbier, 2008). Marine fungi produce sporulating structures such as ascomata, basidiomata, and conidia-bearing structures (anamorphic stage) in/on the substrates in which they grow actively in the form of hyphae. These substrates include mangrove wood, allochthonous wood, lignocellulosic materials such as coral, decaying leaves, macroalgae, cuttlebone of squids, and exoskeletons of crustaceans.

Culture-dependent approaches are a powerful tool with benefits in manipulating individual isolates, elucidating physiological properties, metabolic interactions among microorganisms and the surroundings, and accordingly provide statistics for their potential ecological roles in ecosystems (Otlewska et al., 2014).

Nonetheless, there are some drawbacks when using culture-dependent methods, e.g., culture-dependent approaches allow the isolation of only a small portion of the overall fungal diversity in an environment (Schabereiter-Gurtner et al., 2001; Jeewon and Hyde, 2007). Many metabolically active strains occur in the environment in the state of anabiosis, being viable but non-culturable (VBNC) and these are left out from being documented (Salma et al., 2013; Otlewska et al., 2014; Schottroff et al., 2018). Therefore, most of the fungal strains in environmental samples cannot be cultured and the culture-dependent methods provide only limited information on the biodiversity of microorganisms from that area (Poli et al., 2017). Moreover, culture-dependent methods are time-consuming due to long culture periods and elaborate culture techniques (Jany and Barbier, 2008; Stefani et al., 2015). To overcome these limitations, culture-independent approaches have recently been developed to explore and access the uncultured microbial community (Overy et al., 2019).

Culture-Independent Techniques

Culture-independent techniques describe taxa more comprehensively than culture-dependent techniques. They are primarily focused on the use of next-generation sequencing (NGS), which along with phylogenetic data offers phenomenal evidence on relationships of species (Raimundo et al., 2018). It plays an important role in exploring marine fungal ecology, nutrient cycling, stress responses, and ecological niche construction (Dupont et al., 2007). NGS technologies have allowed unexplored marine ecosystems to be examined, e.g., hydrothermal vents, Mid-Atlantic Ridge, South Atlantic Ocean, where Ascomycota and Basidiomycota members were dominant, with various new phylotypes (Xu et al., 2017). High throughput sequencing (HTS) study of sediment samples from high Arctic fjord revealed 113 fungal OTUs by using ITS region, defined with a 97% sequence similarity cut-off (Zhang et al., 2015; Rämä et al., 2017).

Besides, genomic sequencing of model organisms from natural communities elucidates their biodiversity that promotes ecological structure, evolution events, taxonomic interactions, life history, and physiological biodiversity. For example, genomic sequencing-based characterization of a shared genomic element (nucleotide transporter) between *Rozella allomycis* and endoparasitic *Microsporidia* suggests that they share a common ancestor and *Rozella* leads a host-dependent lifestyle, where it depends on the host for essential metabolic genes (James et al., 2013).

Studies comparing culture-dependent and NGS techniques revealed wide variations in fungal community composition (Romão et al., 2017). Where, the culture-based approach reported low and variable levels of the species while the NGS methods (ITS1/ITS4 primers) revealed that the whole fungal population included *Purpureocillium lilacinum* in one study (Romão et al., 2017). Ecological roles of the dark matter fungi (DMF) in organic matter cycling have also been studied through environmental DNA sequences (SSU) (Grossart et al., 2016). DMFs are parasitic and saprophytic, they have not been cultured before and are missing from the taxonomy of the fungi.

Next-generation sequencing methods such as Illumina, Ion Torrent, and Pyrosequencing, are mostly used (Schlaeppi et al., 2016; Sanka Loganathachetti et al., 2017; Lendemer et al., 2019). Due to low cost, fast speed, and lack of cloning step in examining fungal diversity, Pyrosequencing was a preferred method (Lim et al., 2010). In a single run, it generates millions of short reads (300–500 nt) with a low error rate (Margulies et al., 2005; Buée et al., 2009). Despite this, it was discontinued in 2013, due to its non-competitiveness. In addition, the Illumina sequencing technology is widely used; though, it yields shorter reads than pyrosequencing, as this is improving rapidly (Snyder et al., 2010). In comparison to other NGS technologies, Illumina has a vast amount of sequencing depth. Further, Ion Torrent uses semiconductors, resulting in a higher number of sequence reads and faster processing times. It is being used to investigate fungal populations in mangrove and deep-sea soil compartments, where, *Aspergillus*, *Penicillium*, red-pigmented basidiomycetous yeasts, psychrotrophic fungi, and other uncultured deep-sea taxa were discovered (Nagano et al., 2017; Sanka Loganathachetti et al., 2017).

Large sequence reads are generated by technologies like PacBio, allowing complete lengths of barcode genes to be accessed (Kyaschenko et al., 2017; Tedersoo et al., 2017). It has been effectively used for the investigation of fungi metabarcoding. Despite this, researchers have been hesitant to use this approach due to its poor throughput, high error rate, and ever-changing bioinformatics methods (Tedersoo et al., 2017). However, advances in data processing algorithms applied to NGS data have reduced sequencing data preferences error rate, rendering them more accurate (Edgar, 2013).

Molecular methods such as DNA metabarcoding, in addition to NGS methods, allow to identify significantly greater taxonomic biodiversity within the samples (Stat et al., 2017). However, for effective fungal identification, the error rate is still too high (Li et al., 2019). Amplified rDNA restriction analysis, amplified ribosomal intergenic spacer analysis, denaturing gradient gel electrophoresis (DGGE), temporal temperature gradient gel electrophoresis (TTGE) and single-strand conformation polymorphism are some of the other techniques that can be used (Pang and Mitchell, 2005; Jany and Barbier, 2008). For example, DGGE was used to investigate fungal diversity in coastal areas where Ascomycota, Basidiomycota, Chytridiomycota, and novel environmental fungal clades predominated (Gao et al., 2010; Cury et al., 2011). However, since most research focuses on non-fungal microbial diversity, only a handful of these approaches have been used in the study of marine habitats (Raghukumar, 2008). As a result, the scope for using culture-independent approaches in marine habitats is enormous, and it holds great promise for revealing heretofore unknown fungal diversity (Raghukumar, 2017).

TAXONOMY AND EVOLUTION OF MARINE FUNGI

Fungi transitioned multiple times from marine to terrestrial environments, and vice versa (Amend et al., 2019). Numerous

reports have recommended that fungi with plants were the first eukaryotes to inhabit the land, with mycorrhizal symbioses allowing this to happen (Lutzoni et al., 2018). Fungi inhabited land during the Cambrian to Ordovician periods, according to molecular dating (542–488.3 and 488.3–443.7 Mya) (Pisani et al., 2004; Dunn, 2011; **Figure 1**). The members in phylum Glomeromycota play a pivotal role in this colonization process (Hyde et al., 2017; Lutzoni et al., 2018). This has led to the theory that Glomeromycota members who lived in symbiotic relationships with cyanobacteria or algae eventually were symbionts of early land plants (Schüßler, 2002).

Studies proposed that marine ascomycetes diverged from many independent migrations of terrestrial and freshwater lineages to the sea (Jones et al., 2019). Where, bituniciates ascomycetes prefer tropical mangrove environments and unitunicates prefer temperate oceanic climates (Sharpe et al., 2015). Several freshwaters and terrestrial fungal genera, such as *Anthostomella*, *Didymella*, *Leptosphaeria*, *Lophiostoma*, *Massarina*, *Mycosphaerella*, *Passeriniella*, *Phaeosphaeria*, *Phomatospora*, *Saccardoella*, *Savoryella*, and *Trematosphaeria*, have marine members, indicating various land-sea transitions (Kohlmeyer and Kohlmeyer, 1979a; Hyde et al., 2000; Vijaykrishna et al., 2006; Jones et al., 2009; Suetrong et al., 2009; Sakayaroj et al., 2011).

Furthermore, Chytridiomycota and Rozellomycota (syn. Cryptomycota) are both marine early-diverging lineages. Chytridiomycota is flagellated fungi that diverged around 750 Mya (Chang et al., 2015). In terrestrial fungi that produce non-motile spores, the existence of flagellated zoospores was eventually lost (Sekimoto et al., 2011). Chytrids are parasitic organisms that infest phytoplankton and cyanobacteria. They may be saprobes or necrotrophs (Planktothrix) found mainly in nearshore and sediment samples (Comeau et al., 2016; Agha et al., 2018). Rozellomycota, on the other hand, has been found in anoxic marine environments, with an estimated divergence time between 408 and 1078 Mya (Grossart et al., 2016; Li et al., 2016; Tedersoo et al., 2018). Rozellomycota includes Rozella species as well as LKM-11 cluster sequences distributed in anoxic, aquatic, and marine habitats (Raghukumar, 2017).

What role do early diverging fungi play in the evolution of marine fungi, and what contribution do they make to the evolutionary system? The ARM (Aphelidomycota, *Rozella*, *Microsporidia*) clade includes early diverging communities such as Rozellomycota, Chytridiomycota, Mucoromycota, and Microsporidia, which are also pathogens of various other eukaryotes, such as amoebae, algae, and other fungi (James and Berbee, 2012; Tedersoo et al., 2018). The ARM clade can be dated back to the adaptation to intracellular parasitism (Corsaro et al., 2014). Aphelids, on the other hand, have more transitional characteristics than fungi and represent an earlier lineage in the holomycotan clade (Corsaro et al., 2014; Karpov et al., 2017). Since these classes have marine members, Aphelidiomycota (*Pseudaphelidium*), Chytridiomycota, and Rozellomycota would be important in understanding the evolution of marine fungi in the future (Comeau et al., 2016; Hassett and Gradinger, 2016; Karpov et al., 2017; Jones et al., 2019). However, they suffer from insufficient taxon sampling (Tedersoo et al., 2018).

ARM clade has a sister group called the Mycetaen fungi, which evolved between 760 Mya–1.06 Bya (Gingras et al., 2011; Beraldi-Campesi, 2013; Raghukumar, 2017; Jones et al., 2019). The Holomycota is the reference to both groups (1.108 Bya) (Jones, 2011; Raghukumar, 2017; Tedersoo et al., 2018). Mycetaen fungi and Metazoa (Animalia) are members of the Opisthokonta Division (1.240–1.481 Bya) and share a common ancestor that split during the Neoproterozoic era (Cohen et al., 2009; Parfrey et al., 2011; Doglioni et al., 2016; Tedersoo et al., 2018). The division of extant fungi and metazoans from a single ancestor is estimated in the early and mid-Neoproterozoic (Berney and Pawlowski, 2006; Sharpe et al., 2015). Whereas the age of the last eukaryotic common ancestor (LECA) is estimated to be between 1007 (943–1102) and 1898 (1655–2094) Mya (Eme et al., 2017; **Figure 1**).

During the Neoproterozoic era, salinity dropped with major cooling events, resulting in glaciations, and allowing dissolved oxygen (O_2) into the ocean (Knauth, 2005). The Precambrian was responsible for increasing O_2 accumulation in the atmosphere before the Neoproterozoic, but eukaryotes had only started to evolve during this period (Knauth, 2005; Doglioni et al., 2016). The key concern was the amount of dissolved oxygen in the water, which improved during Neoproterozoic. Based on divergence time analyses, the origin of fungi dates to 888 and 966 Mya, where Blastocladiomycota, Chytridiomycota, and Rozellomycota diverged around 750 Mya, indicating that early development of fungi occurred in the aquatic habitat (Sanderson et al., 2004; Chang et al., 2015; Tedersoo et al., 2018). However, this raises a debate about whether marine fungi originated from marine or freshwater habitats. To answer, we want to mention a few points, (1) salinity decreased in Neoproterozoic and O_2 levels increased; (2) in the Precambrian era there was an increasing atmospheric level of O_2, with atmospheric oxygenation, more diluted waters, such as lakes, rivers, and streams oxygenated well ahead of the ocean. Therefore, during the evolution in Precambrian and Neoproterozoic, lowering of temperature and salinity, and increased dissolved O_2 were the main determinants that suggest the origin of the fungi in freshwater (James et al., 2006; Lücking et al., 2009; Raghukumar, 2017). The ocean allowed non-marine early fungi to transition to the current marine niche due to the significant decrease in salinity and a rise in O_2. This claim thus indicates the origin of existing marine fungi in freshwater. However, the oceanic origin of existing marine fungi is still discussed (le Calvez et al., 2009; Raghukumar, 2017).

The concern about the fungal origin is most likely answered by high-throughput sequencing of genetic markers in several freshwaters and aquatic environments (Lear et al., 2018). As we learn more about the early-diverging fungal species, we would better understand the underlying mechanism of multicellularity, fungal colonization on land, and the origin of marine fungi.

CONSEQUENCES AND ADAPTATION OF MARINE FUNGI TO CLIMATE CHANGE

In the marine ecosystem, almost all organisms depend on fungi, for the decomposition and recycling of carbon and minerals. To

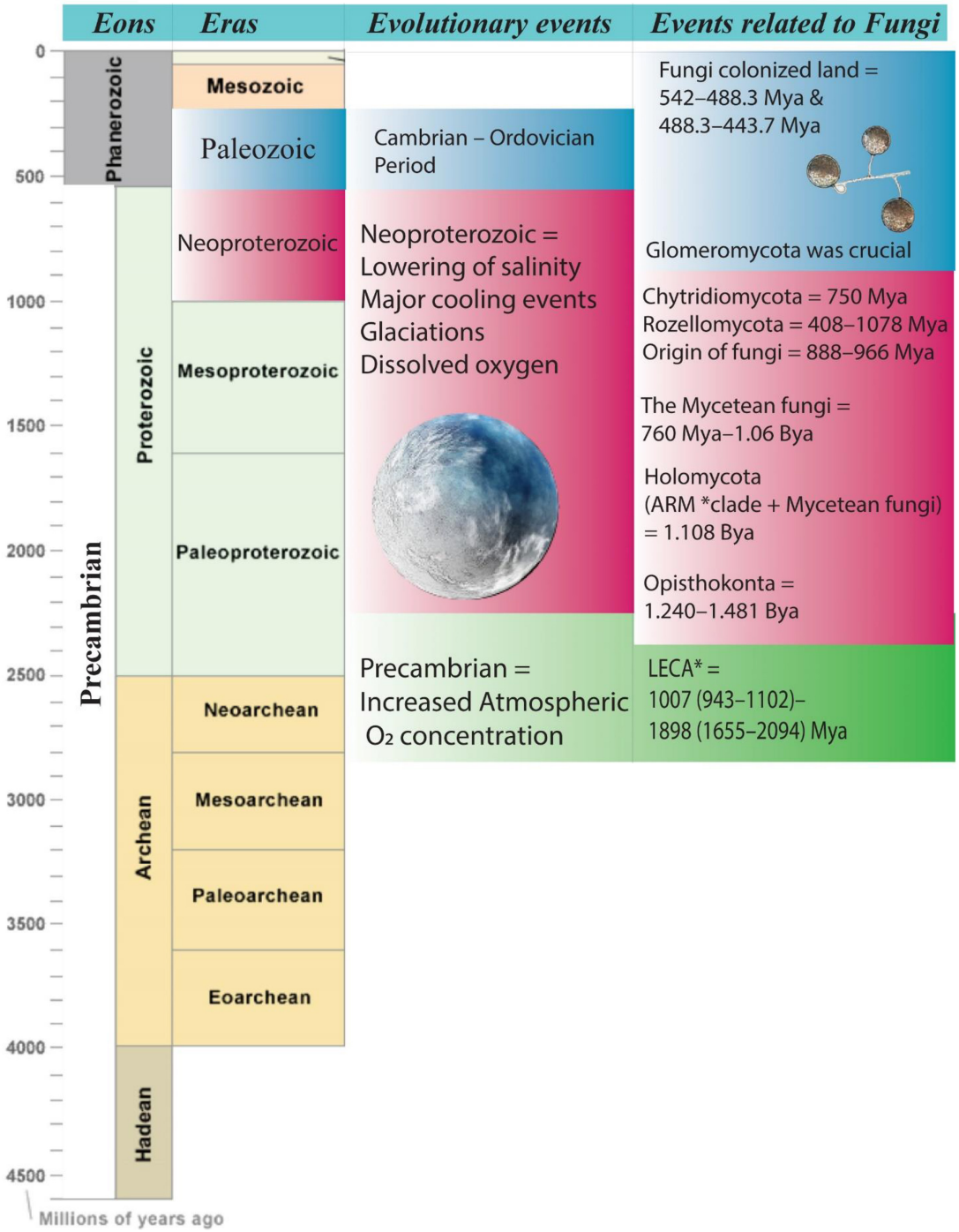


FIGURE 1 | Schematic representation of evolutionary timeline related to fungi. *ARM, Aphelidomycota, *Rozella*, *Microsporidia*; *LECA, last eukaryotic common ancestor.

understand the consequences of climate change it is essential to understand the critical response of fungi toward it. In fungal growth, fruiting, and distribution in marine environments, climates play a dynamic and critical role. However, marine environments are getting progressively fragile due to various natural and anthropogenic stressors, including increased human population pressures, pollution, habitat loss, and degradation (Crain et al., 2008). Numerous biotic and abiotic factors influence the composition and distribution of marine fungal species (Hyde et al., 2000; Jones, 2011). Here we discussed several issues concerning the impact and adaptation of marine fungi to climate change.

Increased CO_2 Levels

Carbon dioxide is one of the most dissolved gas in seawater. In addition to salinity and aridity, greater concentrations of CO_2 can be damaging to various marine fungi contributing to significant changes in vegetation (Sandilyan and Kathiresan, 2012). Increased degrees of atmospheric CO_2 influence both the host and the associated fungal communities (Maček et al., 2019). Increased mycelium growth has been documented because of high CO_2, including in arbuscular mycorrhizal fungi (Maček et al., 2019). Change has been observed in the community of Basidiomycetes on the coastal scrub oak forest soil after a 5-year treatment of soil with elevated CO_2 (Klamer et al., 2002). This suggests, with the elevation in CO_2 level, the composition of the soil fungal community changes significantly (Tu et al., 2015). However, the impact of CO_2 on the fungal population from various environments should not be overlooked as it alters to some degree the community structure.

Rising Temperatures

Temperature is a significant element affecting the worldwide distribution of marine fungi (Jones, 2000). Temperature fluctuations can significantly impact marine ecosystems and marine organisms, e.g., mangrove forests that are integral to marine ecosystems need a photosynthesis temperature of 28–32°C, leaf temperatures at 38–40°C can dramatically decrease the growth of mangrove trees, thus affecting net productivity. Increased temperature in fungal communities by 4–8°C leads to compositional changes that favor various classes of decomposers and promote the degradative succession of fungi (Venkatachalam et al., 2019). Species that are already under stress suffer the most due to habitat loss. Some obligate, as well as facultative marine fungi, follow the "Phoma-pattern" and prefer higher temperatures to grow (Ritchie, 1957; Venkatachalam et al., 2019), allowing mycetaen marine fungi to grow better at increased salinities and temperatures. This may lead to false-positive results while evaluating the fungal diversity of species involved in decomposition and restricted to temperature regimes (Chen et al., 2011; Xu et al., 2014).

Rising Sea-Levels

Sea-level has a major influence on the climate and fungal diversity. Studies have shown the potential to disrupt marine ecosystems through rising sea levels. Strong cyclones may destroy mangroves by defoliation, uprooting, and tree mortality because of the accelerated increase in sea level. An increase in sea levels raises surface water and groundwater salinity by 1–33% in 25 consecutive years through saltwater intrusion which potentially influences the aquatic food web, food security, and expansion of salt-affected arable lands (Rahman et al., 2018; Ullah et al., 2021). Alongside this, the characteristics of soil sediments also change, which affects the fungal demography (Woodruff et al., 2013; Li et al., 2016; Tisthammer et al., 2016). Fungi which are dependent on host plants may also be affected when plants are affected in coastal ecosystems. Nevertheless, there have been no formal threats against marine fungi due to rises in sea levels. Studies relating the evaluation of marine fungal assemblages and environmental data with human development are thus required to forecast the response to climate change and identify directions for future coordinated management of marine ecosystems.

Adaptation of Marine Fungi to Climate Change

Studies on the adaptive capacity of certain organisms to climate change and in extreme habitats could alleviate the detrimental impacts predicted by future climate change. The potential of some fungi for their phenotypic and genetic adaptation in response to extreme habitat and climate change has been acknowledged here. Several fungi have characterized themselves to adapt in environments with low water activity and high concentrations of toxic ions by complex molecular and cellular adaptation (Gostinèar et al., 2011; Gladfelter et al., 2019; Romeo et al., 2020). Spores of some marine fungi have developed strategies of sheath and appendages to attach, float, and adapt to a new environment. Similarly, some marine arenicolous fungi have developed subiculum to attach to sand grains and tolerate extreme conditions like high temperature, variation in salinity, and desiccation (Jones, 2000). These adaptation mechanisms and their effects are highly dependent on different levels of biological organization and follow a cascade of events (molecular and cellular, whole organisms, population, and community). Fungi isolated from hypersaline marine environments, deep-sea hydrothermal vents, and deep-sea sediments have different molecular and cellular mechanisms for adaptation, e.g., gene expression, high osmolarity glycerol (HOG) signaling pathway, melanization of the cell wall, composition and accumulation of ions, enzymes involved in fatty acid modifications, and plasma membrane composition (Turk and Plemenitaš, 2002; Kogej et al., 2006; Romeo et al., 2020). Fungi with these characters have been described for genera such as *Aspergillus*, *Cladosporium*, *Emericella*, *Eurotium*, *Hortaea*, *Trimmatostroma*, and *Wallemia*. Here we have discussed what molecular and cellular mechanisms determine the adaptation of *Hortaea werneckii* to marine habitat and how *H. werneckii* deals with climatic stressors such as high salinity (3–4.5 M) and temperature (Kejžar et al., 2015; Gunde-Cimerman et al., 2018; Romeo et al., 2020; Selbmann et al., 2020). Also, we have addressed the adaptations of *Aspergillus terreus* to the extreme conditions of a hydrothermal vent (Pang et al., 2020).

In yeasts (*Saccharomyces cerevisiae*), cells respond to the stress signals through sensing the environmental stimuli by mitogen-activated protein kinases (MAPKs) which ensure adaptation to the current environment (Plemenitaš et al., 2014). The sensor on the plasma membrane binds stimulus to the central MAPK cascade through membrane proteins, tyrosine kinase receptors, G-protein-coupled receptors, and histidine-aspartic phosphorylation sensors. The signal from sensors is passed to the MAPK kinase kinase (MAPKKKs), which phosphorylates the MAPK kinase (MAPKKs), activation of MAPKKs leads to the activation of MAPKs. MAPKs are then translocated to the nucleus to activate multiple factors for the adaptive transcriptional response. One such MAPK cascade called the HOG signaling pathway, has been characterized and conserved in *H. werneckii* (HwHOG) which acts as a

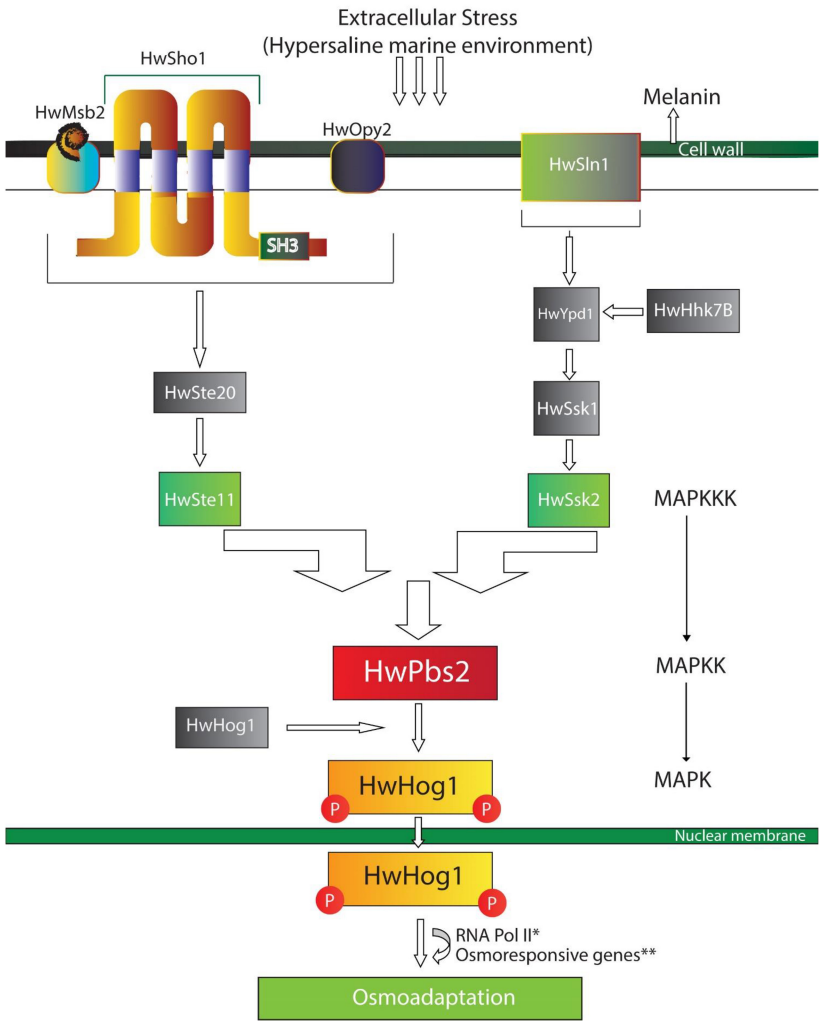


FIGURE 2 | A schematic model of HOG – signaling pathway in *Hortaea werneckii*. Here, HwSho1 along with HwMsb2 and HwOpy2 interact with external stimuli (changes in salinity and osmolarity) and transmits the signal to HwHog1 (MAPK) mediated pathway through HwSte11 (MAPKKK) and HwPbs2 (MAPKK). In *H. werneckii*, the phosphorelay (HwSln1-HwYpd1-HwSsk1) interacts with cytosolic histidine-kinase (Hwhhk7) to transmit the external signal to HwSsk2. Signals from both sensors (HwSho1 and HwSln1) converge at HwPsb2, HwPsb2 activates HwHog1. Activated HwHog1 translocates to the nucleus and interacts with osmoresponsive genes and RNA Pol II for adaptation in fluctuating and hyperosmotic environments. *May or may not interact, **genes responsible for glycerol production, protein synthesis, amino acid metabolism, lipid metabolism.

survival and adaptation tool in hypersaline, marine, and deep-sea environments (Plemenitaš et al., 2014; Gunde-Cimerman et al., 2018; Romeo et al., 2020).

Hortaea werneckii is the most studied eukaryotic model organism in adaptive extremophiles, it can grow with or without salt and there is plenty of literature available on it. *H. werneckii* is placed in the family Teratosphaeriaceae (Capnodiales, Dothideomycetes), and can be found from beach soil, microbial mats, environments with low water activity, salty food, seawater, wood immersed in hypersaline waters, and rocks in tropical or subtropical coastal areas (Kejžar et al., 2015; Gunde-Cimerman et al., 2018; Selbmann et al., 2020).

However, *H. werneckii* is still not considered a marine fungus (Jones et al., 2015). Despite it has recently been isolated from different depths of the Mediterranean Sea and shallow hydrothermal vent as a common fungus (De Leo et al., 2019; Pang et al., 2019; Romeo et al., 2020). Pang et al. (2019) also considered that *H. werneckii* could be a marine fungus. Recently, based on observed phylogenomic differences between different strains, it has been found that the marine *H. werneckii* strains are derived by intraspecific hybridization, suggesting that marine strains are adapting and evolving in this environment (Romeo et al., 2020).

Whole genome sequencing of a marine *H. werneckii* strains reported the up regulations and activation of genes related to stress-activated MAPK cascade (GO: 0051403), MAPKKK activity (GO: 0004709), cellular response to osmotic stress, heat, and oxidative stress (GO: 0071470, GO: 0034605, GO: 0034599), and regulation of mitotic cell cycle (GO: 0007346). Besides this, several Heat Shock Proteins (Hsps) were also activated such as HSP88, HSP78 mitochondrial, and HSP

DnaJ (Romeo et al., 2020). Suggesting that cells of *H. werneckii* respond and adapt to fluctuating NaCl concentrations, heat, and extreme conditions in its environment by immediate responses, without requiring the synthesis of new proteins but modulate the pre-existing ones in metabolism and membrane transport mechanisms. Such modulation requires the activation of signal transduction pathways by stress signals (Kejžar et al., 2015; Gunde-Cimerman et al., 2018; Román et al., 2020). The combination of signaling pathways enables cells to resume growth and adapt to the conditions of marine and other extreme environments, leading toward their adaptation (Marchetta et al., 2018; Román et al., 2020; Romeo et al., 2020; **Figure 2**).

Similarly, in *A. terreus* (Aspergillaceae, Eurotiales); genes related to MAPKs and the HOG signaling pathway were up-regulated along with several other pathways, during intracellular osmotic balance and stress tolerance. Pang et al. (2020) investigated the growth of *A. terreus* at different pH, temperature, and salinity, mimicking the stressed environment, also the molecular adaptation of *A. terreus* was studied with the transcriptome analysis. At higher temperature, salinity, and low pH, *A. terreus* was able to grow optimally. This was validated by the transcriptome results, where, due to high temperature higher expressions of Hsps were observed along with that high temperature-induced reactive oxygen species (ROS), to counter this, genes related to catalases and superoxide dismutase were up-regulated (Abrashev et al., 2014; Pang et al., 2020).

Several proteins are associated with fungal reactions to acidic pH, such as up-regulation of phenylalanine ammonia-lyase (PAL), ATP-binding cassette transporters (ABC), and gamma-aminobutyric acid, while pH binding transcription factor (Pac C) and acetyl xylan esterase were down-regulated (Pang et al., 2020). To tolerate the salinity stress, genes related to arginine metabolism, HOG-pathway, MAPK, and linoleic acid were up-regulated, suggesting their involvement in intracellular osmotic balance. Pang et al. (2020), suggest that marine *Aspergillus* species are able to tolerate a range of environmental stress with the help of their stress-related genes and they could be a great source of such genes for transgenic studies.

Apart from these species, marine fungi such as *Aspergillus aculeatus*, *Microascus brevicaulis*, *Penicillium oxalicum*, and *Trichoderma harzianum* are also able to grow at different environmental conditions including high temperature and salinity, suggesting that they could also adapt to different changes in the environmental conditions (Jones et al., 2015; Pang et al., 2016, 2020).

CONCLUSION AND FUTURE PROSPECTS

1. Adaptation of marine fungi to climate change and extreme environment is an unending topic and there are more to be understood regarding it. In this review, we discussed how climate change leads to changes in the physicochemical properties of marine habitats, which alters the ecological structure, function, physiology, and population of individual species. Also, how certain marine fungal groups have adapted to these conditions. We have also discussed the molecular ecology and evolution of marine fungi and their origin based on the existing literature.
2. Cellular and molecular alterations due to climate change could be either as a response to the changing dynamics of the marine environment or due to the direct influence of stressors on the fungi. To counter stressors like osmotic and temperature fluctuations, *H. werneckii* utilizes HOG-pathway for osmoadaptation. Osmoadaptation includes adjustments in metabolism, cell surface properties, cell morphogenesis, growth and proliferation, and cellular protectant production, such as glycerol, erythritol, arabitol, mannitol. The regulation occurs either by activation and/or recruitment of specific transcriptional factors or associating with RNA polymerase-II, or both. *H. werneckii* produces melanin, which accumulates on the outer cell wall, forms a dense shield-like layer to protect the cell from UV. At optimal NaCl concentration, melanized cell walls help to retain glycerol in the cell.
3. Similarly, *A. terreus* isolated from shallow hydrothermal vents were able to grow at 45°C, pH 3, and 30% salinity. Along with this, genes related to stress tolerance were also shown to be up-regulated, suggesting the molecular adaptation of *A. terreus* to extreme conditions and environmental changes (Pang et al., 2020).
4. Marine fungi can adapt to the high salinity, temperature, and severe pH levels, which provides them with greater variety in biotechnological applications and offers an important biological advantage over terrestrial fungi. The understanding of the adaptation of marine fungi in extreme environments would help researchers to develop transgenic plants that can grow in such environments and provide greater flexibility during changing climatic conditions.
5. Studies on marine fungal ecology and evolution are rare, this could be due to (1) fewer marine fungal taxa were recorded than terrestrial habitats; (2) large geographical areas were still not explored; (3) some marine fungal taxa had recently been investigated with a lot of work yet to be carried out; (4) convergent evolution may have masked evolutionary relationships, and (5) a huge amount of marine water dilutes any evidence of environmental genetic material available. Through using molecular techniques with culture-based approaches, new environmental sequences from many marine environments will make a significant contribution to fungal diversity (Richards et al., 2012; Zhang et al., 2014).
6. There are continuing discussions about whether sequence-based species description can be accepted and recognized for taxon identification? (Hawksworth and Lücking, 2017; Hongsanan et al., 2018; Thines et al., 2018). In our opinion, the classification by sequence is not sufficient but may provide leads to the fungal diversity of unresolved taxa from the provided environmental samples.
7. We also addressed some of the current hypotheses concerning the origin and development of marine fungi

with their rationales. We suggest, the extant marine fungi were originated from freshwater and subsequently moved from land and, because of the strong natural selection, the species evolved and adapted to the sea.

AUTHOR CONTRIBUTIONS

VK designed this study with VS and G-FH. VK collected the literature and prepared the first draft of the manuscript. KT and VS critically revised the manuscript. J-JH and X-YL provided their valuable insights to the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

This work was supported in part by the National Key R&D Program (2018YFD0200100) and the National Natural Science Foundation of China (21772059, 91853127, 31960546, and 31960548). Program of Introducing Talents of Discipline to Universities of China (111 Program, D20023).

ACKNOWLEDGMENTS

VK would like to thank the reviewers for their time and valuable suggestions on the manuscript.

REFERENCES

Abrashev, R., Stoitsova, S., Krumova, E., Pashova, S., Paunova-Krasteva, T., Vassilev, S., et al. (2014). Temperature-stress tolerance of the fungal strain *Aspergillus niger* 26: physiological and ultrastructural changes. *World J. Microbiol. Biotechnol.* 30, 1661–1668. doi: 10.1007/s11274-013-1586-8

Agha, M., Ennen, J. R., Bower, D. S., Nowakowski, A. J., Sweat, S. C., and Todd, B. D. (2018). Salinity tolerances and use of saline environments by freshwater turtles: implications of sea level rise. *Biolog. Rev.* 93, 1634–1648. doi: 10.1111/brv.12410

Amend, A., Burgaud, G., Cunliffe, M., Edgcomb, V. P., Ettinger, C. L., Gutiérrez, M. H., et al. (2019). Fungi in the marine environment: Open questions and unsolved problems. *mBio* 10, e1189–e1118. doi: 10.1128/mBio.01189-18

Balabanova, L., Slepchenko, L., Son, O., and Tekutyeva, L. (2018). Biotechnology potential of marine fungi degrading plant and algae polymeric substrates. *Front. Microbiol.* 9:1527. doi: 10.3389/fmicb.2018.01527

Beraldi-Campesi, H. (2013). Early life on land and the first terrestrial ecosystems. *Ecol. Proc.* 2, 1–17. doi: 10.1186/2192-1709-2-1

Berney, C., and Pawlowski, J. (2006). A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record. *Proc. R. Soc. B* 273, 1867–1872. doi: 10.1098/rspb.2006.3537

Bonugli-Santos, R. C., Vasconcelos, M. R., dos, S., Passarini, M. R. Z., Vieira, G. A. L., Lopes, V. C. P., et al. (2015). Marine-derived fungi: Diversity of enzymes and biotechnological applications. *Front. Microbiol.* 6:269. doi: 10.3389/fmicb.2015.00269

Buée, M., Reich, M., Murat, C., Morin, E., Nilsson, R. H., Uroz, S., et al. (2009). 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytologist*. 184, 449–456. doi: 10.1111/j.1469-8137.2009.03003.x

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., et al. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology*. 17, 569–586. doi: 10.1038/s41579-019-0222-5

Chang, Y., Wang, S., Sekimoto, S., Aerts, A. L., Choi, C., Clum, A., et al. (2015). Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. *Genome Biol. Evol.* . 7, 1590–1601. doi: 10.1093/gbe/evv090

Chen, H. Y., Xue, D. S., Feng, X. Y., and Yao, S. J. (2011). Screening and production of ligninolytic enzyme by a marine-derived fungal pestalotiopsis sp. J63. *Appl. Biochem. Biotechnol.* 165, 1754–1769. doi: 10.1007/s12010-011-9392-y

Chi, Z., Ma, C., Wang, P., and Li, H. F. (2007). Optimization of medium and cultivation conditions for alkaline protease production by the marine yeast Aureobasidium pullulans. *Bioresour. Technol.* 98, 534–538. doi: 10.1016/j.biortech.2006.02.006

Chróst, R. J. (1991). Environmental Control of the Synthesis and Activity of Aquatic Microbial Ectoenzymes. *Aquat. Env.* 1991:317. doi: 10.1007/978-1-4612-3090-8_3

Cohen, P. A., Knoll, A. H., and Kodner, R. B. (2009). Large spinose microfossils in Ediacaran rocks as resting stages of early animals. *Proc. Natl. Acad. Sci. U S A.* 16, 6519–6524. doi: 10.1073/pnas.0902322106

Comeau, A. M., Vincent, W. F., Bernier, L., and Lovejoy, C. (2016). Novel chytrid lineages dominate fungal sequences in diverse marine and freshwater habitats. *Sci. Rep.* 6, 1–6. doi: 10.1038/srep30120

Corsaro, D., Walochnik, J., Venditti, D., Steinmann, J., Müller, K. D., and Michel, R. (2014). Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. *Parasitol. Res.* 113, 1909–1918. doi: 10.1007/s00436-014-3838-4

Crain, C. M., Kroeker, K., and Halpern, B. S. (2008). Interactive and cumulative effects of multiple human stressors in marine systems. *Ecol. Lett.* 11, 1304–1315. doi: 10.1111/j.1461-0248.2008.01253.x

Cury, J. C., Araujo, F. V., Coelho-Souza, S. A., Peixoto, R. S., Oliveira, J. A. L., Santos, H. F., et al. (2011). Microbial diversity of a Brazilian coastal region influenced by an upwelling system and anthropogenic activity. *PLoS One* 6:e16553. doi: 10.1371/journal.pone.0016553

Damare, S., and Raghukumar, C. (2008). Fungi and macroaggregation in deep-sea sediments. *Microb. Ecol.* 56, 168–177. doi: 10.1007/s00248-007-9334-y

Dayarathne, M. (2020). Morpho-molecular characterization of microfungi associated with marine based habitats. *Mycosphere* 11, 1–188. doi: 10.5943/mycosphere/11/1/1

De Leo, F., Lo Giudice, A., Alaimo, C., De Carlo, G., Rappazzo, A. C., Graziano, M., et al. (2019). Occurrence of the black yeast Hortaea werneckii in the Mediterranean Sea. *Extremophiles* 23, 9–17. doi: 10.1007/s00792-018-1056-1

Devadatha, B., Jones, E. B. G., Pang, K. L., Abdel-Wahab, M. A., Hyde, K. D., Sakayaroj, J., et al. (2021). *Occurrence and geographical distribution of mangrove fungi*. Netherlands: Springer, 106, 137–227. doi: 10.1007/s13225-020-00468-0

Doglioni, C., Pignatti, J., and Coleman, M. (2016). Why did life develop on the surface of the Earth in the Cambrian? *Geosci. Front.* 7, 865–873. doi: 10.1016/j.gsf.2016.02.001

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. *Ann. Rev. Mar. Sci.* 4, 11–37. doi: 10.1146/annurev-marine-041911-111611

Dunn, R. (2011). How fungi made us hot blooded. *New Sci.* 212, 50–53. doi: 10.1016/S0262-4079(11)62984-9

Dupont, S., Wilson, K., Obst, M., Sköld, H., Nakano, H., and Thorndyke, M. C. (2007). Marine ecological genomics: When genomics meets marine ecology. *Mar. Ecol. Prog. Series* 332, 257–273. doi: 10.3354/meps332257

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Eme, L., Spang, A., Lombard, J., Stairs, C. W., and Ettema, T. J. G. (2017). Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* 15:711. doi: 10.1038/nrmicro.2017.133

Gao, Z., Johnson, Z. I., and Wang, G. (2010). Molecular characterization of the spatial diversity and novel lineages of mycoplankton in Hawaiian coastal waters. *ISME J.* 4, 111–120. doi: 10.1038/ismej.2009.87

Gessner, M. O., Swan, C. M., Dang, C. K., McKie, B. G., Bardgett, R. D., Wall, D. H., et al. (2010). Diversity meets decomposition. *Trends Ecol. Evol.* 25, 372–380. doi: 10.1016/j.tree.2010.01.010

Gingras, M., Hagadorn, J. W., Seilacher, A., Lalonde, S. V., Pecoits, E., Petrash, D., et al. (2011). Possible evolution of mobile animals in association with microbial mats. *Nat. Geosci.* 4, 372–375. doi: 10.1038/ngeo1142

Gladfelter, A. S., James, T. Y., and Amend, A. S. (2019). Marine fungi. *Curr. Biol.* 29, 191–195. doi: 10.1016/j.cub.2019.02.009

Gleason, F. H., Carney, L. T., Lilje, O., and Glockling, S. L. (2012). Ecological potentials of species of Rozella (Cryptomycota). *Fungal Ecol.* 5, 651–656. doi: 10.1016/j.funeco.2012.05.003

Gostinèar, C., Grube, M., and Gunde-Cimerman, N. (2011). Evolution of Fungal Pathogens in Domestic Environments? *Fungal Biol.* 115, 1008–1018. doi: 10.1016/j.funbio.2011.03.004

Grossart, H. P., Wurzbacher, C., James, T. Y., and Kagami, M. (2016). Discovery of dark matter fungi in aquatic ecosystems demands a reappraisal of the phylogeny and ecology of zoosporic fungi. *Fungal Ecol.* 19, 28–38. doi: 10.1016/j.funeco.2015.06.004

Gunde-Cimerman, N., Plemenitaš, A., and Oren, A. (2018). Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiol. Rev.* 42, 353–375. doi: 10.1093/femsre/fuy009

Gutiérrez, M. H., Jara, A. M., and Pantoja, S. (2016). Fungal parasites infect marine diatoms in the upwelling ecosystem of the Humboldt current system off central Chile. *Env. Microbiol.* 18, 1646–1653. doi: 10.1111/1462-2920.13257

Hassett, B. T., and Gradinger, R. (2016). Chytrids dominate arctic marine fungal communities. *Env. Microbiol.* 18, 2001–2009. doi: 10.1111/1462-2920.13216

Hawksworth, D. L., and Lücking, R. (2017). Fungal diversity revisited: 2.2 to 3.8 million species. *Fungal King.* 2017, 79–95. doi: 10.1128/9781555819583.ch4

Hongsanan, S., Jeewon, R., Purahong, W., Xie, N., Liu, J. K., Jayawardena, R. S., et al. (2018). Can we use environmental DNA as holotypes? *Fungal Div.* 92, 1–30. doi: 10.1007/s13225-018-0404-x

Hunter-Cevera, J. C., Karl, D., and Buckley, M. (2005). Marine microbial diversity: the key to Earth's habitability. *Am. Acad. Microbiol.* 2005:32687283.

Hyde, K. D., Norphanphoun, C., Abreu, V. P., Bazzicalupo, A., Thilini Chethana, K. W., Clericuzio, M., et al. (2017). Fungal diversity notes 603–708: taxonomic and phylogenetic notes on genera and species. *Fungal Div.* 87, 1–235. doi: 10.1007/s13225-017-0391-3

Hyde, K. D., Venkateswara Sarma, V., and Jones, E. B. G. (2000). Morphology and taxonomy of higher marine fungi. *Mar. Mycol.* 2000, 172–204.

James, T. Y., and Berbee, M. L. (2012). No jacket required - new fungal lineage defies dress code: Recently described zoosporic fungi lack a cell wall during trophic phase. *BioEssays* 34, 94–102. doi: 10.1002/bies.201100110

James, T. Y., Letcher, P. M., Longcore, J. E., Mozley-Standridge, S. E., Porter, D., Powell, M. J., et al. (2006). A molecular phylogeny of the flagellated fungi (Chytridiomycota) and description of a new phylum (Blastocladiomycota). *Mycologia* 98, 860–871. doi: 10.1080/15572536.2006.11832616

James, T. Y., Pelin, A., Bonen, L., Ahrendt, S., Sain, D., Corradi, N., et al. (2013). Shared signatures of parasitism and phylogenomics unite cryptomycota and microsporidia. *Curr. Biol.* 23, 1548–1553. doi: 10.1016/j.cub.2013.06.057

Jany, J. L., and Barbier, G. (2008). Culture-independent methods for identifying microbial communities in cheese. *Food Microbiol.* 25, 839–848. doi: 10.1016/j.fm.2008.06.003

Jeewon, R., and Hyde, K. D. (2007). *Detection and diversity of fungi from environmental samples: traditional versus molecular approaches. In Advanced techniques in soil microbiology*. Berlin: Springer, 1–15.

Jones, E. B. G. (2000). Marine fungi: Some factors influencing biodiversity. *Fungal Div.* 4, 53–73.

Jones, E. B. G. (2011). Are there more marine fungi to be described? *Botanica Marina* 54, 343–354. doi: 10.1515/BOT.2011.043

Jones, E. B. G., Devadatha, B., Abdel-Wahab, M. A., Dayarathne, M. C., Zhang, S. N., Hyde, K. D., et al. (2020). Phylogeny of new marine Dothideomycetes and Sordariomycetes from mangroves and deep-sea sediments. *Botanica Mar.* 63, 155–181. doi: 10.1515/bot-2019-0014

Jones, E. B. G., Pang, K. L., Abdel-Wahab, M. A., Scholz, B., Hyde, K. D., Boekhout, T., et al. (2019). An online resource for marine fungi. *Fungal Div.* 96, 347–433. doi: 10.1007/s13225-019-00426-5

Jones, E. B. G., Sakayaroj, J., Suetrong, S., Somrithipol, S., and Pang, K. L. (2009). Classification of marine Ascomycota, anamorphic taxa and Basidiomycota. *Fungal Div.* 35, 1–187.

Jones, E. B. G., Suetrong, S., Sakayaroj, J., Bahkali, A. H., Abdel-Wahab, M. A., and Boekhoutet, T. (2015). Classification of marine Ascomycota, Basidiomycota, Blastocladiomycota and Chytridiomycota. *Fungal Div.* 73, 1–72. doi: 10.1007/s13225-015-0339-4

Karpov, S. A., Torruella, G., Moreira, D., Mamkaeva, M. A., and López-García, P. (2017). Molecular Phylogeny of Paraphelidium letcheri sp. nov. (Aphelida, Opisthosporidia). *J. Euk. Microbiol.* 64, 573–578. doi: 10.1111/jeu.12389

Kejžar, A., Grötli, M., Tamás, M. J., Plemenitaš, A., and Lenassi, M. (2015). HwHog1 kinase activity is crucial for survival of Hortaea werneckii in extremely hyperosmolar environments. *Fungal Genet. Biol.* 74, 45–58. doi: 10.1016/j.fgb.2014.11.004

Klamer, M., Roberts, M. S., Levine, L. H., Drake, B. G., and Garland, J. L. (2002). Influence of elevated CO2 on the fungal community in a coastal scrub oak forest soil investigated with terminal-restriction fragment length polymorphism analysis. *Appl. Env. Microbiol.* 68:4370. doi: 10.1128/AEM.68.9.4370-4376.2002

Knauth, L. P. (2005). Temperature and salinity history of the Precambrian ocean: Implications for the course of microbial evolution. *Geobiology* 2005, 53–69. doi: 10.1016/B978-0-444-52019-7.50007-3

Kogej, T., Gorbushina, A. A., and Gunde-Cimerman, N. (2006). Hypersaline conditions induce changes in cell-wall melanization and colony structure in a halophilic and a xerophilic black yeast species of the genus Trimmatostroma. *Mycolog. Res.* 110, 713–724. doi: 10.1016/j.mycres.2006.01.014

Kohlmeyer, J., and Kohlmeyer, E. (1979a). Fungi on Mangroves and Other Tropical Shoreline Trees. *Mar. Mycol.* 1979, 92–104. doi: 10.1016/b978-0-12-418350-6.50017-6

Kohlmeyer, J., and Kohlmeyer, E. (1979b). Key to the Filamentous Higher Marine Fungi. *Mar. Mycol.* 1979, 188–211. doi: 10.1016/b978-0-12-418350-6.50029-2

Kyaschenko, J., Clemmensen, K. E., Hagenbo, A., Karltun, E., and Lindahl, B. D. (2017). Shift in fungal communities and associated enzyme activities along an age gradient of managed Pinus sylvestris stands. *ISME J.* 11, 863–874. doi: 10.1038/ismej.2016.184

le Calvez, T., Burgaud, G., Mahé, S., Barbier, G., and Vandenkoornhuyse, P. (2009). Fungal diversity in deep-sea hydrothermal ecosystems. *Appl. Env. Microbiol.* 11, 1588–1600. doi: 10.1128/AEM.00653-09

Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H., Buckley, T., et al. (2018). Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *New Zeal. J. Ecol.* 42, 10–50. doi: 10.20417/nzjecol.42.9

Lendemer, J. C., Keepers, K. G., Tripp, E. A., Pogoda, C. S., McCain, C. M., and Kane, N. C. (2019). A taxonomically broad metagenomic survey of 339 species spanning 57 families suggests cystobasidiomycete yeasts are not ubiquitous across all lichens. *Am. J. Bot.* 106, 1090–1095. doi: 10.1002/ajb2.1339

Li, P. D., Jeewon, R., Aruna, B., Li, H. Y., Lin, F. C., and Wang, H. K. (2019). Metabarcoding reveals differences in fungal communities between unflooded versus tidal flat soil in coastal saline ecosystem. *Sci. Total Environ.* 647, 1230–1238. doi: 10.1016/j.scitotenv.2019.06.473

Li, W., Wang, M., Bian, X., Guo, J., and Cai, L. (2016). A high-level fungal diversity in the intertidal sediment of Chinese seas presents the spatial variation of community composition. *Front. Microbiol.* 7:2098. doi: 10.3389/fmicb.2016.02098

Lim, Y. W., Kim, B. K., Kim, C., Jung, H. S., Kim, B. S., Lee, J. H., et al. (2010). Assessment of soil fungal communities using pyrosequencing. *J. Microbiol.* 48, 284–289. doi: 10.1007/s12275-010-9369-5

Lücking, R., Huhndorf, S., Pfister, D. H., Plata, E. R., and Lumbsch, H. T. (2009). Fungi evolved right on track. *Mycologia* 101, 810–822. doi: 10.3852/09-016

Lutzoni, F., Nowak, M. D., Alfaro, M. E., Reeb, V., Miadlikowska, J., Krug, M., et al. (2018). Contemporaneous radiations of fungi and plants linked to symbiosis. *Nat. Comm.* 9, 1–11. doi: 10.1038/s41467-018-07849-9

Maček, I., Clark, D. R., Šibanc, N., Moser, G., Vodnik, D., Müller, C., et al. (2019). Impacts of long-term elevated atmospheric CO2 concentrations on communities of arbuscular mycorrhizal fungi. *Mole. Ecol.* 28, 3445–3458. doi: 10.1111/mec.15160

Marchetta, A., van den Ende, B. G., Al-Hatmi, A. M. S., Hagen, F., Zalar, P., Sudhadham, M., et al. (2018). Global molecular diversity of the halotolerant fungus Hortaea werneckii. *Life* 8:31. doi: 10.3390/life8030031

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 441, 120–120. doi: 10.1038/nature03959

Nagano, Y., Miura, T., Nishi, S., Lima, A. O., Nakayama, C., Pellizari, V. H., et al. (2017). Fungal diversity in deep-sea sediments associated with asphalt seeps at the Sao Paulo Plateau. *Deep-Sea Res. II: Top. Stud. Oceanogr.* 146, 59–67. doi: 10.1016/j.dsr2.2017.05.012

Naranjo-Ortiz, M. A., and Gabaldón, T. (2019). Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. *Biolog. Rev*. 94, 2101–2137. doi: 10.1111/brv.12550

Otlewska, A., Adamiak, J., and Gutarowska, B. (2014). Application of molecular techniques for the assessment of microorganism diversity on cultural heritage objects. *Acta Biochim. Polon.* 61:2014. doi: 10.18388/abp.2014_1889

Overy, D. P., Berrue, F., Correa, H., Hanif, N., Hay, K., Lanteigne, M., et al. (2014). Sea foam as a source of fungal inoculum for the isolation of biologically active natural products. *Mycology* 5, 130–144. doi: 10.1080/21501203.2014.931893

Overy, D. P., Rämä, T., Oosterhuis, R., Walker, A. K., and Pang, K. L. (2019). The neglected marine fungi, sensu stricto, and their isolation for natural products' discovery. *Mar. Drugs* 17:42. doi: 10.3390/md17010042

Pang, K. L., Chiang, M. W. L., Guo, S. Y., Shih, C. Y., Dahms, H. U., Hwang, J. S., et al. (2020). Growth study under combined effects of temperature, pH and salinity and transcriptome analysis revealed adaptations of *Aspergillus terreus* NTOU4989 to the extreme conditions at Kueishan Island Hydrothermal Vent Field, Taiwan. *PLoS One* 15:e0233621. doi: 10.1371/journal.pone.0233621

Pang, K. L., Guo, S. Y., Chen, I. A., Burgaud, G., Luo, Z. H., Dahms, H. U., et al. (2019). Insights into fungal diversity of a shallow-water hydrothermal vent field at Kueishan Island, Taiwan by culture-based and metabarcoding analyses. *PLoS One* 14:e0226616. doi: 10.1371/journal.pone.0226616

Pang, K. L., and Mitchell, J. I. (2005). Molecular approaches for assessing fungal diversity in marine substrata. *Botan. Mar.* 48, 332–347. doi: 10.1515/BOT.2005.046

Pang, K. L., Overy, D. P., Jones, E. G., da Luz, Calado, M., Burgaud, G., et al. (2016). 'Marine fungi'and 'marine-derived fungi'in natural product chemistry research: toward a new consensual definition. *Fungal Biol. Rev*. 30, 163–175. doi: 10.1016/j.fbr.2016.08.001

Parfrey, L. W., Lahr, D. J. G., Knoll, A. H., and Katz, L. A. (2011). Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci. U S A*. 108, 13624–13629. doi: 10.1073/pnas.1110633108

Pisani, D., Poling, L. L., Lyons-Weiler, M., and Hedges, S. B. (2004). The colonization of land by animals: Molecular phylogeny and divergence times among arthropods. *BMC Biol*. 2, 1–10. doi: 10.1186/1741-7007-2-1

Plemenitaš, A., Lenassi, M., Konte, T., Kejžar, A., Zajc, J., Gostinèar, C., et al. (2014). Adaptation to high salt concentrations in halotolerant/halophilic fungi: A molecular perspective. *Front. Microb*. 5:199. doi: 10.3389/fmicb.2014.00199

Poli, A., Finore, I., Romano, I., Gioiello, A., Lama, L., and Nicolaus, B. (2017). Microbial diversity in extreme marine habitats and their biomolecules. *Microorganisms* 5:25. doi: 10.3390/microorganisms5020025

Raghukumar, C. (2008). Marine fungal biotechnology: An ecological perspective. *Fungal Div*. 31, 19–35.

Raghukumar, C., Raghukumar, S., Sheelu, G., Gupta, S. M., Nagender Nath, B., and Rao, B. R. (2004). Buried in time: Culturable fungi in a deep-sea sediment core from the Chagos Trench, Indian Ocean. *Deep-Sea Res. Oceanogr. Res. Papers* 51, 1759–1768. doi: 10.1016/j.dsr.2004.08.002

Raghukumar, S. (2017). *Fungi in coastal and oceanic marine ecosystems: Marine fungi*. New York, NY: Springer, doi: 10.1007/978-3-319-54304-8

Raghukumar, C., Damare, S. R., and Singh, P. (2010). A review on deep-sea fungi: occurrence, diversity and adaptations. *Bot. Mar.* 53, 479–492. doi: 10.1515/bot.2010.076

Rahman, A. K. M. M., Ahmed, K. M., Butler, A. P., and Hoque, M. A. (2018). Influence of surface geology and micro-scale land use on the shallow subsurface salinity in deltaic coastal areas: a case from southwest Bangladesh. *Environ. Earth Sci*. 77:423. doi: 10.1007/s12665-018-7594-0

Raikar, M. T., Raghukumar, S., Vani, V., David, J. J., and Chandramohan, D. (2001). Thraustochytrid protists degrade hydrocarbons. *Indian J. Mar. Sci*. 30, 139–145.

Raimundo, R. L. G., Guimarães, P. R., and Evans, D. M. (2018). Adaptive Networks for Restoration Ecology. *Trends Ecol. Evol*. 33, 664–675. doi: 10.1016/j.tree.2018.06.002

Rämä, T., Hassett, B. T., and Bubnova, E. (2017). Arctic marine fungi: from filaments and flagella to operational taxonomic units and beyond. *Bot. Mar.* 60, 433–452. doi: 10.1515/bot-2016-0104

Rédou, V., Navarri, M., Meslet-Cladière, L., Barbier, G., and Burgaud, G. (2015). Species richness and adaptation of marine fungi from deep-subseafloor sediments. *Appl. Env. Microbiol.* 81, 3571–3583. doi: 10.1128/AEM.04064-14

Richards, T. A., Jones, M. D. M., Leonard, G., and Bass, D. (2012). Marine fungi: their ecology and molecular diversity. *Ann. Rev. Mar. Sci*. 35, 283–298. doi: 10.1146/annurev-marine-120710-100802

Ritchie, D. (1957). Salinity Optima for Marine Fungi Affected by Temperature. *Am. J. Bot.* 1957, 870–874. doi: 10.2307/2438907

Román, E., Correia, I., Prieto, D., Alonso, R., and Pla, J. (2020). The HOG MAPK pathway in Candida albicans: more than an osmosensing pathway. *Internat. Microbiol*. 23, 23–29. doi: 10.1007/s10123-019-00069-1

Romão, D., Staley, C., Ferreira, F., Rodrigues, R., Sabino, R., Veríssimo, C., et al. (2017). Next-generation sequencing and culture-based techniques offer complementary insights into fungi and prokaryotes in beach sands. *Mar. Pollut. Bull.* 119, 351–358. doi: 10.1016/j.marpolbul.2017.04.036

Romeo, O., Marchetta, A., Giosa, D., Giuffrè, L., Urzì, C., and De Leo, F. (2020). Whole Genome Sequencing and Comparative Genome Analysis of the Halotolerant Deep Sea Black Yeast *Hortaea werneckii*. *Life* 10:229. doi: 10.3390/life10100229

Sakayaroj, J., Pang, K. L., and Jones, E. B. G. (2011). Multi-gene phylogeny of the Halosphaeriaceae: Its ordinal status, relationships between genera and morphological character evolution. *Fungal Div*. 46, 87–109. doi: 10.1007/s13225-010-0072-y

Salma, M., Rousseaux, S., Sequeira-Le Grand, A., Divol, B., and Alexandre, H. (2013). Characterization of the viable but nonculturable (VBNC) state in *Saccharomyces cerevisiae*. *PLoS One* 8:e77600. doi: 10.1371/journal.pone.0077600

Sanderson, M. J., Thorne, J. L., Wikström, N., and Bremer, K. (2004). Molecular evidence on plant divergence times. *Am. J. Bot*. 91, 1656–1665. doi: 10.3732/ajb.91.10.1656

Sandilyan, S., and Kathiresan, K. (2012). Mangrove conservation: A global perspective. *Biodiv. Conserv.* . 21, 3523–3542. doi: 10.1007/s10531-012-0388-x

Sanka Loganathachetti, D., Poosakkannu, A., and Muthuraman, S. (2017). Fungal community assemblage of different soil compartments in mangrove ecosystem. *Sci. Rep*. 7, 1–9. doi: 10.1038/s41598-017-09281-3

Sarma, V. V., and Hyde, K. D. (2001). A review on frequently occurring fungi in mangroves. *Fungal Div*. 8, 1–34.

Sarma, V. V., and Jeewon, R. (2019). Marine fungal ecology in the molecular era. *Adv. Front. Mycol. Mycotechnol.* 2019, 143–180. doi: 10.1007/978-981-13-9349-5_6

Schabereiter-Gurtner, C., Piñar, G., Lubitz, W., and Rölleke, S. (2001). Analysis of fungal communities on historical church window glass by denaturing gradient gel electrophoresis and phylogenetic 18S rDNA sequence analysis. *J. Microbiol. Methods* 47, 345–354. doi: 10.1016/S0167-7012(01)00227-5

Schlaeppi, K., Bender, S. F., Mascher, F., Russo, G., Patrignani, A., Camenzind, T., et al. (2016). High-resolution community profiling of arbuscular mycorrhizal fungi. *New Phytolog*. 212, 780–791. doi: 10.1111/nph.14070

Schottroff, F., Fröhling, A., Zunabovic-Pichler, M., Krottenthaler, A., Schlüter, O., and Jäger, H. (2018). Sublethal injury and viable but non-culturable (VBNC) state in microorganisms during preservation of food and biological materials by non-thermal processes. *Front. Microbiol*. 9:2773. doi: 10.3389/fmicb.2018.02773

Schüßler, A. (2002). Molecular phylogeny, taxonomy, and evolution of Geosiphon pyriformis and arbuscular mycorrhizal fungi. *Plant Soil* 244, 75–83. doi: 10.1023/A:1020238728910

Sekimoto, S., Rochon, D., Long, J. E., Dee, J. M., and Berbee, M. L. (2011). A multigene phylogeny of Olpidium and its implications for early fungal evolution. *BMC Evol. Biol*. 11, 1–10. doi: 10.1186/1471-2148-11-331

Selbmann, L., Benkō, Z., Coleine, C., de Hoog, S., Donati, C., Druzhinina, I., et al. (2020). Shed light in the dark lineages of the fungal tree of life—stres. *Life* 10:362. doi: 10.3390/life10120362

Sharpe, S. C., Eme, L., Brown, M. W., and Roger, A. J. (2015). Timing the Origins of Multicellular Eukaryotes Through Phylogenomics and Relaxed Molecular Clock Analyses. *Evol. Trans. Multicell. Life* 2015, 3–29. doi: 10.1007/978-94-017-9642-2_1

Sinsabaugh, R. (2005). Fungal Enzymes at the Community Scale. *Mycol. Series* 23:349. doi: 10.1201/9781420027891.ch17

Snyder, M., Du, J., and Gerstein, M. (2010). Personal genome sequencing: current approaches and challenges. *Genes Dev*. 24, 423–431. doi: 10.1101/gad.1864110

Sridhar, K. R., and Maria, G. L. (2006). Fungal diversity on mangrove woody litter Rhizophora mucronata (Rhizophoraceae). *Indian J. Mar. Sci*. 35, 318–325.

Stat, M., Huggett, M. J., Bernasconi, R., Dibattista, J. D., Berry, T. E., Newman, S. J., et al. (2017). Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Sci. Rep*. 7:12240. doi: 10.1038/s41598-017-12501-5

Stefani, F. O., Bell, T. H., Marchand, C., de la Providencia, I. E., El Yassimi, A., St-Arnaud, M., et al. (2015). Culture-dependent and-independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. *PLoS One* 10:e0128272. doi: 10.1371/journal.pone.0128272

Stocker, T. (ed.) (2014). *Climate Change 2013: the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge university press.

Suetrong, S., Schoch, C. L., Spatafora, J. W., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Sakayaroj, J., et al. (2009). Molecular systematics of the marine Dothideomycetes. *Stud. Mycol*. 64, 155–173. doi: 10.3114/sim.2009.64.09

Tedersoo, L., Bahram, M., Puusepp, R., Nilsson, R. H., and James, T. Y. (2017). Novel soil-inhabiting clades fill gaps in the fungal tree of life. *Microbiome* 5, 1–10. doi: 10.1186/s40168-017-0259-5

Tedersoo, L., Sánchez-Ramírez, S., Kõljalg, U., Bahram, M., Döring, M., Schigel, D., et al. (2018). High-level classification of the Fungi and a tool for evolutionary ecological analyses. *Fungal Div*. 90, 135–159. doi: 10.1007/s13225-018-0401-0

Thines, M., Crous, P. W., Aime, M. C., Aoki, T., Cai, L., Hyde, K. D., et al. (2018). Ten reasons why a sequence-based nomenclature is not useful for fungi anytime soon. *IMA Fungus* 9, 177–183. doi: 10.5598/IMAFUNGUS.2018.09.01.11

Tisthammer, K. H., Cobian, G. M., and Amend, A. S. (2016). Global biogeography of marine fungi is shaped by the environment. *Fungal Ecol*. 19, 39–46. doi: 10.1016/j.funeco.2015.09.003

Tu, Q., Yuan, M., He, Z., Deng, Y., Xue, K., Wu, L., et al. (2015). Fungal communities respond to long-term CO2 elevation by community reassembly. *Appl. Env. Microbiol*. 81:2445. doi: 10.1128/AEM.04040-14

Turk, M., and Plemenitaš, A. (2002). The HOG pathway in the halophilic black yeast Hortaea werneckii: Isolation of the HOG1 homolog gene and activation of HwHog1p. *FEMS Microbiol. Lett*. 216, 193–199. doi: 10.1016/S0378-1097(02)01033-9

Ullah, A., Bano, A., and Khan, N. (2021). Climate Change and Salinity Effects on Crops and Chemical Communication Between Plants and Plant Growth-Promoting Microorganisms Under Stress. *Front. Sustain. Food Syst.* 5:618092. doi: 10.3389/fsufs.2021.618092

Venkatachalam, M., Gérard, L., Milhau, C., Vinale, F., Dufossé, L., and Fouillaud, M. (2019). Salinity and temperature influence growth and pigment production in the marine-derived fungal strain talaromyces albobiverticillius 30548. *Microorganisms* 7:10. doi: 10.3390/microorganisms7010010

Vijaykrishna, D., Jeewon, R., and Hyde, K. D. (2006). Molecular taxonomy, origins and evolution of freshwater ascomycetes. *Fungal Div*. 23, 351–390.

Woodruff, J. D., Irish, J. L., and Camargo, S. J. (2013). Coastal flooding by tropical cyclones and sea-level rise. *Nature* 504, 44–52. doi: 10.1038/nature12855

Xu, W., Guo, S., Pang, K. L., and Luo, Z. H. (2017). Fungi associated with chimney and sulfide samples from a South Mid-Atlantic Ridge hydrothermal site: Distribution, diversity and abundance. *Deep-Sea Res. : Oceanogr. Res. Papers* 123, 48–55. doi: 10.1016/j.dsr.2017.03.004

Xu, W., Pang, K. L., and Luo, Z. H. (2014). High Fungal Diversity and Abundance Recovered in the Deep-Sea Sediments of the Pacific Ocean. *Microb. Ecol*. 68, 688–698. doi: 10.1007/s00248-014-0448-8

Zhang, X. Y., Tang, G. L., Xu, X. Y., Nong, X. H., and Qi, S. H. (2014). Insights into deep-sea sediment fungal communities from the East Indian ocean using targeted environmental sequencing combined with traditional cultivation. *PLoS One* 8:e82131. doi: 10.1371/journal.pone.0109118

Zhang, T., Wang, N. F., Zhang, Y. Q., Liu, H. Y., and Yu, L. Y. (2015). Diversity and distribution of fungal communities in the marine sedi-ments of Kongsfjorden, Svalbard (high arctic). *Sci. Rep*. 5:14524. doi: 10.1038/srep14524

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 September 2021
doi: 10.3389/fmicb.2021.717176

Comparative Genome Analysis and Spore Heat Resistance Assay Reveal a New Component to Population Structure and Genome Epidemiology Within *Clostridium perfringens* Enterotoxin-Carrying Isolates

Kaisa Jaakkola[1], Kira Virtanen[2,3], Päivi Lahti[4], Riikka Keto-Timonen[1], Miia Lindström[1] and Hannu Korkeala[1]*

[1] Department of Food Hygiene and Environmental Health, University of Helsinki, Helsinki, Finland, [2] Department of Bacteriology and Immunology, Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland, [3] Northern Finland Laboratory Centre NordLab, Oulu, Finland, [4] City of Helsinki, Unit of Environmental Services, Helsinki, Finland

OPEN ACCESS

Edited by:
Maria Elena Martino, University of Padua, Italy

Reviewed by:
Jihong Li, University of Pittsburgh, United States
Marjorie Pizarro-Guajardo, Texas A&M University, United States

***Correspondence:**
Kaisa Jaakkola
kaisa.w.jaakkola@helsinki.fi

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: *30 May 2021*
Accepted: *10 August 2021*
Published: *08 September 2021*

Citation:
Jaakkola K, Virtanen K, Lahti P, Keto-Timonen R, Lindström M and Korkeala H (2021) Comparative Genome Analysis and Spore Heat Resistance Assay Reveal a New Component to Population Structure and Genome Epidemiology Within Clostridium perfringens Enterotoxin-Carrying Isolates. Front. Microbiol. 12:717176. doi: 10.3389/fmicb.2021.717176

Clostridium perfringens causes a variety of human and animal enteric diseases including food poisoning, antibiotic-associated diarrhea, and necrotic enteritis. Yet, the reservoirs of enteropathogenic enterotoxin-producing strains remain unknown. We conducted a genomic comparison of 290 strains and a heat resistance phenotyping of 30 *C. perfringens* strains to elucidate the population structure and ecology of this pathogen. *C. perfringens* genomes shared a conserved genetic backbone with more than half of the genes of an average genome conserved in >95% of strains. The *cpe*-carrying isolates were found to share genetic context: the *cpe*-carrying plasmids had different distribution patterns within the genetic lineages and the estimated pan genome of *cpe*-carrying isolates had a larger core genome and a smaller accessory genome compared to that of 290 strains. We characterize *cpe*-negative strains related to chromosomal *cpe*-carrying strains elucidating the origin of these strains and disclose two distinct groups of chromosomal *cpe*-carrying strains with different virulence characteristics, spore heat resistance properties, and, presumably, ecological niche. Finally, an antibiotic-associated diarrhea isolate carrying two copies of the enterotoxin *cpe* gene and the associated genetic lineage with the potential for the emergence of similar strains are outlined. With *C. perfringens* as an example, implications of input genome quality for pan genome analysis are discussed. Our study furthers the understanding of genome epidemiology and population structure of enteropathogenic *C. perfringens* and brings new insight into this important pathogen and its reservoirs.

Keywords: *Clostridium perfringens*, core genome, enterotoxin (CPE), spore heat resistance, gastrointestinal disease, comparative genome analysis

INTRODUCTION

Clostridium perfringens is an anaerobic, spore former causing gas gangrene, wound infections, and a variety of human and animal diseases involving the gastrointestinal (GI) system (Kiu and Hall, 2018). Based on the recent estimation, this pathogen originated 40–80 000 years ago, and strains similar to the current isolates have been isolated from a 12 000-year-old mummified puppy and a 5,000-year-old mummified human corpse (Lugli et al., 2017; Feng et al., 2020). Presently, *C. perfringens* is a common human enteric pathogen: approximately 15% of reported antibiotic-associated diarrhea (AAD) cases are caused by *C. perfringens* (Asha et al., 2006), and the amount of *C. perfringens* food-poisoning cases in the European Union has been projected to be between 850 000 and 5 million cases per year (Silva et al., 2015).

Genome sequencing and comparative analysis have shown that *C. perfringens* has a large accessory genome reflecting the wide variety of caused diseases and produced toxins (Petit et al., 1999; Rood et al., 2018). Recently, five phylogenetic lineages (I–V) of *C. perfringens* have been identified (Kiu et al., 2017; Feng et al., 2020; Abdel-Glil et al., 2021). There is currently no consensus on naming these lineages, and different nomenclature has been used in previous studies. Feng et al. (2020) estimated that the core genome (genes present in 95% of genomes) shared by these five lineages comprised only one-third of the full genome of each lineage, emphasizing the size and diversity of this species' accessory genomes.

Identified genetic lineages are not yet established in use for the classification of *C. perfringens* strains. Instead, five toxins are used to toxinotype strains to types A–G (Rood et al., 2018). Strains of different toxinotypes are associated with different host preferences and cause different diseases: for example, toxinotype B strains are associated with sheep dysentery and toxinotype G strains carrying NetB toxin with necrotic enteritis in chickens (Rood et al., 2018). Both food-poisoning and AAD cases of *C. perfringens* diarrhea are primarily caused by type F strains (previously *cpe*-carrying type A strains), which produce pore-forming *C. perfringens* enterotoxin (CPE) but not β-toxin, ε-toxin, or ι-toxin (Rood et al., 2018). The sole chromosomal gene in the toxinotyping scheme is alpha (*plc*), present in all *C. perfringens* strains, while the other toxinotyping toxins are carried on transposable elements or a family of conjugative plasmids (Li et al., 2013). Toxinotyping, therefore, does not reflect the phylogenetic lineage of strains. Recently, a virulence gene profile scheme (I–XV), including chromosomal genes for typing of foodborne isolates (Abdelrahim et al., 2019), has been proposed.

The current understanding is that all *C. perfringens* types can carry the *cpe* gene and produce CPE, but that only approximately 5% of the strains are *cpe*-positive (Miyamoto et al., 2006). The 319 aa *cpe* gene is located either in pCPF5603 or pCPF4969 plasmid (plasmid-mediated *cpe*, p-cpe strains) or on a transposable element Tn*5565* integrated into the chromosome (chromosomal *cpe*, c-cpe strains) (Cornillot et al., 1995; Brynestad et al., 1997; Miyamoto et al., 2006; Li et al., 2010). Foodborne outbreaks are caused by both p-cpe and c-cpe strains (Lahti et al., 2008). Additionally, a 325 aa variant of the *cpe* gene of unknown clinical relevance has been described in the pCPBB1 plasmid (Miyamoto et al., 2011). Known *cpe*-carrying plasmids are all conjugative and horizontally transferable (Miyamoto et al., 2006).

The current typing practices do not reveal the genetic lineage, and likewise, the source of enteric *C. perfringens* infection is often not identified. Transmission routes and ecology of *C. perfringens* are poorly understood, and the reservoirs of enteropathogenic strains remain unknown. *C. perfringens* is known as an environmental bacteria associated with soil, water, sewage, and dust, but also the GI tract of humans and animals (Hatheway, 1990). Suggested reservoirs for *cpe*-positive strains include healthy animals and humans, sludge, soil, and retail meat (Wen and McClane, 2004; Heikinheimo et al., 2006; Li et al., 2007; Lindström et al., 2011; Li et al., 2013; Lahti et al., 2012; Hu et al., 2018). C-cpe strains have been solely isolated from food items or food poisoning associated samples, while p-cpe strains are often associated with non-food-borne human disease and animal enteric disease (Collie and McClane, 1998; Sparks et al., 2001). Comparative genome analysis results have suggested that c-cpe strains are clonal (Deguchi et al., 2009; Xiao et al., 2012) and specialized to yet unknown environmental niche. The GI tract has been suggested to be the adapted niche for p-cpe strains (Lahti et al., 2012). Particularly, pCPF5603 plasmid-carrying strains are associated with care home outbreaks and AAD (Miyamoto et al., 2006; Kiu et al., 2019). A recent study on human enteric isolates suggested that *C. perfringens* strains from a single source might also persist and cause longitudinal outbreaks spanning several years (Kiu et al., 2019), supporting the hypothesis of healthy humans serving as reservoirs for *cpe*-carrying strains (Heikinheimo et al., 2006).

Some *C. perfringens* strains produce heat-resistant spores which enable them to proliferate in foods after cooking if the cold chain is compromised (Sarker et al., 2000; Li and McClane, 2006b; Lindström et al., 2011). The heat resistance of *C. perfringens* spores has been studied widely (Weiss and Strong, 1967; Ando et al., 1985; Sarker et al., 2000; Raju and Sarker, 2005; Li and McClane, 2006a,b, 2008; Xiao et al., 2015), and the main finding is that c-cpe strains produce heat-resistant spores with the $D_{99°C}$ value average of 53.2 min (Weiss and Strong, 1967; Sarker et al., 2000). The known key mechanism for this spore heat resistance phenotype is the ability of strains to produce a variant of small, acid-soluble protein (Ssp4), which binds strongly to the spore DNA to protect it (Li et al., 2009). Contrarily, p-cpe strains and cpe negative strains produce heat-sensitive spores with low $D_{99°C}$ values (on average 1.0 min) (Sarker et al., 2000). The link between genotype and phenotype has not been studied previously, and only three strains with known D values have been sequenced. The growth temperatures of different strains have not been widely studied.

To elucidate the ecology and reservoirs of *cpe*-carrying *C. perfringens* strains, we determined the heat resistance phenotypes and growth temperatures for 30 sequenced *C. perfringens* strains. To understand the population structure and genetic context of this pathogen, the pangenome, cgMLST typing, and phylogenetic tree of 290 *C. perfringens* strains

were created. Our results suggest that the presence of the *cpe* gene is associated with certain genetic backgrounds instead of just horizontally transferable enterotoxin. We also describe two distinct groups of chromosomal *cpe*-carrying strains with different virulence characteristics, spore heat resistance properties, and, presumably, ecological niche. The importance of input genome quality for pangenome analysis is discussed as a larger core genome, and a smaller pangenome of *C. perfringens* compared to recent estimations is presented.

MATERIALS AND METHODS

Bacterial Strains and Genomes

A total of 30 *C. perfringens* strains isolated from various sources during 1984–2007 were sequenced and used for heat resistance and growth temperature assays (**Supplementary Table 1**). The strains included food poisoning associated isolates from foods (n = 8) or feces (n = 6); human cases of antibiotic-associated diarrhea (n = 3); feces of healthy people (n = 8); and feces of healthy production animals (n = 2), soil (n = 1), and sludge (n = 2). The strains, their isolation source, and type of *cpe* locus have been previously reported (Heikinheimo et al., 2006; Lahti et al., 2008, 2012). Twenty-seven isolates carried the *cpe* gene, and three were cpe-negative. The *cpe* gene was present in the chromosome of 11 strains (c-cpe) and a plasmid in 16 strains (p-cpe). The *cpe* genetic location on p-cpe strains was either on pCPF4969/IS1470-like (n = 9) or pCPF5603/IS1151 (n = 8) plasmid (Heikinheimo et al., 2006; Lahti et al., 2008 2012).

The *cpe*-negative *C. perfringens* strains ATCC 13124 and str. 13, and c-cpe strain SM101 were used as reference genomes (Shimizu et al., 2002; Myers et al., 2006). For phylogenetic and pangenome analyses also publicly available *C. perfringens* genomes (n = 260), representing all described genetic lineages, and including 56 *cpe*-carrying isolates, were included (**Supplementary Table 2**). Genomes were downloaded from the Patric database (Gillespie et al., 2011).

Genome Sequencing and Annotation

Genomic DNA of 30 *C. perfringens* strains was extracted (Keto-Timonen et al., 2006), and whole-genome sequencing was performed using PacBio RSII (Institute of Biotechnology, Helsinki, Finland). Sequenced genomes were assembled using HGAP3 and checked for circularity using Gap4 (Staden et al., 2003; Chin et al., 2013). To improve the draft assembly, Illumina MiSeq reads and Pilon tool were used for genome polishing (Walker et al., 2014). Sequenced *C. perfringens* genomes were deposited in the GenBank (accession numbers listed in **Supplementary Table 3**). Both sequenced and downloaded genomes were annotated using Prokka (Seemann, 2014). Protein functional annotation and conserved domain predictions were refined using CD-search and CDD (Li and Godzik, 2006; Marchler-Bauer et al., 2009).

cgMLST

To determine the cgMLST target gene set and to create a genome-wide gene-by-gene comparison, ChewBBACA was used (Silva et al., 2018). Both custom-made schema and schema created by Abdel-Glil (2019) available at https://www.cgmlst.org (accessed 10.01.2021) were used. The custom-made schema was created using closed *C. perfringens* genome ATCC 13124 as a training genome and selected genomes as initial query (**Supplementary Table 1**). Orthologous coding sequences (CDSs) were called using blast score ratio equal to or greater than 0.6 as the threshold. Paralogs, genes shorter than 200 bp, and genes without start or stop codons were excluded, and only the longest allele was kept in the list. Listed genes were used to call alleles from query genomes (**Supplementary Table 2**). In both called schemes, a threshold of 95 was selected for further analysis (9 genomes removed). Called alleles were once again pruned for paralogs, and the cgMLST gene list was created by filtering out alleles present in less than 95% of genomes.

Pangenome and Comparative Genome Analyses

For the core genome and pangenome analysis, Roary (Page et al., 2015) and Panaroo (Tonkin-Hill et al., 2020) pipelines were used. The genomes were submitted to the Roary pangenome pipeline v3.6.1 using parameters to split paralogs and 95% minimum percentage identity for blastp. The Panaroo pipeline was run in strict mode and using the default parameter to split paralogs. Similar parameters were used for the creation of pangenomes for 86 *cpe*-carrying strains and 63 lineage IV strains. The core genome clusters were functionally annotated with Eggnog 5.0 (Huerta-Cepas et al., 2019). Scoary was used to associate gene clusters with observed phenotypes (Brynildsrud et al., 2016).

Initially, the Roary pipeline produced a noticeably small core genome, and to produce comparable results with the Panaroo pipeline, a set of quality criteria for input genomes was applied. Genomes were required to meet two or more quality indicators to be included in pangenome analysis (Parrello et al., 2019). Used quality criteria were (1) Patric Genome Quality coarse consistency over 97.0%; (2) Patric Genome Quality fine consistency over 92.0%; (3) less than 100 alleles over 20% shorter than reference sequence in Abdel-Glil cgMLST scheme (**Supplementary Tables 2**, **10**). In total, seven genomes (BER-NE33, W1319, PBS5, PBD1, PC5, K473, and T3381) were excluded from the Roary analysis of 283 strains based on these criteria.

Additional genome comparisons were performed using bi-directional BLASTP with the standard scoring matrix BLOSUM62 and an initial E-value cut-off of $1e^{-05}$. Two genes were acknowledged orthologs if a reciprocal best blast hit existed among them, and both hits had over 90% similarity and 80% coverage. On average identified bi-directional best hits between any two strains shared over 95% identity (data not shown). When manual curation revealed putative orthologs with conserved synteny but sequence identity below 80%, a term divergent ortholog is used. Genomes were aligned and visually compared with progressive Mauve (Darling et al., 2010) and SEED Viewer 2.0 (Overbeek et al., 2014). Comparison tools available at Patric database (Gillespie et al., 2011; Wattam et al., 2016) were also used for manual curation.

Phylogenetic Analysis

Maximum likelihood phylogenetic trees were constructed using fasttree 2 (Price et al., 2010). Minimal spanning tree was calculated with Grapetree (Zhou et al., 2018) using MSTreeV3. Phylogenetic trees based on Roary calculated core genomes and cgMLST were visualized and annotated with Grapetree and iTOL v3[1].

Heat-Resistance Assay for Spores and Vegetative Cells

The heat-resistance phenotype and estimated D values of 30 *C. perfringens* spores and vegetative cells were determined according to established procedures (Weiss and Strong, 1967; Ando et al., 1985; Sarker et al., 2000). All strains were prepared as described by Duncan and Strong (1968) and Ando et al. (1985). Briefly, Duncan-Strong (DS) medium cultures grown for 24 h at 37°C were heat-shocked at 75°C for 15 min to kill the vegetative cells and facilitate spore germination. A 100 μl sample was withdrawn, diluted, and plated to determine the initial spore count. The remainder of each heat-shocked DS medium culture was then heated at either 89 or 99°C for 1 min to 4 h or until the spore count reached 0 (depending on the individual isolate and the temperature being used). Sampling occurred every 30 s for 0–5 min, then every 60 s until 10 min, and 11–240 min every 30 min. At each time point, culture was mixed and a 100 μl time point sample was withdrawn and diluted (dilution range, 10^{-2} to 10^{-7}). For vegetative cells, FTG (fluid thioglycolate) medium cultures grown for 24 h at 37°C were heated to 60°C for a time ranging from 1 min to 2 h. The time point samples were withdrawn as explained above. Finally, to determine the number of viable spores or cells present per milliliter (CFU/ml), the dilutions were plated on BHI. The logarithmic counts of viable spores and cells of every isolate were graphed against the heating time to determine the slope of the survival curve. The estimation of the D value was determined from this curve. P values were calculated using ANOVA analysis of variance (IBM SPSS Statistics for Macintosh, Version 27.0). Strains (955/85, CPI 57K-1, C269) did not produce spores with the methods used here.

Maximum and Minimum Growth Temperatures

The growth temperatures of 30 *C. perfringens* strains were examined using the Gradiplate W10 incubator (Biodata Oy, Helsinki, Finland) placed in an anaerobic workstation (MK III, Don Whitley Scientific, Ltd., Shipley, United Kingdom) to determine the minimum and maximum growth temperatures (Korkeala et al., 1990; Derman et al., 2011). Briefly, strains were cultured in FTG at 37°C for 24 h. Strains were refreshed and grown at 37°C until turbidity OD_{600} reached 0.6–0.9. At least two biological replicates were cultured and analyzed. Cultures were diluted 1:10 in peptone water, and a sample of 25 microliters was transferred to a Gradiplate cuvette. Tested temperature gradients for minimum and maximum growth temperatures were 8–18°C and 47–57°C, respectively. Strains were incubated for 2 days for maximum temperatures and 21 days for minimum temperatures. Growth boundaries were determined by using a stereomicroscope, and the growth temperature threshold was determined as the boundary where dense bacterial growth was discontinued. Minimum and maximum growth temperatures were calculated as described by Korkeala et al. (1990) and Derman et al. (2011) using the formula $T = T_{low} + d * g$, where d is the distance (in millimeters) of the growth boundary to the measurement point of T_{low} and g is the temperature gradient.

[1]http://itol.embl.de/

RESULTS

Phylogenetic Analysis Revealed Dispersal Patterns of *cpe*-Plasmids

Analyzed strains (n = 290) represented all five genetic lineages (I–V) and were isolated from a variety of animals as well as from humans and the environment (**Supplementary Tables 1**, **2**). Here, lineages are numbered based on tree topology and are consistent with that used by Feng et al. (2020).

Three of the genetic lineages were well presented in sequenced strains: III (34 strains), IV (63 strains), and V (183 strains), and all *cpe*-carrying isolates belonged to these three lineages (**Table 1**). The majority of sequenced genomes (56.9%) were clinical isolates associated with veterinary (126/290) or human disease (39/290), and only 2.8% of genomes (eight, including the three environmental samples sequenced in this study) had been isolated from an environmental source such as soil, water, and sludge (**Figure 1**). Veterinary clinical isolates were scattered across all lineages, while all environmental isolates belonged to lineage V. Considering how common the AAD caused by *C. perfringens* is, it is surprising assemblies or closed genomes of AAD strains have not been published previously. In this study, three AAD isolates were included. The c-cpe genomes and one *cpe*-negative strain sequenced in this study belonged to lineage IV (n = 11), one AAD isolate to lineage III, and the rest to lineage V (n = 18). The minimal spanning tree of *C. perfringens* strains and their *cpe* type and isolation source based on cgMLST results are represented in **Figure 1**.

TABLE 1 | Distribution of gastrointestinal (GI) disease isolates of human or animal origin and *cpe*-carrying isolates within lineages I–V.

Lineage	No. of genomes/ isolates	% of isolates associated with GI disease	% of *cpe*-carrying genomes	*cpe* types (no. of genomes)
I	8	0	0	None (0)
II	2	50	0	None (0)
III	34	88	91	III/pCPPB1 (4), III/pCPF5603 (27)
IV	63	41	49	Chromosomally inserted *cpe* (21), unconfirmed *cpe* location on a draft genome (10)
V	183	44	14	V/pCPF4969 (10), V/pCPF5603 (15)

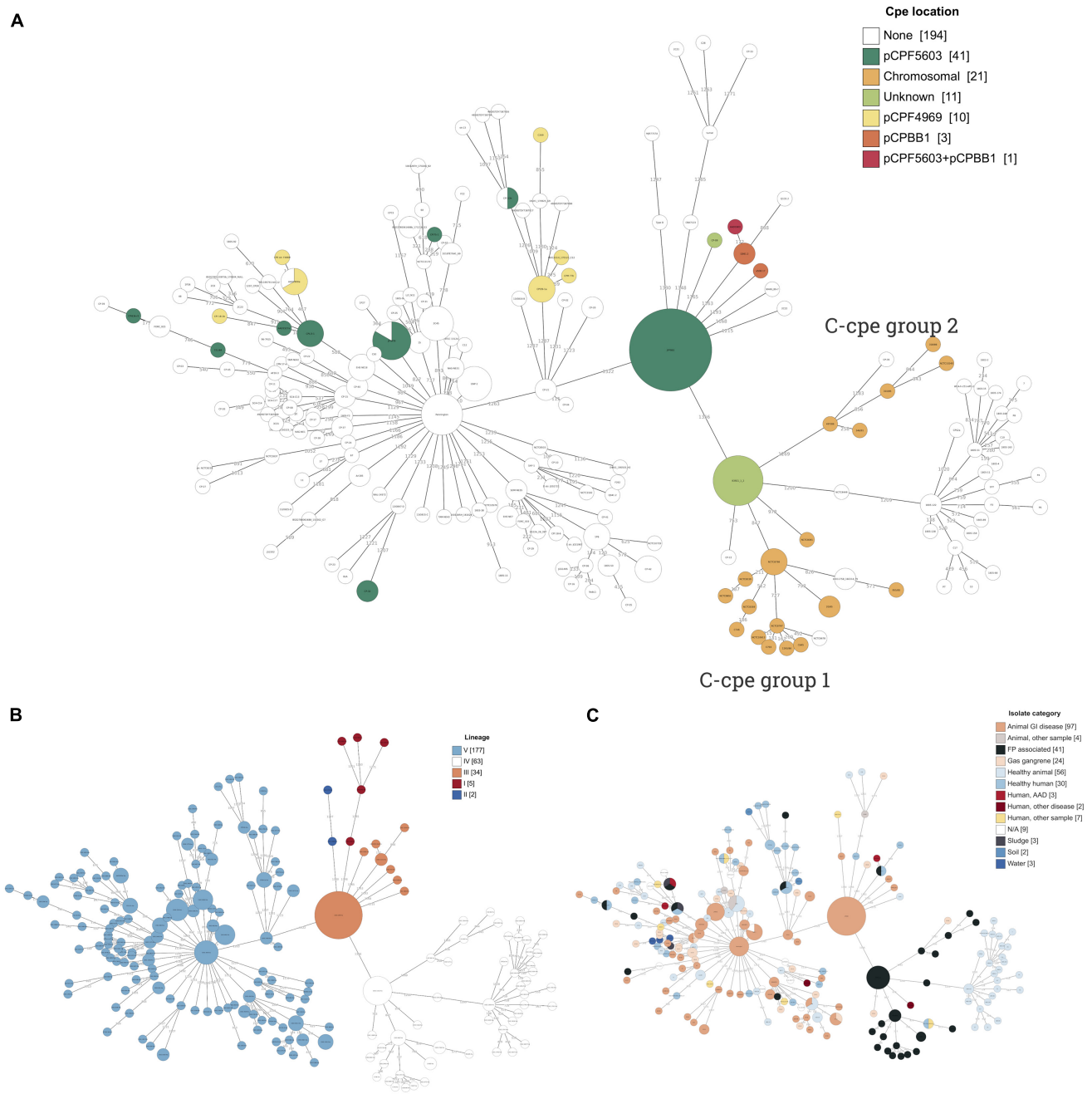


FIGURE 1 | A minimal spanning tree of *Clostridium perfringens* strains based on cgMLST (scheme by Abdel-Glil, 2019) results annotated with **(A)** *cpe* location, **(B)** lineages I–V (Feng et al., 2020), and **(C)** isolate category based on their isolation source. Branches with a distance of less than 100 are collapsed. Plasmid pCPF5603 was widely distributed phylogenetically and present in two lineages, III and V. Contrastingly, pCPF4969 plasmids were present in two branches of lineage V. Lineage IV included 32 *cpe*-negative strains together with the food poisoning-related isolates with *cpe* gene inserted in their chromosome.

C-cpe strains (n = 31) all belonged to lineage IV. Phylogenetic analysis further revealed two subgroups of c-cpe strains (from now on referred to as the c-cpe group 1 and 2) and a closely related subgroup of *cpe*-negative isolates. In total, lineage IV included 32 (51%) *cpe*-negative isolates isolated from a variety of sources and geographic locations. These included 26 isolates from healthy swine and chicken (including the here sequenced European swine isolate 2a), stool of a healthy human (North America), two cases of gas gangrene in animals (Asia), and two food poisoning isolates (Europe). All c-cpe genomes had been isolated from food or stool samples linked to food poisonings in Europe.

Contrastingly, p-cpe strains (n = 52) belonged to lineages III and V and had been isolated from Europe, Asia, and North

America. **Figure 1A** shows the dispersal patterns of *cpe*-carrying plasmids (pCPF5603, pCP4969, and pCPBB1) within genetic lineages. Plasmid pCPF5603 (n = 42) was the most widely distributed and present in both lineages III and V. Plasmid pCPF4969 (n = 10) was present in two branches of lineage V, which were phylogenetically related with each other's and with lineage IV (**Supplementary Figure 1**). All strains carrying pCPPB1 plasmid (n = 4) with a variant of *cpe* toxin clustered together within lineage III (**Table 1** and **Figure 1A**). The presence of toxin plasmids in other sequenced bacteria was looked up, but all three toxin plasmids were only found from *C. perfringens* genomes.

Core Genome and Pangenome of *C. perfringens*

The core genome of the 290 strains consisted of, depending on the clustering method used, 1170–1660 coding sequences (CDSs), and the pangenome amounted to 14 306–23 148 CDSs (**Table 2**). The pangenome was open, and after 270 genomes had been added, the pangenome still expanded at a pace of approximately 10 new genes per added genome. The number of unique CDSs per genome varied from 133 (CP-21) to 0 (several genomes). The core genome presented 7–12% of the pangenome and 38–53% of the CDSs of an average genome (3120 CDSs).

The pangenomes of *cpe*-carrying strains (n = 86) and 290 strains were compared (**Figure 2** and **Table 2**). For the *cpe*-carrying strains, the core genome covered 51–62% (1577-1941/3120) of the average genome CDSs and the pangenome amounted to 7721 CDSs. The core genome of *cpe*-carrying genomes was 17% (core genome present in 99% of strains) or 88% (core genome present in 100% of strains) larger compared to the core genome of all strains. The accessory genome of analyzed *cpe*-carrying genomes was also remarkably constrained (5780 CDSs) compared to that of all strains (12 646 CDSs). The most frequent gene cluster of the *cpe*-carrying genomes was the core genome (25%) (**Figure 2**).

Most of the genes (94%, 1044/1114) associated with *cpe*-carrying ($p < 0.001$, Benjamini-Hochberg adjusted) represented the accessory genome rather than shared characteristics between

TABLE 2 | Core genome and pangenome estimations of *Clostridium perfringens*.

Study	No. of input strains (subgroup)	Pangenome Pipeline and chosen parameters	No. of core CDSs per core threshold Present in% of genomes			No. of CDSs in the pangenome	The proportion (%) of core genome (99%) of the pangenome and the average genome[b]
			100	99	95[a]		
This study	290	Panaroo, strict	1034	1660	2059	14306	11.6, 53.2
This study	290	Roary, 95% identity, split paralogs	231	590	1078	23148	2.5, 18.9
This study	283[c]	Roary, 95% identity, split paralogs	696	1170	1581	16875	6.9, 37.5
This study	86 (*cpe*-carrying strains)	Roary, 95% identity, split paralogs	1577	1577	1797	7835	20.1, 50.7
This study	86 (*cpe*-carrying strains)	Panaroo, strict	1941	1941	2101	7721	25.1, 63.2
This study	63 (lineage IV strains)	Panaroo, strict	1840	1840	2090	5917	31.0, 65.4
Kiu et al., 2019	110 (human enteric isolates)	Roary, 90%, don't split paralogs	N/A	1965	N/A	∼6300	31.2, 63.2
Kiu et al., 2017	56	Roary, 95% identity, split paralogs	1470	1470	N/A	11667	12.6, 47.3
Abdel-Glil, 2019	76	Roary, 90%, split paralogs	N/A	N/A	2057	10098	20.3, 66.1[d]
Feng et al., 2020	173	Roary, parameters not specified	N/A	N/A	1020	26954	3.8, 32.8[d]

[a] *To enable comparison with previous results the core genome present in 95% of strains is given. Due to the concerns explained in the text, these should not be considered biologically relevant*

[b] *The average CDSs of analyzed genomes: 3120 CDSs for 290 strains, 3,071 for cpe-carrying strains, and 2814 for lineage IV strains.*

[c] *Confounding strains BER-NE33, K473, T3381, W1319, PC5, PBD1, and PBS5 were excluded from the analysis of 283 strains.*

[d] *95% core threshold was used instead of 99%.*

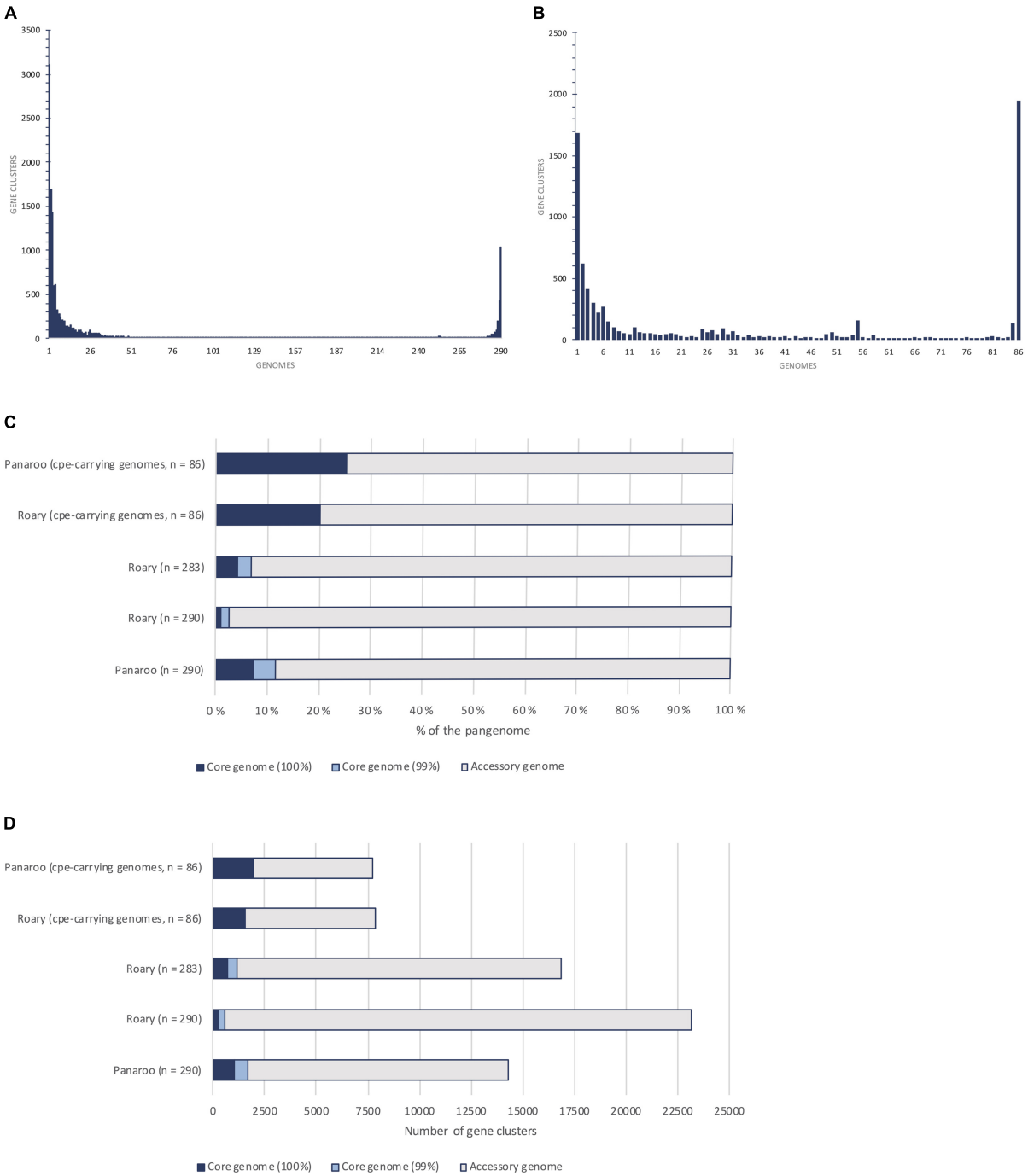


FIGURE 2 | Pangenome and core genome comparison of *Clostridium perfringens* using Roary and Panaroo pipelines. **(A,B)** show the frequency of gene clusters within the pangenome, i.e., how many genomes were the gene clusters present in. **(A)** shows the gene cluster size distribution within the pangenome of 290 strains and **(B)** the gene distribution within the pangenome of 86 *cpe*-carrying isolates. Both **(A,B)** are based on Panaroo results. Chart **(C,D)** visualize the proportion of the predicted core genome to the accessory genome in different analyses based on the percentage **(C)** and the actual number of genes **(D)**.

all *cpe*-carrying strains. Functional comparison of core genomes of 290 strains and *cpe*-carrying strains still revealed that the latter group shared more genes in 18 out of 24 clusters of orthologous groups (COGs). Especially genes belonging to COGs linked with amino acid metabolism and transport (n = 12), transcription (n = 18), signal transduction (n = 10), and unknown (n = 38) functions were more conserved in the core genome of *cpe*-carrying strains than the core genome of all *C. perfringens* strains. Operons related to riboflavin biosynthesis and sucrose metabolism, and three genes

associated with drug resistance, bacitracin resistance gene *uppP* (CPR_1021), *mepA* (CPR_1218), and vancomycin resistance gene *yoaR* (CPR_1209), were fixed within the core genome of *cpe*-carrying strains, but not within the core genome (99%) of all *C. perfringens* strains.

The Panaroo pipeline produced a larger core genome with a smaller pangenome compared to the Roary pipeline in all settings (**Figure 2** and **Table 2**). Based on manual curation of results, the presence of split CDSs or significantly truncated alleles generated most of this difference. Based on the conserved synteny the Panaroo clustered the broken sequences together with the conserved alleles, resulting in higher conservation counts for genes. In the clustering step for the Roary pipeline, the highly identical split CDSs fractionated the entire protein cluster to two or more separate clusters and consequently dropped these outside both the core genome of 99% strains as well as the soft-core present in 95% strains. Due to this, Roary did not return a reliable count of soft-core genes (present in 95% of genomes) for *C. perfringens* and stringent 99% or 100% core gene estimations were considered instead of 95% results (**Table 2**). Seven of the genomes had a notably high number of split reading frames/truncated alleles, and once these were removed from the analysis, the Roary pipeline produced a 98% larger core genome (83 CDSs per removed genome) (**Figure 2** and **Table 2**).

Conserved *cpe* Gene and a *C. perfringens* Strain With Two Plasmidial Copies of *cpe*

All but one of the strains (82/83) with a 319 aa *cpe* carried an identical copy of the enterotoxin gene either chromosomally or on a pCPF5603 or pCPF4969 plasmid. A single nucleotide mutation or sequencing error resulting in one amino acid change was observed in a previously sequenced strain (JFP981). Additionally, in all assembled c-cpe genomes (n = 21), the *cpe* gene was inserted between uracil-xanthine permease and *nadC* genes together with mobile elements, a transposase, and a hypothetical protein (75 aa) (Brynestad et al., 1997). Our study revealed no deletion scars of the *cpe*-carrying mobile element in the corresponding intergenic region of *cpe*-negative lineage IV strains.

In addition to 83 genomes carrying the 319 aa *cpe* gene, 5 genomes carried a 325 aa variant of the *cpe* gene. Noteworthily, one sequenced strain (AAD1903) contained both the 319 aa *cpe* gene and a 325 aa variant of the *cpe* gene. The 319 aa *cpe* was harbored on a pCPF5603 plasmid and the variant *cpe* on a pCPPB1 plasmid. Genetic regions surrounding the enterotoxin genes on plasmids pCPF5603 and pCPPB1 have been previously described (Miyamoto et al., 2011; Li et al., 2013) and are not repeated here. In addition to AAD1903, strains Q061.2 (pCPPB1), CP-09, a508.17 (pCPPB1), and a515.17 (pCPPB1) also harbored a 325 aa variant of *cpe* but did not carry the 319 aa *cpe* gene.

All five strains with the variant *cpe* clustered together in phylogenetic analysis (**Figure 1A**) and belonged to lineage III. Lineage III included 27 veterinary clinical isolates, mainly from necrotizing enteritis cases from foals and dogs from North America (n = 24), in addition to six other isolates (China, Great Britain, Finland, and France). Lineage III strains had larger genomes (p-value <0.001, ANOVA), lower GC% (p-value <0.001, ANOVA), and more CDSs (p-value <0.01, ANOVA) compared to lineages IV and V. All lineage III strains carried two copies of perfringolysin (PfoA). This variant of *pfoA* has an 85.7% nucleotide identity to the typical *pfoA* (Abdel-Glil et al., 2021).

Lineage IV Showed Signs of Reductive Evolution, and the Strains Represented Two Distinct Gene Profiles

Lineage IV strains (n = 63) had a smaller genome (3.04 Mb, p-value <0.001, ANOVA) with fewer coding sequences (2814 ± 160 genes, p-value <0.001, ANOVA) than other lineages (n = 227, 3.43 Mb, 3204 ± 305 CDSs). Operons and genes absent from lineage IV were common in other lineages, suggesting that the ancestors of c-cpe strains carried these genetic characteristics and the genes have been lost on the course of evolution. Mobile genetic elements were more prevalent in lineage IV as annotation of 30 here sequenced strains revealed that lineage IV strains carried on average 99 mobile elements against the 27 found in p-cpe and related strains ($p < 0.001$, ANOVA), 172 repeat elements against 87 ($p < 0.001$, ANOVA), and 17 transposon sequences against 6 ($p < 0.001$, ANOVA) (**Supplementary Table 3**).

Comparative genome analysis showed that c-cpe strains carried fewer virulence genes than other strains. Within lineage V and III, the chromosomal virulence gene profiles were similar (**Figure 3**), but two groups of c-cpe strains (c-cpe group 1 and 2) had distinguishable gene profiles different from other lineages and the other c-cpe group. C-cpe group 2 encoded a more diverse virulence gene profile than c-cpe group 1 (**Figure 3**). Sialidase gene *nanJ*, μ-toxin homolog *nagJ*, glycoside hydrolase endo-β-N-acetylglucosaminidase, and hyaluronidase *nagL* are present in c-cpe group 2 but absent from c-cpe group 1. These genes were located chromosomally, and the genetic context around these genes was otherwise conserved in c-cpe and p-cpe strains, suggesting these genes had been deleted from c-cpe group 1.

In addition to differences in known virulence factors, c-cpe group 1 and 2 genomes were distinguishable by several genes and operons with predicted functions. The comparison revealed that c-cpe group 2 had lost, amongst others, arginine deiminase pathway operon, arsenic resistance operon, citrate metabolism operon, a paralog of iron uptake system FeoAB, iron transporting operon Fhu, and one homolog of [FeFe]-hydrogenase (**Figure 3**). C-cpe group 1 strains, on the other hand, had lost the above-mentioned virulence genes and also enzymes, fucose utilization operon, and carried a broken version of DNA repair mechanism (AP endonuclease family 2 protein) (**Figure 3**).

Comparative genome analysis by Lahti et al. (2012) revealed that c-cpe strains do not carry myo-inositol and ethanolamine utilization operon and also lack an operon for biotin synthesis. Our results add to this list a cold shock protein (CspLA, CPF_1452) and sigma V (SigV, CPF_0539) factor. SigV was absent from 25/31 c-cpe strains and 100% conserved in p-cpe

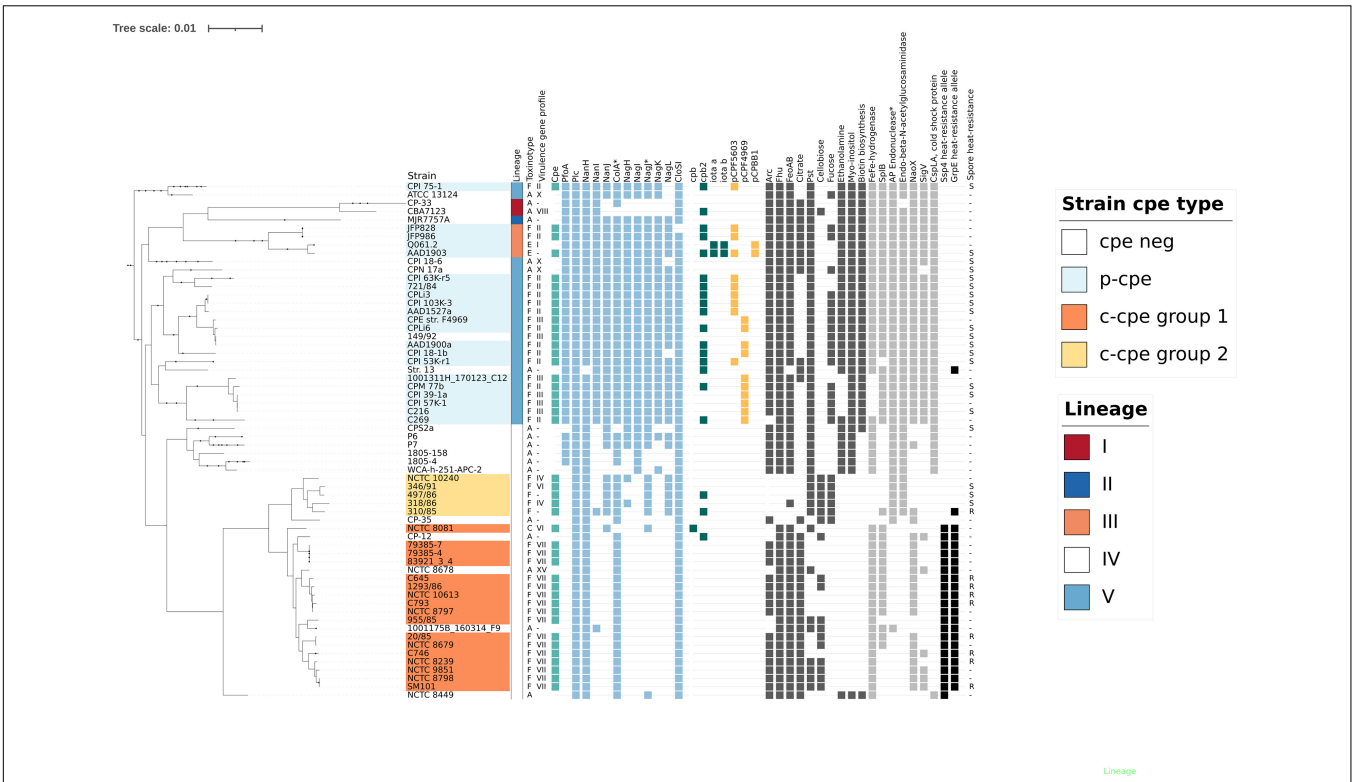


FIGURE 3 | Lineages (I–V) (Feng et al., 2020), toxinotype **(A–F)** (Rood et al., 2018), virulence gene profile (I–XV) (Abdelrahim et al., 2019), together with selected genetic features (cpe, chromosomal, and plasmid-mediated virulence genes, toxin plasmids, operons, and genes differing between c-cpe strains) are visualized here to demonstrate the results of comparative genome analysis. The maximum-likelihood phylogenetic tree is based on the core genome of 290 strains, and sequenced 30 *Clostridium perfringens* genomes are shown together with representatives from each lineage. Squares signify the presence of operons or genes in each strain. The spore heat-resistance phenotype based on our results is indicated: "R" signifies heat-resistant phenotype ($D_{89°C}$ over 100 min), "S" heat-sensitive phenotype ($D_{89°C}$ 10 min or less), and "-" that spore heat-resistance phenotype for the strain is unknown. Arc = arginine deiminase pathway, Fhu = ferrichrome uptake system, FeoAB = iron uptake system, Citrate = citrate metabolism, Pst = phosphate ABC-transporter, Cellobiose = cellobiose metabolism operon, Fucose = Fucose metabolism operon, Ethanolamine = ethanolamine utilization operon, Myo-inositol = myo-inositol utilization operon, Biotin biosynthesis. **Genes** (light gray): SplB – spore photoproduct lyase, NaoX – CoA-disulfide reductase, SigV – sigma V. **Alleles and SNPs** (black): Ssp4 heat resistance allele = gene allele with Asp at residue 36 and Asn at residue 72 (J. Li and McClane, 2008), GrpE allele (CPF_0240) putatively associated with spore heat resistance (CPF_0240) with hydrophobic valine (V) at residue 37. Plasmid-mediated toxin genes *netB*, *etx*, *tpeL*, or *lam* were not present in any genomes presented here and have been omitted from this figure.
* = Some genomes carried divergent ortholog, gene considered absent.

strains (52/52). CspLA was conserved in 100% of p-cpe strains but absent from all c-cpe strains. Both c-cpe group 1 and group 2 carried very few phylogroup-specific accessory genes, 30 and 0, respectively. Genes unique to, and conserved within, c-cpe strains included 106 coding sequences, of which 26 were mobile elements or transposases, 49 hypothetical proteins, and 12 broken or truncated copies of genes (**Supplementary Table 4**). The most notable set of specific genes for lineage IV was the cellobiose utilization operon carried by 30% (19/63) of lineage IV strains and but absent from all but one lineage V and III strains. Cellobiose operon was flanked by phage and transposon-related genetic elements.

Group of Chromosomal *cpe*-Carrying Strains in Lineage IV Produced Heat-Sensitive Spores

To characterize the phenotypes of two different c-cpe gene profiles (**Figure 3**), the sequenced strains were assayed for heat resistance properties of their spores. C-cpe group 1 strains produced highly heat resistant spores with $D_{89°C}$ values of over 100 min (average 226 min) and $D_{99°C}$ values with an average of 36 min (**Figure 4**). C-cpe group 2 strains produced relatively heat-sensitive spores with $D_{89°C}$ values of 3–9 min (average without strain 310/85 5.5 min) and $D_{99°C}$ values with an average of 2.7 min. Strain 310/85 was an outlier in c-cpe group 2 producing distinctly heat resistant spores ($D_{89°C}$ value 204 min). One representative of cpe negative group 3 strains was assayed and produced heat-sensitive spores. P-cpe strains and cpe-negative strains from lineage V and III produced spores with $D_{89°C}$ values below 8 min (average 2.8 min) (**Figure 4**). Calculated D values of strains are available in **Supplementary Table 6**.

Genes Associated With the Spore Heat-Resistance Phenotype

All c-cpe group 1 genomes carried the known heat resistance allele of Ssp4 (Asp at residue 36, Asn at residue 72), and all

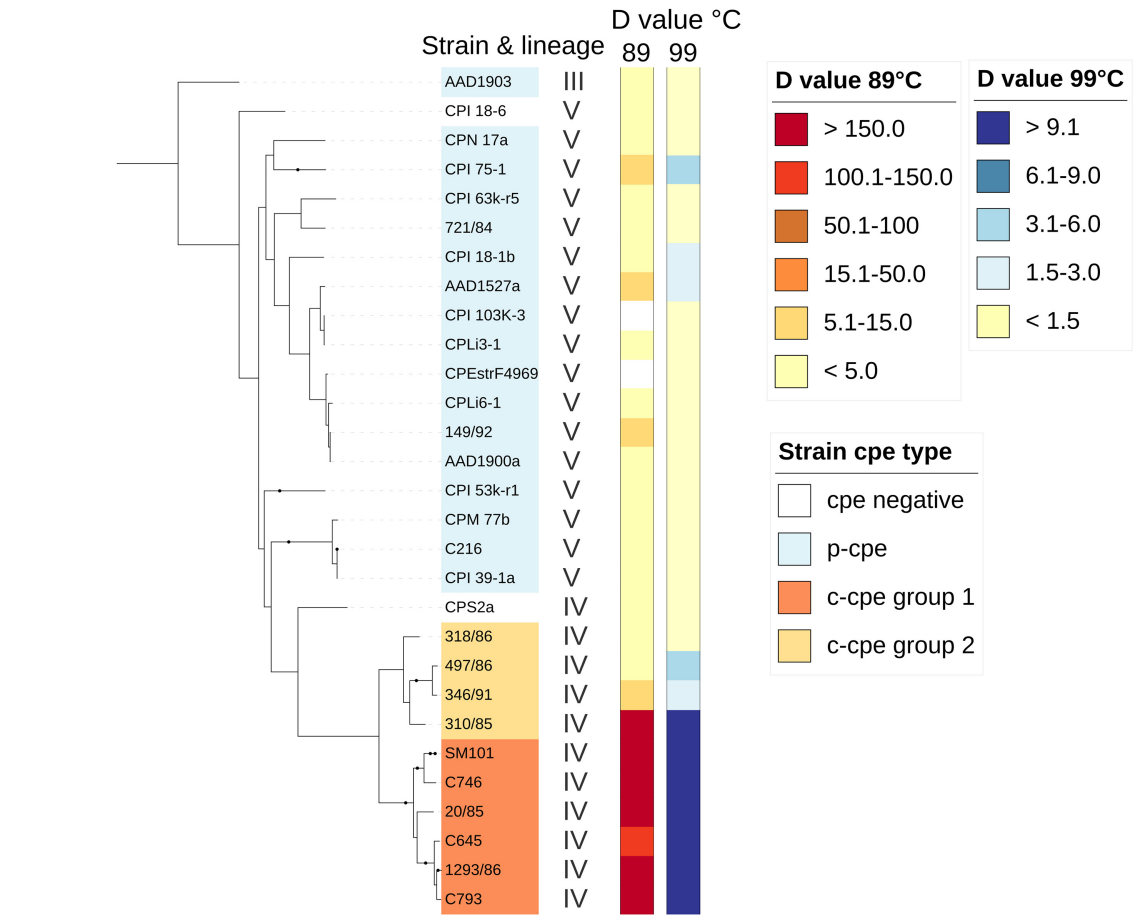


FIGURE 4 | Spore heat-resistance results of *Clostridium perfringens* strains as $D_{89°C}$ and $D_{99°C}$ values. Strains from lineages III, V, and IV were assayed. Groups of chromosomal *cpe*-carrying strains produced heat-sensitive spores (c-cpe group 2) and were phylogenetically distinct from other c-cpe strains (c-cpe group 1). Measured D values (min) for *C. perfringens* strains are shown here against a maximum-likelihood phylogenetic tree of 29 *C. perfringens* strains (27 strains sequenced and heat resistance tested in this study and two reference strains: SM101 and CPE str. F4969). For SM101 and CPE str. F4969, D values were cited from previous publications (Li and McClane, 2008; Sarker et al., 2000). Strain 310/85 was an outlier in c-cpe group 2 producing distinctly heat resistant spores ($D_{89°C}$ value 204 min).

tested c-cpe group 1 strains produced heat-resistant spores. None of the tested strains that produced heat-sensitive spores carried the heat resistance allele of Ssp4 in their genome (**Figure 3**). However, one c-cpe group 2 strain (310/85) produced highly heat-resistant spores ($D_{89°C}$ value 204 min) despite carrying the heat sensitivity associated allele (Gly instead of Asp, Lys instead of Asn) of Ssp4. It is, therefore, possible for *C. perfringens* to produce extremely heat-resistant spores without the heat resistance allele of Ssp4.

Differences between heat-sensitive and heat-resistant c-cpe strains were further analyzed. Orthologs that differed the most between heat sensitive and heat resistant c-cpe strains are summarized in **Supplementary Table 7**. Most of the sporulation-associated genes were highly conserved between strains, and the most divergent sporulation genes between heat sensitive and heat resistant c-cpe strains are listed in **Supplementary Table 8**. HGT elements were not associated with heat resistance. These comparisons revealed a cluster of 5 genes present in heat resistant strains (CPR_0713-0717) but absent from heat-sensitive c-cpe strains. Genes in this cluster encode for CoA-disulfide reductase NaoX and related permease and putative transcriptional regulator (**Figure 3**). Additionally, heat-sensitive c-cpe strains had lost a copy of GraRS two-component system, and type IV pilus operon genes *pilW*, *pilA*, and *pulG* were divergent (identity below 80%) between heat sensitive and heat resistant c-cpe strains.

Single amino acid substitutions (SNP) between c-cpe group 1 and group 2 strains were present in many sporulation and heat resistance-related genes (*ssp1*, *ssp2*, *spp3*, *clpX*, *clpC*, *clpP*, *grpE*, *groESL*, ribosome heat shock protein, *dnaJ*, *clpB*, spore coat protein S, *cotS*, *cotJC*, sporulation protein B) in addition to *ssp4*. Observed SNPs between c-cpe group 1 and 2 strains were consistent with phylogeny and did not differentiate 310/85 from the other c-cpe group 2 strains (data not shown). One

exception to this was an SNP within a heat shock protein *grpE* (CPF_0240) that had an alanine substituted with hydrophobic valine on position 37 in all strains producing heat resistant spores including outlier strain 310/85 (**Figure 3**).

Heat Resistance, Minimum, and Maximum Growth Temperatures of Vegetative Cells

Phenotypes of vegetative cells were characterized for heat resistance and growth temperatures. $D_{60°C}$ values for vegetative cells ranged from 1 to 3.8 min (average 1.3 min), with no significant differences in heat resistance between c-cpe strains and p-cpe strains, or between the two c-cpe groups (**Supplementary Table 6**). The observed minimum and maximum growth temperatures for the strains ranged from 13.1 to 17.7°C and 49.6 to 52.0°C (**Supplementary Table 6**), respectively. None of the observed differences between genetic lineages or strain Cpe types were statistically significant.

DISCUSSION

In this study, we performed phenotypic, phylogenetic, and genomic analysis of enteropathogenic *C. perfringens* and showed that plasmidial *cpe*-gene is present in strains with certain genetic backgrounds. We also identified a new group of c-cpe strains and provided insight into the reservoirs and origin of c-cpe strains.

All *C. perfringens* strains have been presumed to be able to acquire horizontally transferable enterotoxin gene and produce enterotoxin. Based on our results, the *cpe*-carrying isolates can be divided into three main groups with two subgroups each: the lineage V isolates carrying either pCPF4969 or pCPF5603 plasmids, the lineage III isolates carrying either pCPF5063 or pCPPB1, and the lineage IV isolates with c-cpe group 1 and group 2 isolates. Isolates carrying plasmids pCPF4969 and pCPPB1 were phylogenetically close to their kin. Additionally, the accessory genome within the *cpe*-carrying isolates was smaller and the core genome larger than that of all strains. These results suggest that not all *C. perfringens* strains are as likely or perhaps even able to carry enterotoxin. One possible explanation for the here observed genetic context of *cpe*-carrying could be a biased transfer rate of conjugative plasmids favoring the plasmid transfer to the kin. The bias of the transfer rate of plasmids toward related strains has been described within natural populations of *E. coli* (Dimitriu et al., 2019). Alternatively, the drug resistance genes and other genes fixed within the core genome of *cpe*-carrying isolates might reflect the capacities required for strains to flourish at the reservoirs where toxin plasmids are shared. Further studies on enterotoxin plasmids and their incompatibilities in *C. perfringens* are warranted to understand the dispersal of the *cpe* gene between different strains.

Interestingly, c-cpe group 1 and c-cpe group 2 included also *cpe*-negative isolates, and the sequence comparison revealed that all c-cpe genomes had the *cpe* gene inserted in the same genetic location. This supports the notion of the clonal nature of c-cpe strains and suggests that *cpe* insertion to chromosome precedes diversification of c-cpe group 1 and group 2 strains. The *cpe*-negative strains have either persisted together with *cpe*-positive strains or, at some point, lost the chromosomal *cpe* gene. Since chromosomal *cpe* insert is carried in *Tn5565* (Brynestad et al., 1997), the latter explanation seems more probable. Proof for deletion of a *cpe*-carrying mobile element was not observed in the region between uracil-xanthine permease and *nadC* genes in *cpe*-negative lineage IV strains, but this should be further researched. For now, we hypothesize *Tn5565* has been cleanly removed through circulation from chromosomes of *cpe*-negative c-cpe group 1 and group 2 strains.

Noteworthily, here sequenced antibiotic-associated diarrhea strain AAD1903 carried two copies of plasmid-mediated c*pe* gene, one the well-known *cpe* gene of 319 aa sequence length (in pCPF5603 plasmid) and the other a 325aa variant (in pCPPB1 plasmid). The *cpe* variant in pCPPB1 plasmid is expressed (Miyamoto et al., 2011), but the clinical relevance of this variant *cpe* gene remains unknown. The presence of both versions of *cpe* in strain AAD1903 suggests these variants of enterotoxin might have different biological roles. To our understanding, this is the first time a single *C. perfringens* strain has been reported to carry two copies of the enterotoxin gene and two different enterotoxin plasmids. As all pCPPB1 plasmids were sequenced from isolates belonging to double *pfoA*-carrying lineage III, there seems to be potential within lineage III for the emergence of other double *cpe*, double *pfoA* strains similar to the strain AAD1903 described here.

C. perfringens genomes share a conserved genetic backbone, and more than half of the genes of an average genome are present in >95% of other strains. Two different methods for core genome and pangenome estimation were used in this study to probe the pangenome of this microbe. Pseudogenes and truncated alleles seem to be a typical mode of reductive evolution for *C. perfringens*, and the presence of broken CDSs creates issues when clustering-based algorithms are used to identify orthologs. Our results show that the exclusion of confounding genomes dramatically improved the pangenome prediction results. The Panaroo pipeline analyzing synteny handled broken CDSs better in the presence of confounding strains but risked the overestimation of functional core genome as pseudogenes and significantly truncated alleles were predicted as present. Despite the emergence of pangenome pipelines such as Roary, no standardized exclusion methods for poor-quality genomes are available. Here, we utilized genome sequence quality (Parrello et al., 2019) and frequency of truncated alleles in the cgMLST schema as inclusion criteria. In particular, truncated allele counts were useful in detecting genomes with a high frequency of broken reading frames (**Supplementary Table 10**).

The phylogenetic analysis here revealed that swine and chicken farm isolates sequenced by Li et al. (2021) are closely related with c-cpe strains and with a swine isolate sequenced here. Previously, c-cpe strains have been isolated solely from food and food poisoning related human samples, and their reservoirs and origin have been unknown. These findings are not yet directly indicative of the c-cpe reservoirs, but it is clear that the *cpe*-negative swine and chicken commensal strains are the

closest known relatives of c-cpe strains. *Cpe*-negative members of lineage IV have not gathered much attention previously, and they have been rarely reported. However, the sequencing results of Li et al. (2021) suggest that lineage IV *cpe*-negative strains colonize the GI tract of domestic animals and are not rare, but rather underrepresented as they have been classified as type A strains. Sequencing of type A strains will likely help to uncover the reservoirs of c-cpe strains. Our results also show that p-cpe strains carrying pCPF4969 plasmid are phylogenetically closest to lineage IV (**Supplementary Figure 1**). Particularly, a group of lineage V strains mainly isolated from healthy humans (11/15), and carrying pCPF4969 (6/15), and having a smaller genome compared to other lineage V strains ($p < 0.05$, ANOVA) were phylogenetically close, suggesting that lineage IV and these putative members of human commensal GI population share ancestry.

Since the diversification from other lineages, lineage IV has undergone reductive evolution and niche adaptation. Supporting this observation, the c-cpe genomes have smaller genomes, fewer CDSs, few lineage-specific CDSs, and more HGT material compared to the other strains. Our results show that evolution within c-cpe strains has taken at least two directions and the two groups of c-cpe strains described here are likely adapted to different niches. C-cpe group 1 strains, but not group 2 strains, carry operons or genes linked to survival in harsh conditions and autolytic lifestyle (Arc operon, [FeFe]-hydrogenase, arsenic resistance/detoxification operon), competitive iron uptake (Fhu operon, FeoAB), and growth on citrate (Cit operon). An abundance of citrate is known to inhibit iron uptake through normal transport and instead, high-affinity siderophores such as ferrichrome (Fhu operon) are used (Pollack et al., 1970). In general, operons such as Fhu, Cit, and Arc are associated with acidophilic lifestyle, and paralogs of [FeFe]-hydrogenases are known to provide an advantage in changing pH conditions. C-cpe group 1 strains seem to be, compared to c-cpe group 2, better equipped for changing pH and sometimes acidic, high-temperature environments where iron uptake is competitive, and citrate utilization is beneficial. Previously, the reservoir for c-cpe group 1 strains has been proposed to be compost or sludge by Lahti et al. (2012). Both sludge and compost are acidic, prone to pH and temperature changes, abundant in citrate and iron sparse fitting to the profile described here. On the other hand, the loss of the aforementioned genes and operons makes it easy to speculate that pH changes and acidic stress are not characteristic of the adapted niche of c-cpe group 2.

Both c-cpe groups have lost virulence genes and genes common in intestinal bacteria and important for host colonization (Stahl et al., 2011; Becerra et al., 2015). C-cpe group 2 has, compared to c-cpe group 1, retained some genetic traits that likely provide a competitive advantage for host colonization and commensal lifestyle (fucose utilization operon, hyaluronidases *nagJ, nagL,* and sialidase *nanJ*). Adaptation to intestinal colonization is one possible explanation why the arginine deiminase pathway (Arc operon), despite its many uses in stressful conditions, has been deleted from c-cpe group 2 strains. Similar deletion of Arc operon has been suggested to be a trade-off between adaptation to urinary tract compared to the intestine in ESBL *E. coli*; Arc operon was beneficial for strains in the urinary tract, but a burden for mouse intestine colonization (Billard-Pomares et al., 2019). C-cpe group 2 genomes suggest adaptation to a less acidic environment, where the improved acquisition of anaerobic iron does not provide a competitive advantage. Their more diverse virulence gene profile is likely beneficial for commensal life and host colonization and group 2 strains may have retained a niche within the gastrointestinal tract of yet unknown host reservoir(s).

One possible big theme for reductive evolution within c-cpe strains could be an adaptation to planktonic growth. Soncini et al. (2020) have studied the gene expression of *C. perfringens* str. 13 growing on a plate, and many of the identified differentially expressed genes have been deleted from c-cpe strains. This means that the transcription of c-cpe strains growing on a plate differs greatly from those researched by Soncini et al. (2020). C-cpe strains also lack *pfoA*, important for biofilm formation (Vidal et al., 2015), and type IV pilus genes, important for quorum sensing and adhesion, are highly divergent between p-cpe, c-cpe group 1, and c-cpe group 2 (**Supplementary Table 7**). We hypothesize that c-cpe group 1 strains are adapted to environmental liquid niches and prefer planktonic growth. The preferences for liquid versus plated growth between *C. perfringens* strains should be further studied.

We studied heat-resistance and growth temperatures of *cpe*-carrying isolates and concluded that only c-cpe group 1 produced heat-resistant spores, while c-cpe group 2 and p-cpe spores are heat-sensitive. C-cpe strains have been characterized as heat resistant and clonal in previous studies, and to our understanding, this is the first time two separate groups of c-cpe strains with different virulence and functional characteristics have been described.

We also report one c-cpe strain to produce extremely heat-resistant spores without the presence of heat resistance associated allele of Ssp4. Li and McClane have previously shown that a single amino acid substitution in small acid-soluble spore protein beta or Ssp4 significantly affects spore heat resistance (Li and McClane, 2008; Li et al., 2009). Furthermore, genome analysis of 290 strains revealed that this heat resistance allele of Ssp4 was only found from c-cpe group 1 strains indicating this allele has most likely originated within this phylogroup. The capacity to produce highly heat-resistant spores seems to be characteristic and a putative adaptation for c-cpe group 1 strains. However, our results also indicate that other strains have the potential to acquire the capacity to produce highly heat-resistant spores and that heat resistance is not reliant on one allele but more likely to be affected by several genes.

The number of isolates in this study is not sufficient to deduce clear associations between heat-resistant phenotype and genotype within c-cpe strains. We identified two genes potentially associated with spore heat resistance in c-cpe strains: a certain allele of GrpE and the presence of CoA-disulfide reductase NaoX. GrpE in strains producing heat-resistant spores had a point mutation in the N-terminal part before helix structure. This N-terminal part has been proposed to act as a thermometer in *E. coli* (Gelinas et al., 2002), and alterations in this part may

affect the thermal response of strain. CoA-disulfide reductase gene *naoX* has been deleted from c-cpe strains producing heat-sensitive spores. CoA is a common substance in dormant spores, and CoA-disulfide reductases have also been implied to have a role in the heat resistance of dormant spores (Setlow and Setlow, 1977; Hamilton et al., 2014).

To understand the ecology and epidemiology of this important pathogen, lineage-distinguishing typing method to identify closely related toxin-negative strains is needed. Lack of sequenced or lineage-typed environmental isolates compared to clinical isolates is likely contorting our view of *C. perfringens*, which after all is an environmental bacterium common in soil and water.

CONCLUSION

Our study adds valuable genomic and phenotypic data to the publicly available collection of assembled and complete *C. perfringens* strains. The genomes include antibiotic-associated diarrhea isolates, *cpe*-positive environmental isolates, antibiotic-associated diarrhea isolate carrying double enterotoxin gene, and food poisoning isolates including strains representing the here described new subgroup of chromosomal *cpe*-carrying (c-cpe) isolates.

We established the signs of reductive evolution within lineage IV containing the c-cpe strains and elucidated their potential evolutive origin. C-cpe strains share ancestry with contemporary commensal strains isolated from swine and chicken. Comparative genome analysis revealed niche adaptation within the lineage of c-cpe strains. One subgroup produced heat-resistant spores and retained a genetic toolset suitable for survival in harsh environments. Meanwhile, the other subgroup produced heat-sensitive spores, lacked relevant operons for stress survival, but had conserved genes associated with host colonization and commensal lifestyle.

Focus on the rarely reported *cpe*-negative strains of lineage IV is likely to provide more insight into the ecology of c-cpe strains. Currently, the lack of genomic information about environmental isolates and the toxinotyping based on mobile elements is hindering research on ecology and reservoirs of *C. perfringens*.

Based on our results, it is apparent that spore heat resistance of *C. perfringens* strains is likely affected by multiple genes and the capacity to produce heat-resistant spores has developed primarily within one subgroup of chromosomal *cpe*-carrying strains.

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**.

AUTHOR CONTRIBUTIONS

HK and KJ designed the study. KV performed the experiments. KJ performed the genome and data analysis and drafted the manuscript. KJ, PL, ML, and HK contributed to the data analysis and interpretation. KV, PL, RK-T, ML, and HK contributed to the manuscript revision. All authors have read and approved the final manuscript.

FUNDING

This work was funded by the Finnish Foundation of Veterinary Research.

ACKNOWLEDGMENTS

We thank Yagmur Derman and Kirsi Ristkari from the University of Helsinki and Pia Laine from the Institute of Biotechnology for technical assistance and advice.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.717176/full#supplementary-material

Supplementary Figure 1 | Phylogenetic tree of 290 *C. perfringens* strains.

Supplementary Table 1 | *Clostridium perfringens* strains sequenced ad reference genomes used in this study: the cpe genotype, isolation source, and isolation year.

Supplementary Table 2 | Genomes included in the pangenome study.

Supplementary Table 3 | Selected key factors of sequenced *Clostridium perfringens* genomes including plasmids and deposition keys.

Supplementary Table 4 | Genes unique and conserved for c-cpe strains (bi-directional best hit orthologs).

Supplementary Table 5 | Genes unique and conserved for c-cpe strain subgroups (bi-directional best hit orthologs).

Supplementary Table 6 | Phenotyping results, heat-resistance assay results including D values for spores and vegetative cells, and minimum and maximum growth temperatures.

Supplementary Table 7 | Genes divergent between in heat-resistant and heat-sensitive c-cpe strains.

Supplementary Table 8 | Sporulation related genes (Xiao et al., 2015) divergent between c-cpe *Clostridium perfringens* strains and their identity against reference genes from *C. perfringens* SM101.

Supplementary Table 9 | Presentation of genetic lineages within analyzed genomes, GC%, and genome size per lineage.

Supplementary Table 10 | Number of truncated alleles per genome compared to cgMLST scheme.

REFERENCES

Abdel-Glil, M. (2019). *In Silico Genome Analysis and Molecular Typing of Clostridium Perfringens*. Berlin: Freie Univ., Diss., Mensch & Buch Verlag.

Abdel-Glil, M. Y., Thomas, P., Linde, J., Busch, A., Wieler, L. H., Neubauer, H., et al. (2021). Comparative in silico genome analysis of *Clostridium perfringens* unravels stable phylogroups with different genome characteristics and pathogenic potential. *Sci. Rep.* 11:6756. doi: 10.1038/s41598-021-86148-8

Abdelrahim, M. A., Radomski, N., Delannoy, S., Djellal, S., Le Négrate, M., Hadjab, K., et al. (2019). Large-scale genomic analyses and toxinotyping of *Clostridium perfringens* implicated in foodborne outbreaks in France. *Front. Microbiol.* 10:777. doi: 10.3389/fmicb.2019.00777

Ando, Y., Toshibumi, T., Sunagawa, H., and Oka, S. (1985). Heat resistance, spore germination, and enterotoxigenicity of *Clostridium perfringens*. *Microbiol. Immunol.* 29, 317–326. doi: 10.1111/j.1348-0421.1985.tb00830.x

Asha, N. J., Tompkins, D., and Wilcox, M. H. (2006). Comparative analysis of prevalence, risk factors, and molecular epidemiology of antibiotic-associated diarrhea due to *Clostridium difficile*, *Clostridium perfringens*, and *Staphylococcus aureus*. *J. Clin. Microbiol.* 44, 2785–2791. doi: 10.1128/JCM.00165-06

Becerra, J. E., Yebra, M. J., and Monedero, V. (2015). An L-fucose operon in the probiotic *Lactobacillus rhamnosus* GG is involved in adaptation to gastrointestinal conditions. *Appl. Environ. Microbiol.* 81, 3880–3888. doi: 10.1128/AEM.00260-15

Billard-Pomares, T., Clermont, O., Castellanos, M., Magdoud, F., Royer, G., Condamine, B., et al. (2019). The arginine deiminase operon is responsible for a fitness trade-off in extended-spectrum-β-lactamase-producing strains of *Escherichia coli*. *Antimicrob. Agents Chemother.* 63:e00635-19. doi: 10.1128/AAC.00635-19

Brynestad, S., Synstad, B., and Granum, P. E. (1997). The *Clostridium perfringens* enterotoxin gene is on a transposable element in type a human food poisoning strains. *Microbiology* 143, 2109–2115. doi: 10.1099/00221287-143-7-2109

Brynildsrud, O., Bohlin, J., Scheffer, L., and Eldholm, V. (2016). Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol.* 17:238. doi: 10.1186/s13059-016-1108-8

Chin, C.-S., David, H. A., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10, 563–569. doi: 10.1038/nmeth.2474

Collie, R. E., and McClane, B. A. (1998). Evidence that the enterotoxin gene can be episomal in *Clostridium perfringens* isolates associated with non-food-borne human gastrointestinal diseases. *J. Clin. Microbiol.* 36, 30–36. doi: 10.1128/jcm.36.1.30-36.1998

Cornillot, E., Saint-Joanis, B., Daube, G., Katayama, S., Granum, P. E., Canard, B., et al. (1995). The enterotoxin gene (Cpe) of *Clostridium perfringens* can be chromosomal or plasmid-borne. *Mol. Microbiol.* 15, 639–647. doi: 10.1111/j.1365-2958.1995.tb02373.x

Darling, A. E., Mau, B., and Perna, N. T. (2010). Progressivemauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147

Deguchi, A., Miyamoto, K., Kuwahara, T., Miki, Y., Kaneko, I., Li, J., et al. (2009). Genetic characterization of type a enterotoxigenic *Clostridium perfringens* strains. *PLoS One* 4:e0005598. doi: 10.1371/journal.pone.0005598

Derman, Y., Lindström, M., Selby, K., and Korkeala, H. (2011). Growth of group II *Clostridium botulinum* strains at extreme temperatures. *J. Food Protect.* 74, 1797–1804. doi: 10.4315/0362-028X.JFP-11-187

Dimitriu, T. A., Lauren, M., Angus, B., and Ben, R. (2019). Bacteria from natural populations transfer plasmids mostly towards their kin. *Proc. R. Soc. B* 286:20191110. doi: 10.1098/rspb.2019.1110

Duncan, C. L., and Strong, D. H. (1968). Improved medium for sporulation of *Clostridium perfringens*. *Appl. Microbiol.* 16, 82–89. doi: 10.1128/aem.16.1.82-89.1968

Feng, Y., Fan, X., Zhu, L., Yang, X., Liu, Y., Gao, S., et al. (2020). Phylogenetic and genomic analysis reveals high genomic openness and genetic diversity of *Clostridium perfringens*. *Microb. Genom.* 6:e000441. doi: 10.1099/mgen.0.000441

Gelinas, A. D., Langsetmo, K., Toth, J., Bethoney, K. A., Stafford, W. F., and Harrison, C. J. (2002). A Structure-based interpretation of *E. coli* grpe thermodynamic properties. *J. Mol. Biol.* 323, 131–142. doi: 10.1016/S0022-2836(02)00915-4

Gillespie, J. J., Wattam, A. R., Cammer, S. A., Gabbard, J. L., Shukla, M. P., Dalay, O., et al. (2011). PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. *Infect. Immun.* 79, 4286–4298. doi: 10.1128/IAI.00207-11

Hamilton, C. J., Arbach, M., Groom, M. (2014). "Beyond glutathione: Different low molecular weight thiols as mediators of redox regulation and other metabolic functions in lower organisms", *Recent Advances in Redox Active Plant and Microbial Products: From Basic Chemistry to Widespread Applications in Medicine and Agriculture* (Netherlands: Springer), 291–320.

Hatheway, C. L. (1990). Toxigenic clostridia. *Clin. Microbiol. Rev.* 3, 66–98. doi: 10.1128/cmr.3.1.66-98.1990

Heikinheimo, A., Lindström, M., Granum, P. E., and Korkeala, H. (2006). Humans as reservoir for enterotoxin gene-carrying *Clostridium perfringens* type A. *Emerg. Infect. Dis.* 12, 1724–1729. doi: 10.3201/eid1211.060478

Hu, W. S., Kim, H., and Koo, O. K. (2018). Molecular genotyping, biofilm formation and antibiotic resistance of enterotoxigenic *Clostridium perfringens* isolated from meat supplied to school cafeterias in South Korea. *Anaerobe* 52, 115–121. doi: 10.1016/j.anaerobe.2018.06.011

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-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

Keto-Timonen, R., Heikinheimo, A., Eerola, E., and Korkeala, H. (2006). Identification of *Clostridium* species and DNA fingerprinting of *Clostridium perfringens* by amplified fragment length polymorphism analysis. *J. Clin. Microbiol.* 44, 4057–4065. doi: 10.1128/JCM.01275-06

Kiu, R., Caim, S., Alexander, S., Pachori, P., and Hall, L. J. (2017). Probing genomic aspects of the multi-host pathogen *Clostridium perfringens* reveals significant pangenome diversity, and a diverse array of virulence factors. *Front. Microbiol.* 8:2485. doi: 10.3389/fmicb.2017.02485

Kiu, R., Caim, S., Painset, A., Pickard, D., Swift, C., Dougan, G., et al. (2019). Phylogenomic analysis of gastroenteritis-associated *Clostridium perfringens* in england and wales over a 7-year period indicates distribution of clonal toxigenic strains in multiple outbreaks and extensive involvement of enterotoxin-encoding (Cpe) plasmids. *Microb. Genom.* 5:e000297. doi: 10.1099/mgen.0.000297

Kiu, R., and Hall, L. J. (2018). An update on the human and animal enteric pathogen *Clostridium perfringens*. *Emerg. Microbes Infect.* 7:141. doi: 10.1038/s41426-018-0144-8

Korkeala, H., Mäkelä, P., and Suominen, H. (1990). Growth temperatures of ropy slime-producing lactic acid bacteria. *J. Food Protect.* 53, 793–794. doi: 10.4315/0362-028X-53.9.793

Lahti, P., Heikinheimo, A., Johansson, T., and Korkeala, H. (2008). *Clostridium perfringens* type a strains carrying a plasmid-borne enterotoxin gene (Genotype IS1151-Cpe or IS1470-like-Cpe) as a common cause of food poisoning. *J. Clin. Microbiol.* 46, 371–373. doi: 10.1128/JCM.01650-07

Lahti, P., Lindström, M., Somervuo, P., Heikinheimo, A., and Korkeala, H. (2012). Comparative genomic hybridization analysis shows different epidemiology of chromosomal and plasmid-borne Cpe-carrying *Clostridium perfringens* type A. *PLoS One* 7:e46162. doi: 10.1371/journal.pone.0046162

Li, J., Adams, V., Bannam, T. L., Miyamoto, K., Garcia, J. P., Uzal, F. A., et al. (2013). Toxin plasmids of *Clostridium perfringens*. *Microbiol. Mol. Biol. Rev.* 77, 208–233. doi: 10.1128/mmbr.00062-12

Li, J., and McClane, B. A. (2006a). Comparative effects of osmotic, sodium nitrite-induced, and ph-induced stress on growth and survival of *Clostridium perfringens* type a isolates carrying chromosomal or plasmid-borne enterotoxin genes. *Appl. Environ. Microbiol.* 72, 7620–7625. doi: 10.1128/AEM.01911-06

Li, J., and McClane, B. A. (2006b). Further comparison of temperature effects on growth and survival of *Clostridium perfringens* type a isolates carrying a chromosomal or plasmid-borne enterotoxin gene. *Appl. Environ. Microbiol.* 72, 4561–4568. doi: 10.1128/AEM.00177-06

Li, J., and McClane, B. A. (2008). A novel small acid soluble protein variant is important for spore resistance of most *Clostridium perfringens* food poisoning isolates. *PLoS Pathog.* 48:e1000056. doi: 10.1371/journal.ppat.1000056

Li, J., Miyamoto, K., Sayeed, S., and McClane, B. A. (2010). Organization of the Cpe locus in CPE-positive *Clostridium perfringens* type C and D isolates. *PLoS One* 5:e0010932. doi: 10.1371/journal.pone.0010932

Li, J., Paredes-Sabja, D., Sarker, M. R., and McClane, B. A. (2009). Further characterization of *Clostridium perfringens* small acid soluble protein-4 (Ssp4) properties and expression. *PLoS One* 4:e6249. doi: 10.1371/journal.pone.0006249

Li, J., Sayeed, S., and McClane, B. A. (2007). Prevalence of enterotoxigenic *Clostridium perfringens* isolates in pittsburgh (pennsylvania) area soils and home kitchens. *Appl. Environ. Microbiol.* 73, 7218–7224. doi: 10.1128/AEM.01075-07

Li, J., Zhou, Y., Yang, D., Zhang, S., Sun, Z., Wang, Y., et al. (2021). Prevalence and antimicrobial susceptibility of *Clostridium perfringens* in chickens and pigs from Beijing and Shanxi, China. *Vet. Microbiol.* 24:252. doi: 10.1016/j.vetmic.2020.108932

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158

Lindström, M., Heikinheimo, A., Lahti, P., and Korkeala, H. (2011). Novel insights into the epidemiology of *Clostridium perfringens* type a food poisoning. *Food Microbiol.* 28, 192–198. doi: 10.1016/j.fm.2010.03.020

Lugli, G. A., Milani, C., Mancabelli, L., Turroni, F., Ferrario, C., Duranti, S., et al. (2017). Ancient bacteria of the ötzi's microbiome: a genomic tale from the copper age. *Microbiome* 5:5. doi: 10.1186/s40168-016-0221-y

Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., et al. (2009). CDD: specific functional annotation with the conserved domain database. *Nucleic Acids Res.* 37, D205–D210.

Miyamoto, K., Fisher, D. J., Li, J., Sayeed, S., Akimoto, S., and McClane, B. A. (2006). Complete sequencing and diversity analysis of the enterotoxin-encoding plasmids in *Clostridium perfringens* type a non-food-borne human gastrointestinal disease isolates. *J. Bacteriol.* 188, 1585–1598. doi: 10.1128/JB.188.4.1585-1598.2006

Miyamoto, K., Yumine, N., Mimura, K., Nagahama, M., Li, J., McClane, B. A., et al. (2011). Identification of novel *Clostridium perfringens* type e strains that carry an iota toxin plasmid with a functional enterotoxin gene. *PLoS One* 6:e0020376. doi: 10.1371/journal.pone.0020376

Myers, G. S., Rasko, D. A., Cheung, J. K., Ravel, J., Seshadri, R., DeBoy, R. T., et al. (2006). Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. *Genome Res.* 16, 1031–1040. doi: 10.1101/gr.5238106

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421

Parrello, B., Butler, R., Chlenski, P., Olson, R., Overbeek, J. C., Pusch, G. D., et al. (2019). A machine learning-based service for estimating quality of genomes using PATRIC. *BMC Bioinformatics* 20:486. doi: 10.1186/s12859-019-3068-y

Petit, L., Gibert, M., and Popoff, M. R. (1999). *Clostridium perfringens*: toxinotype and genotype. *Trends Microbiol.* 7, 104–110. doi: 10.1016/S0966-842X(98)01430-9

Pollack, J. R., Ames, B. N., and Neilands, J. B. (1970). Iron transport in *Salmonella* typhimurium: mutants blocked in the biosynthesis of enterobactin. *J. Bacteriol.* 104, 635–639. doi: 10.1128/jb.104.2.635-639.1970

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490

Raju, D., and Sarker, M. R. (2005). Comparison of the levels of heat resistance of wild-type, Cpe knockout, and Cpe plasmid-cured *Clostridium perfringens* type a strains. *Appl. Environ. Microbiol.* 71, 7618–7620. doi: 10.1128/AEM.71.11.7618-7620.2005

Rood, J. I., Adams, V., Lacey, J., Lyras, D., McClane, B. A., Melville, S. B., et al. (2018). Expansion of the *Clostridium perfringens* toxin-based typing scheme. *Anaerobe* 53, 5–10. doi: 10.1016/j.anaerobe.2018.04.011

Sarker, M. R., Shivers, R. P., Sparks, S. G., Juneja, V. K., and McClane, B. A. (2000). Comparative experiments to examine the effects of heating on vegetative cells and spores of *Clostridium perfringens* isolates carrying plasmid genes versus chromosomal enterotoxin genes. *Appl. Environ. Microbiol.* 66, 3234–3240. doi: 10.1128/AEM.66.8.3234-3240.2000

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

Setlow, B., and Setlow, P. (1977). Levels of acetyl coenzyme A, reduced and oxidized coenzyme A, and coenzyme A in disulfide linkage to protein in dormant and germinated spores and growing and sporulating cells of Bacillus megaterium. *J. Bacteriol.* 132, 444–452. doi: 10.1128/jb.132.2.444-452.1977

Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A., Shiba, T., et al. (2002). Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. *Proc. Natl. Acad. Sci. U.S.A.* 99, 996–1001. doi: 10.1073/pnas.022493799

Silva, F., Da, M. T., Hald, T., Liebana, E., Allende, A., Hugas, M., et al. (2015). Risk ranking of pathogens in ready-to-eat unprocessed foods of non-animal origin (FoNAO) in the EU: initial evaluation using outbreak data (2007-2011). *Int. J. Food Microbiol.* 195, 9–19. doi: 10.1016/j.ijfoodmicro.2014.11.005

Silva, M., Machado, M. P., Silva, D. N., Rossi, M., Moran-Gilad, J., Santos, S., et al. (2018). ChewBBACA: a complete suite for gene-by-gene schema creation and strain identification. *Microb. Genom.* 4:e000166. doi: 10.1099/mgen.0.000166

Soncini, S. R., Hartman, A. H., Gallagher, T. M., Camper, G. J., Jensen, R. V., and Melville, S. B. (2020). Changes in the expression of genes encoding type IV pili-associated proteins are seen when *Clostridium perfringens* is grown in liquid or on surfaces. *BMC Genomics* 21:45. doi: 10.1186/s12864-020-6453-z

Sparks, S. G., Carman, R. J., Sarker, M. R., and McClane, B. A. (2001). Genotyping of enterotoxigenic *Clostridium perfringens* fecal isolates associated with antibiotic-associated diarrhea and food poisoning in North America. *J. Clin. Microbiol.* 39, 883–888. doi: 10.1128/JCM.39.3.883-888.2001

Staden, R., Judge, D. P., and Bonfield, J. K. (2003). Managing sequencing projects in the GAP4 environment. *Introduct. Bioinform.* 4, 327–344. doi: 10.1007/978-1-59259-335-4_20

Stahl, M., Friis, L. M., Nothaft, H., Liu, X., Li, J., Szymanski, C. M., et al. (2011). L-fucose utilization provides *Campylobacter jejuni* with a competitive advantage. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7194–7199. doi: 10.1073/pnas.1014125108

Tonkin-Hill, G., MacAlasdair, N., Ruis, C., Weimann, A., Horesh, G., Lees, J., et al. (2020). Producing polished prokaryotic pangenomes with the *Panaroo pipeline*. *Genome Biol.* 21:180. doi: 10.1186/s13059-020-02090-4

Vidal, J. E., Shak, J. R., and Canizalez-Roman, A. (2015). The CpAL quorum sensing system regulates production of hemolysins CPA and PFO to build *Clostridium perfringens* biofilms. *Infect. Immun.* 83, 2430–2442. doi: 10.1128/IAI.00240-15

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., et al. (2016). Improvement to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res.* 4, D535–D542. doi: 10.1093/nar/gkw1017

Weiss, K. F., and Strong, D. H. (1967). Some properties of heat-resistant and heat-sensitive strains of *Clostridium perfringens*. I. Heat resistance and toxigenicity. *J. Bacteriol.* 93, 21–26. doi: 10.1128/jb.93.1.21-26.1967

Wen, Q., and McClane, B. A. (2004). Detection of enterotoxigenic *Clostridium perfringens* type A isolates in American retail foods. *Appl. Environ. Microbiol.* 70, 2685–2691. doi: 10.1128/AEM.70.5.2685-2691.2004

Xiao, Y., Van Hijum, S. A. F. T., Abee, T., and Wells-Bennik, M. H. J. (2015). Genome-wide transcriptional profiling of *Clostridium perfringens* SM101 during sporulation extends the core of putative sporulation genes and genes determining spore properties and germination characteristics. *PLoS One* 10:e0127036. doi: 10.1371/journal.pone.0127036

Xiao, Y., Wagendorp, A., Moezelaar, R., Abee, T., and Wells-Bennik, M. H. J. (2012). A wide variety of *Clostridium perfringens* type a food-borne isolates that carry a chromosomal Cpe gene belong to one multilocus sequence typing cluster. *Appl. Environ. Microbiol.* 78, 7060–7068. doi: 10.1128/AEM.01486-12

Zhou, Z., Alikhan, N. F., Sergeant, M. J., Luhmann, N., Vaz, C., Francisco, A. P., et al. (2018). Grapetree: visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Res.* 28, 1395–1404. doi: 10.1101/gr.232397.117

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 24 September 2021
doi: 10.3389/fgene.2021.724661

Diversity Scaling of Human Digestive Tract (DT) Microbiomes: The Intra-DT and Inter-individual Patterns

Hongju (Daisy) Chen[1,2,3†], *Bin Yi*[3†], *Qiang Liu*[3], *Xia Xu*[3†], *Lin Dai*[4*] and *Zhanshan (Sam) Ma*[1,2,5*]

[1]*Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China,* [2]*Kunming College of Life Sciences, University of Chinese Academy of Sciences, Kunming, China,* [3]*College of Mathematics, Honghe University, Mengzi, China,* [4]*Faculty of Science, Kunming University of Science and Technology, Kunming, China,* [5]*Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China*

OPEN ACCESS

Edited by:
Ricardo S. Ramiro, Gulbenkian Institute of Science (IGC), Portugal

Reviewed by:
Erida Gjini, Universidade de Lisboa, Portugal
Hong-Yu Zhang, Huazhong Agricultural University, China

***Correspondence:**
Zhanshan (Sam) Ma
ma@vandals.uidaho.edu
Lin Dai
453091926@qq.com

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Genetics

Received: *14 June 2021*
Accepted: *01 September 2021*
Published: *24 September 2021*

Citation:
Chen H, Yi B, Liu Q, Xu X, Dai L and Ma Z (2021) Diversity Scaling of Human Digestive Tract (DT) Microbiomes: The Intra-DT and Inter-individual Patterns. Front. Genet. 12:724661. doi: 10.3389/fgene.2021.724661

The human gut microbiome has been extensively studied, but its diversity scaling (changes or heterogeneities) along the digestive tract (DT) as well as their inter-individual heterogeneities have not been adequately addressed to the best of our knowledge. Here we fill the gap by applying the diversity-area relationship (DAR), a recent extension to the classic species-area relationship (SAR) in biogeography, by reanalyzing a dataset of over 2000 16s-rRNA microbiome samples obtained from 10 DT sites of over 200 individuals. We sketched out the biogeography "maps" for each of the 10 DT sites by cross-individual DAR analysis, and the intra-DT distribution pattern by cross-DT-site DAR analysis. Regarding the inter-individual biogeography, it was found that all DT sites have the invariant (constant) scaling parameter—all sites possessing the same diversity change rate across individuals, but most sites have different *potential* diversities, which include the portions of diversity that may be absent locally but present regionally. In the case of this study, the potential diversity of each DT site covers the total diversity of the respective site from all individuals in the cohort. In terms of the genus *richness*, an average individual hosts approximately 20% of the population-level genus richness (total bacterial genus of a human population). In contrast, in terms of community *biodiversity*, the percentages of individual over population may exceed 90%. This suggests that the differences between individuals in their DT microbiomes are predominantly in the composition of bacterial species, rather than how their abundances are distributed (*i.e.*, biodiversity). Regarding the intra-DT patterns, the scaling parameter (*z*) is larger—suggesting that the intra-DT biodiversity changes are larger than inter-individual changes. The higher intra-DT heterogeneity of bacteria diversity, as suggested by larger intra-DT *z* than the inter-individual heterogeneity, should be expected since the intra-DT heterogeneity reflects the functional differentiations of the DT tract, while the inter-individual heterogeneity (z) reflects the difference of the same DT site across individuals. On average, each DT site contains 21–36% of the genus diversity of the whole DT, and the percentages are even higher in terms of higher taxon levels.

Keywords: Inter-individual microbiome heterogeneity, Intra-DT microbiome heterogeneity, Diversity-area relationship (DAR), Potential diversity, Diversity scaling

INTRODUCTION

There have been extensive studies on various aspects of human gut microbiomes over the last decade or so, particularly on their diversities (for example, HMP Consortium (Human Microbiome Project Consortium), 2012a; HMP Consortium (Human Microbiome Project Consortium), 2012b; Lozupone et al., 2012; Segata et al., 2012). However, certain information on the diversity scaling (changes) across individuals or across various sites of an individual's DT (digestive tract) seems to be still missing. For example, to what extent can an individual's gut microbial species richness (number of species) represent a population? What about species diversity (such as Shannon evenness index) if individual *vs* population is compared? How is the microbiome distributed along one's DT, in terms of species richness or diversity? How about the scaling relationships on higher taxon levels, such as phylum, class, order, and family? Similar questions are traditionally investigated in biogeography. One of the most well-known ecological laws in the field is the so-termed species-area relationship (SAR), which was first discovered in the 19th century (Watson, 1835) and has been extensively investigated since the 1960s (Preston, 1960; Connor and McCoy, 1979; Rosenzweig, 1995; Lomolino, 2000; Drakare et al., 2006; Tjørve and Tjørve, 2008, 2009; Harte et al., 2009; He and Hubbell 2011; Sizling et al., 2011; Storch et al., 2012; Triantis et al., 2012; Helmus et al., 2014). It is considered as "ecology's most general, yet protean pattern" by Lomolino (2000) and Whittaker and Triantis (2012). In practice, SAR has become one of the most important theories and models in conservation biology and biodiversity protection. This is because SAR can be applied to establish a simple functional relationship between the number of species (species richness) and the area of a region. It was found that SAR typically follow a simple power function in the form of $S = cA^z$, where S is the number of species, A is the size of area (in the case of human microbiome, A can be treated as the number of individuals sampled), and z and c are SAR parameters. In particular, z is termed the scaling parameter of SAR, and it is a measure of change (increase) rate of species number over area.

It is generally recognized that the number of species, or formally species richness, is a rather convenient but very rough measure of biodiversity. This is because biodiversity is obviously strongly influenced by both species richness and species abundances of individual species. Several biodiversity metrics that consider both species numbers and abundances have been proposed and widely used in ecology since 1960s. A recent consensus has been that the Hill numbers, which were first introduced by Hill (1973) into ecology but did not receive significant attention until recently, offer the most appropriate metrics for measuring alpha-diversity (Chao et al., 2012; Chao et al., 2014a; Chao et al., 2014b) since most existing diversity metrics such as species richness, Shannon entropy, and Simpson index turned out to be special cases (or functions) of the Hill numbers. To take advantages of the Hill numbers as general biodiversity metrics, Ma (2018a) extended the classic SAR to more general diversity-area relationship (DAR) by substituting the species richness with general diversity measured with Hill numbers. In the present study, we applied the DAR approach to address the previous raised questions regarding the changes of DT microbiome diversity, both across individuals (inter-individual) and across DT sites (intra-individual or intra-DT).

To investigate the inter-individual and intra-individual (intra-DT) diversity scaling patterns with the DAR approach, we used a dataset originally collected by Segata et al (2012). Their study collected ten microbiome samples from each of over 200 individuals' digestive tract [buccal mucosa (BM), keratinized gingiva (KG), hard palate (HP), throat (Th), palatine tonsils (PT), tongue dorsum (TD) and saliva (Sal), supraginval (SupP), subgingival plaques (SubP), and stool (Stool)]. The dataset provides an ideal opportunity for us achieve the objective of this study—analyzing the inter-individual and intra-individual diversity scaling of the human DT microbiomes.

MATERIALS AND METHODS

A Brief Description of the Digestive Tract Microbiome Dataset

The DT microbiome dataset we reanalyzed in this study was first reported by Segata et al. (2012), which is part of the Human Microbiome Project (HMP). A total of 2078 DT microbiome samples were collected from 242 healthy adults aged from 18 to 40 years old, who were enrolled in the HMP. The ten DT sites sampled included seven from the oral cavity (BM, KG, HP, Th, PT, TD, and Sal), two from the oropharynx (SupP and SubP), and one from the gut (Stool). The operational taxonomic unit (OTU) tables and the metadata information on the individuals are available at https://www.hmpdacc.org/, and for more detailed information on the dataset, refer to Segata *et al.* (2012).

Computational Procedures for the DAR Analysis

Definitions of Alpha Diversities

We applied the Hill numbers (Hill 1973; Chao et al., 2012; 2014a) to measure the alpha diversity, which are defined as:

$$ {}^{q}D = \left(\sum_{i=1}^{S} p_i^q \right)^{1/(1-q)} \tag{1} $$

where D is the diversity in Hill numbers, q (=0, 1, 2, ...) is the order number of diversity, S is number of species (or OTUs), and p_i is the relative abundance of OTU i. When $q = 1$, the Hill number is not defined, but we can figure out its limit as q approaches to 1 as follows:

$$ {}^{1}D = \lim_{q \to 1} {}^{q}D = \exp\left(-\sum_{i=1}^{S} p_i \log(p_i) \right) \tag{2} $$

The Hill numbers are a series of diversity measures corresponding to different diversity orders (q), where q determines the weight of relative frequencies of species abundances. When $q = 0$, species abundance is not involved in the calculation, and ${}^{0}D$ is the number of OTUs or the species richness. When $q = 1$, ${}^{1}D$ equals the *exponential* of Shannon entropy, which represents the number of typical or

common species in the community. When $q = 2$, ${}^2D$ is equal to the reciprocal of the Simpson index and represents the number of species with high abundance. Generally, ${}^qD$ represents the diversity of a community with $x = {}^qD$ equally abundant species.

DAR Analysis

According to Ma (2018a), Ma (2018b), Ma (2019), we selected and used two DAR models for the DT microbiome in this study, one is the power law (PL) model, and another is the power law with exponential cutoff (PLEC) model. The PL model is:

$$ {}^qD = cA^z \tag{3}$$

where ${}^qD$ is the diversity measured in the qth order Hill numbers, A is *area*, and c and z are the PL parameters.

The PLEC model is:

$$ {}^qD = cA^z \exp(dA) \tag{4}$$

where d is a third parameter with taper-off effect, and exp (dA) is the exponential decay term that eventually overwhelms the power law behavior when A becomes very large.

We transformed the non-linear **Eqs 3**, **4** into log-linear regression **Equations 5**, **6** to estimate the parameters of PL and PLEC models, respectively:

$$ \ln(D) = \ln(c) + z\ln(A) \tag{5}$$

$$ \ln(D) = \ln(c) + z\ln(A) + dA \tag{6}$$

Four Important DAR Parameters and Corresponding Profiles

According to Ma (2018a) and Ma and Li (2019), there are four important DAR parameters, including the diversity scaling parameter (z), pair-wise diversity overlap (PDO or g), maximal accrual diversity (MAD or $D_{\max}$), and ratio of individual diversity to population accrual diversity (RIP).

i) As the slope or tangent of the PL-DAR model (**Eqs. 3**, **5**, the z-value was termed as the diversity scaling parameter.

ii) The PDO or g was defined as,

$$ g = 2 - 2^z \tag{7}$$

where z is the scaling parameter of the PL-DAR model. The range of g is generally between 0 and 1. If $z = 1$, then $g = 0$ and there is no overlap or similarity. If $z = 0$, then $g = 1$ and there is a total overlap or similarity.

iii) The MAD or $D_{\max}$ was defined based on the PLEC-DAR model (**Eqs. 4**, **6**, that is,

$$ {}^qD_{\max} = c\left(-\frac{z}{d}\right)^z \exp(-z) = cA_{\max}^z \exp(-z) \tag{8}$$

where $A_{\max} = -z/d$ is the number of individuals (microbiome samples) needed to reach the MAD, and c and d are parameters of the PLEC-DAR model.

iv) The RIP was defined as,

$$ {}^qRIP = {}^qc/{}^qD_{\max} \tag{9}$$

Where ${}^qc$ is the parameter of PL-DAR model at diversity order of q, and ${}^qD_{\max}$ is the MAD that can be computed with **Eq. 8**.

Ma (2018a) and Ma and Li (2019) also defined the relationships between these four parameters and the diversity order (q) as the DAR profile, PDO profile, MAD profile and RIP profile, respectively.

Design for DAR Analysis

Our analysis consists of two parts, inter-individual (cross-individual) DAR analysis and intra-individual (cross-DT site) DAR analysis. Based on the inter-individual DAR analysis, we can investigate diversity scaling for each of the 10 DT sites across 200 + individuals. We built PL-DAR and PLEC-DAR models for each DT site, and further conducted permutation tests (randomization test) for the DAR parameters (i.e., z and $D_{\max}$) of different DT sites. The procedure for the randomization test refers to Collingridge (2013), in which the number of permutations or re-samplings was set to 1,000 times. The p-value of randomization test can be used to determine the significance of differences. It is noted that parameter c of the PL model indicates the diversity in the first unit of area to accrue. Thus, to exclude the influence of the accrual order of area unit on parameter c, we randomly permutated the area units to be accumulated each time the DAR model was built. In the inter-individual DAR analysis, we repeated this re-sampling procedure 100 times, and adopted the averages of the model parameters from the 100 times of DAR fittings as the final model parameters of the inter-individual DAR model for the DT site under investigation. The detailed computational procedures can be found in Ma (2018a), Ma (2018b). Based on the intra-individual DAR analysis, we can investigate intra-DT diversity scaling across 10 DT sites. The steps of intra-individual DAR analysis are as follows: 1) We first randomly selected a sample from all samples belonging to the same DT site, and a total of 10 samples from 10 DT sites constitute the intra-DAR samples for an "individual". 2) We built PL-DAR and PLEC-DAR models for the "individual", in which area units (DT sites) were accumulated in order of anatomy structure from the oral cavity to the intestinal tract. 3) We repeated steps (*i*)-(*ii*) 1,000 times, and adopted the averages of the model parameters from the 1,000 times of DAR fittings as the final model parameters of the intra-individual DAR model. In addition, beside genus taxon level, we also analyzed the DAR patterns of other taxa including phylum, class, order, and family taxa.

RESULTS

Inter-Individual DAR Modeling of the Human Gut Microbiome for Each DT Site

At each of the five taxon levels, we built PL-DAR and PLEC-DAR models for each of the 10 sites of the human DT microbiome. **Supplementary Table S1** listed the results of fitting DAR models for each 10 DT sites at the genus taxon level, including the diversity order (q) of the Hill numbers, the mean model parameters (z, c, d, g, $D_{\max}$ and RIP) and measures for

goodness-of-fitting (R and p-value). N is the number of successful fittings out of 100 re-samplings, as explained previously. **Supplementary Tables S2-S5** listed the results at other taxon levels, i.e., phylum, class, order, and family. **Tables 1** and **2** listed results from the permutation tests for the differences in scaling parameter (z) and D_{max} of different DT microbiome sites. From these results, we summarize the following findings:

i) *DAR profile*: At the genus level, the average scaling parameter (z) of the 10 DT sites across diversity order q = 0–3 is z = (0.294, 0.038, 0.020, 0.014), and their standard errors ranged from 0.003 to 0.008 (as shown in **Supplementary Table S1** and **Figure 1**). For each DT microbiome site, the z-values monotonically decreased with the diversity order q. As shown in **Supplementary Tables S2-S5**, the scaling parameter (z) gradually decreased with the taxon level, and for example, the average scaling parameter (z) at phylum level across q = 0–3 is z = (0.108, 0.013, 0.009, 0.007). As shown in **Table 1**, no significant differences in scaling parameter (z) at species-level were detected among 10 DT sites, and the detailed results of randomization tests were listed in **Supplementary Table S7**.

ii) *PDO profile*: PDO or parameter g characterizes the overlap or similarity between pair-wise microbiomes. As shown in **Supplementary Table S1**, at genus level, the average PDO parameter (g) of the 10 DT sites across diversity order q = 0–3 is g = (0.773, 0.973, 0.985, 0.989), and their standard errors ranged from 0.002 to 0.007. In contrast to the diversity scaling parameter (z), the PDO parameter (g) increased with either diversity order q or taxon level (see **Supplementary Tables S2-S5**).

iii) *MAD profile*: As shown in **Supplementary Table S1**, at genus level, the average D_{max} of the 10 DT sites across diversity order q = 0–3 is D_{max} = (288.7, 14.2, 8.2, 6.6). MAD or parameter D_{max} can be considered as a proxy of potential or "dark" diversity, which can be used to estimate microbial biodiversity of a DT sites for a human population. For example, at taxonomic genus, the maximal accrual of species richness (Hill numbers for q = 0) across individuals is around 289. Similar to the DAR profile, the MAD profile of each DT site decreased with diversity order q (**Supplementary Table S1** and **Figure 1**). As shown in **Supplementary Tables S2-S5**, the D_{max} decreased with the taxon level, and for example, the average D_{max} at phylum level across q = 0–3 is D_{max} = (14.7, 4.0, 3.3, 3.1). We test the difference in parameter D_{max} between each pair of DT sites by using randomization test, and results were listed in **Table 2**. At diversity order q = 0, two out of 45 or 4.4% comparisons between 10 DT sites exhibited statistically significant differences. These two comparisons with difference were SAL *vs* Stool and SupP *vs* Stool. At diversity order q = 1-3, there were 73.3% (33/45), 57.8% (26/45) and 68.9% (31/45) comparisons with significant differences, respectively. Please see **Supplementary Table S7** for the detailed results of randomization tests.

iv) *RIP profile*: As shown in **Supplementary Table S1**, at genus level, the average RIP of the 10 DT sites across diversity order q = 0–3 is RIP = (19.1, 83.1, 90.8, 93.4), and their standard errors ranged from 1.0 to 3.3. The RIP profiles of each DT site monotonically increased with q (**Supplementary Table S1** and **Figure 1**). RIP can characterize the relationship between individual-level diversity and population-level diversity. For example, at diversity order q = 0, RIP = 19.1 indicating that an average individual can represent for approximately 19% of population diversity.

Intra-DT Diversity Scaling (Across DT Sites) Analysis With Intra-Individual DAR Models

At each taxon level, we built the PL- and PLEC-DAR models across 10 DT sites to investigate intra-DT distribution pattern. **Supplementary Table S6** and **Figure 2** list the results of fitting intra-DAR models at all five taxon levels, including the same parameters as **Supplementary Table S1**. N is the number of successful fittings out of 1,000 re-samplings, as explained in the materials and methods section. When q = 0, fitting to both intra-DAR models were failed at phylum level. The reason is that different DT sites had the same number of phyla in 890 out of 1,000 re-samplings, which results in the relationship between $\ln(D)$ and $\ln(c)$ being equivalent to a line parallel to the x-axis, and the estimation of goodness-of-fitting R and p-value being failed. From **Supplementary Table S6**, we summarize the following findings:

i) *DAR profile*: At the genus level, the scaling parameter (z) across diversity order q = 0–3 is z = (0.417, 0.512, 0.535, 0.493). DAR profile increased with q at q = 0–3, but slightly decreased at q = 4. Compared with inter-individual DAR profile, there was not much difference in z-values between at q = 0 and other diversity orders in intra-individual diversity scaling. Similar to the inter-individual diversity scaling, the scaling parameter (z) also decreased with the taxon level, but dropped relatively slowly.

ii) *PDO profile*: The trends of PDO profiles over diversity order q were contrary to those of DAR profiles.

iii) *MAD profile*: At genus level, D_{max} across diversity order q = 0–3 is D_{max} = (140.0, 18.2, 12.4, 5.8). MAD or parameter D_{max} offers estimates for the potential microbial diversity in the whole human DT. For example, at taxonomic genus, the theoretical maximal accrual of species richness (Hill numbers for q = 0) across all DT sites is around 140. Similar to the inter-individual MAD profile, D_{max} decreased with diversity order q and taxon level.

iv) *RIP profile*: At genus level, RIP across diversity order q = 0–3 is RIP = (32.5, 28.7, 20.8, 36.1). The RIP profiles of each DT sites monotonically increased with q. Compared with inter-individual DAR models, the trends of RIP over either diversity order q or taxon level were less obvious in intra-individual DAR models.

CONCLUSION AND DISCUSSION

Understanding the biogeography or the spatial distribution of biodiversity is of critical significance both theoretically and practically. Theoretically, biogeography shows a big picture of

TABLE 1 | Summary of the randomization tests from **Supplementary Table S7**: the pair-wise comparisons of the DAR parameter (z) among the ten DT sites of the human digestive microbiome (at Genus taxon level), each digit of the code (*e.g.*, "0,000") represents the result of randomization test for each diversity order $q = 0, 1, 2, 3$. '0 = no significant difference, "1" = significant difference.

Digestive sites	BM	HP	KG	PT	SAL	Stool	SubP	SupP	TD	Th
BM	NA	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
HP	0,000	NA	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
KG	0,000	0,000	NA	0,000	0,000	0,000	0,000	0,000	0,000	0,000
FPT	0,000	0,000	0,000	NA	0,000	0,000	0,000	0,000	0,000	0,000
SAL	0,000	0,000	0,000	0,000	NA	0,000	0,000	0,000	0,000	0,000
Stool	0,000	0,000	0,000	0,000	0,000	NA	0,000	0,000	0,000	0,000
SubP	0,000	0,000	0,000	0,000	0,000	0,000	NA	0,000	0,000	0,000
SupP	0,000	0,000	0,000	0,000	0,000	0,000	0,000	NA	0,000	0,000
TD	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	NA	0,000
Th	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	NA
***Percentage with Significant Differences (%)**	0	0	0	0	0	0	0	0	0	0

**The percentage with significant differences at all diversity orders* q *= 0–3 are the same, i.e., all zeros.*

TABLE 2 | Summary of the randomization tests from **Supplementary Table S7**: the pair-wise comparisons of the DAR parameter (D_{max}: potential diversity) among the ten DT sites of the human digestive microbiome (Genus level) for each diversity order q. The same coding scheme as in **Table 1** was used.

Digestive sites	BM	HP	KG	PT	SAL	Stool	SubP	SupP	TD	Th
BM	NA	0,100	0,000	0,101	0,101	0,000	0,101	0,101	0,101	0,101
HP	0,100	NA	0,100	0,111	0,111	0,000	0,111	0,111	0,111	0,111
KG	0,000	0,100	NA	0,111	0,111	0,000	0,111	0,111	0,111	0,111
PT	0,101	0,111	0,111	NA	0,111	0,100	0,111	0,011	0,000	0,000
SAL	0,101	0,111	0,111	0,111	NA	1,111	0,111	0,000	0,100	0,011
Stool	0,000	0,000	0,000	0,100	1,111	NA	0,111	1,111	0,111	0,100
SubP	0,101	0,111	0,111	0,111	0,111	0,111	NA	0,110	0,111	0,111
SupP	0,101	0,111	0,111	0,011	0,000	1,111	0,110	NA	0,111	0,011
TD	0,101	0,111	0,111	0,000	0,100	0,111	0,111	0,111	NA	0,000
Th	0,101	0,111	0,111	0,000	0,011	0,100	0,111	0,011	0,000	NA
***q* = 0 (%)**	0	0	0	0	11.1	22.2	0	11.1	0	0
***q* = 1 (%)**	77.8	88.9	77.8	66.7	77.8	66.7	100	66.7	77.8	55.6
***q* = 2 (%)**	0	66.7	66.7	55.6	66.7	44.4	88.9	77.8	55.6	55.6
***q* = 3 (%)**	66.7	66.7	66.7	66.7	77.8	44.4	88.9	77.8	66.7	66.7

community/metacommunity patterns on a larger scale, in our case, the inter-individual microbiome distribution in a human population cohort. A somewhat unique quality of this study is the analysis of intra-individual diversity distribution across the DT sites, using the same DAR tools as used for the inter-individual diversity scaling, which have been a norm of biogeography study in the studies of human microbiomes.

Practically, biogeography of human microbiome is of obvious importance for public health and personified medicine. For instance, understanding the inter-individual heterogeneity is essential for studying and implementing microbiome intervention treatments such as fecal transplantation for treating certain microbiome-associated diseases (Kump et al., 2013; Wei et al., 2015; Weingarden and Vaughn 2017; Ishikawa et al., 2017, 2018; Cohen et al., 2019). The inter-individual heterogeneity can help to identify/explain possible differences in treatment effects. Similarly, studying the intra-individual (along the human DT in this study) diversity scaling is also of important significance, for example, in choosing the optimum location of DT sites for treatment intervention. An additional advantage of our study is that we apply the same DAR approach to investigate both inter-individual and intra-DT (intra-individual) diversity scaling (across individual and across DT sites) of the human DT microbiomes, which makes the integrated analysis of both sources of heterogeneities (inter-individuals and intra-individual) implementable using the same set of parameters (such as diversity scaling rate z, potential diversity, and RIP).

In terms of the inter-individual microbial diversity scaling (**Supplementary Table S1** and **Table 1** and **2**), the diversity scaling parameter (z) of all 10 DT sites seems invariant with the site—the z-values of all sites did not show significant statistical differences. At genus level, the average scaling parameter (rate) (z) across 10 sites is 0.294 at diversity order $q = 0$ (species richness), 0.038 at $q = 1$ (Shannon entropy), 0.020 at $q = 2$ (Simpson index), and 0.014 at $q = 3$. The scaling rate of *species richness* (Hill number for $q = 0$) is nearly 10 times larger than those of other diversity orders, *e.g.,* community evenness measured with Shannon entropy. These results suggest that the inter-individual differences in DT microbiome diversity are primarily in the number of microbial genus—species richness, as demonstrated by much higher scaling rate, rather than in general community diversity as demonstrated by Hill

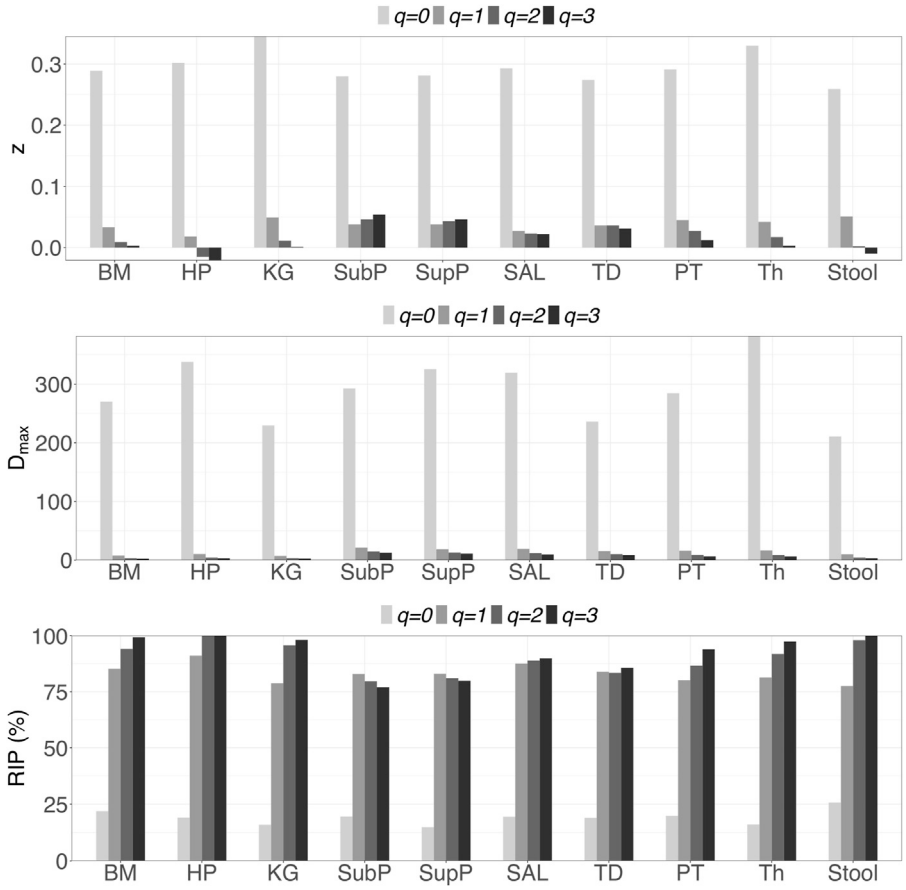


FIGURE 1 | Graphs of the three important profiles from inter-individual DAR models for each of the 10 DT microbiome sites at genus taxon level, including DAR profiles (*z*-*q* patterns), MAD profiles (D_{max}-*q* patterns), and RIP profiles (*RIP*-*q* patterns). Bar color-depth indicates diversity order. The *x*-axis shows the DT sites: buccal mucosa (BM), keratinized gingiva (KG), hard palate (HP), throat (Th), palatine tonsils (PT), tongue dorsum (TD) and saliva (Sal), supraginval (SupP), subgingival plaques (SubP), and stool (Stool).

numbers for $q > 1$. The inter-individual diversity scaling parameters (z) obtained in this study is also consistent with previous study by [Citations: Ma (2018a), Ecology and Evolution, Ma (2018b), Microbial Ecology], in which single gut microbiome diversity scaling was investigated.

Besides inter-individual diversity scaling parameter (z), another DAR parameter RIP (ratio of individual to population level diversity) also revealed that the critical differences between individuals lie in species richness (q = 0), rather than in community diversity ($q > 0$). The average RIP for species richness (q = 0) of 10 DT sites is 19.1% with standard error of 1.0 only, while RIP for general community diversity (q = 1–3) ranged from 83.1 to 93.4%. These RIP numbers indicates that an average individual can host approximately 20% of microbial genus owned by a whole population, while the microbial diversity of an individual may exceed 90% the total diversity of a population from which the individual comes from. Furthermore, all 10 DT exhibited very similar inter-individual diversity scaling as described above, which is evidenced by the rather small standard error of the average z and RIP across the 10 DT sites. To the best of our knowledge, noprevious studies have addressed the RIP of gut microbiomes.

We also systematically investigated the inter-individual diversity scaling on other four taxa including phylum, class, order, and family (**Supplementary Table S2-S5**). The scaling patterns are similar to the previously summarized genus-level scaling, but the diversity scaling parameter (z) generally decreases with the taxon level. That is, the inter-individual differences in diversity decreases with the higher taxonomic orders. This should be expected since higher diversity order such as phyla and classes are more general (rough) classifications and the similarity in diversity should certainly be higher at more general taxonomic scales. Higher similarity in diversity is equivalent to lower diversity scaling parameter (z), *i.e.,* slower scaling rate. To the best of our knowledge, this study should be the first one that studies diversity scaling at taxonomic level beyond species.

In terms of intra-DT diversity scaling patterns (**Supplementary Table S6** and **Figure 2**), at the genus level, the diversity scaling parameter (z) at various diversity orders (q = 0–3) is actually more similar that inter-individual scaling, which is indicated by the relatively

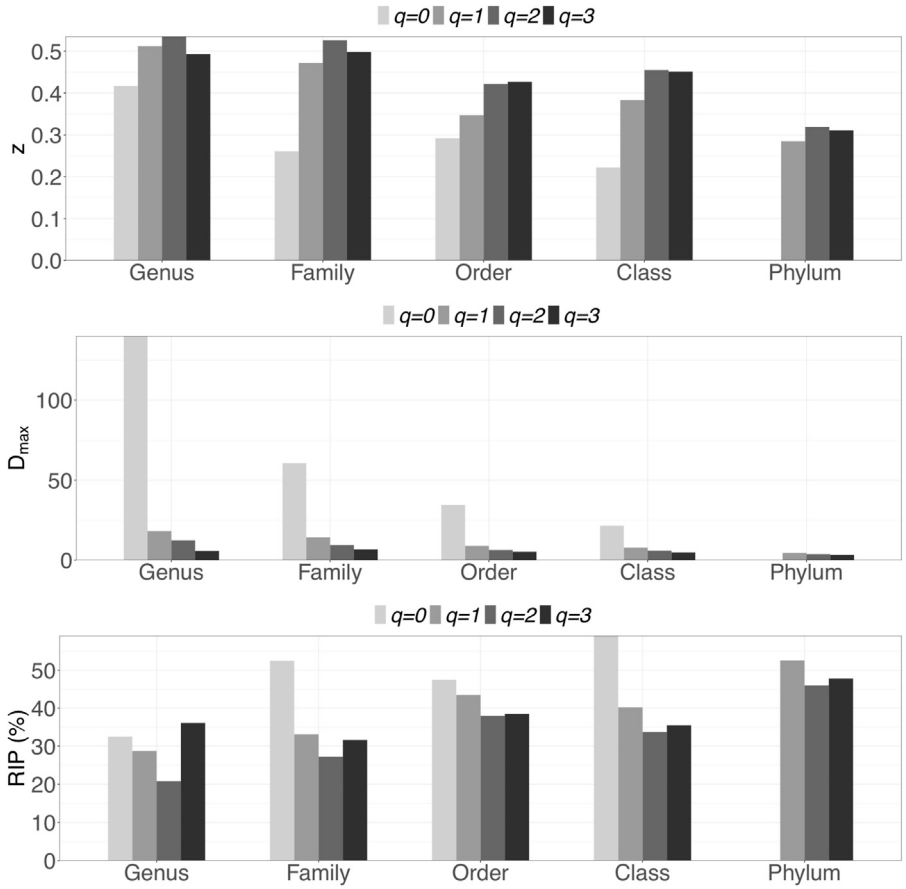


FIGURE 2 | Graphs of three important profiles from intra-individual DAR models for each of the five taxon levels, including DAR profiles (*z*-*q* patterns), MAD profiles (D_{max}-*q* patterns), and RIP profiles (*RIP*-*q* patterns). Bar color indicates diversity order.

narrow range of *z*-values {0.417 (q = 0), 0.512 (q = 1), 0.535 (q = 2), 0.493 (q = 3)}. Similarly, the RIP vector {32.5% (q = 0), 28.7% (q = 1), 20.8% (q = 2), 36.1 (q = 3)} also exhibited a range, compared with previous inter-individual. These findings suggest that intra-DT heterogeneity seems to be universally stronger than the inter-individual heterogeneity. This is obviously determined by the human biology, since for intra-DT diversity scaling, we are comparing "apples" and "oranges" (*e.g.*, oral site *vs* gut), while inter-individual diversity scaling was comparing the "apples from two apple trees". Therefore, comparing them may not be that meaningful. The important insight our study revealed is that 1) each DT site hosts approximately 1/5 to 1/3 of the whole DT diversity, and 2) there should be significantly overlaps (similarity) among the DT sites as inferred by (1). This high similarity in intra-DT diversity can be explained by the biological fact that DT is a continuum in which microbial dispersal occurs routinely. On the other hand, the heterogeneity in diversity scaling can be explained that the DT continuum is not homogenous either. In fact, the DT is differentiated as four different niches hosting some functionally different microbial species as revealed in the original study of Segata *et al.* (2012), upon the datasets of which our study is based.

Similar to the inter-individual diversity scaling, we also investigated the intra-DT diversity scaling on other four taxon levels (phylum, class, order, and family) (**Supplementary Table S6**) beyond genus level. The pattern is similar to previous inter-individual diversity scaling. That is, at higher taxonomic level, the difference becomes smaller or the similarity becomes larger, perhaps like using a telescope to observe remote landscapes. For example, at the taxonomic *class* level, the RIP for q = 0 was 59.1%, suggesting that an average individual can host approximately 60% of the microbial classes of a whole population.

A minor limitation of this study is that the species-level DAR analysis was missing given that the original raw sequencing reads reported in Segata *et al.* (2012) were only binned to genus and above taxon levels. Since in many cases, the species level (or 97% similarity level) OTUs are simply a number appended to genus name, and may be of limited biomedical significances. In the meantime, the annotations at higher taxon levels (genus, family, order, class, and phylum) should be rather stable, and the analyses

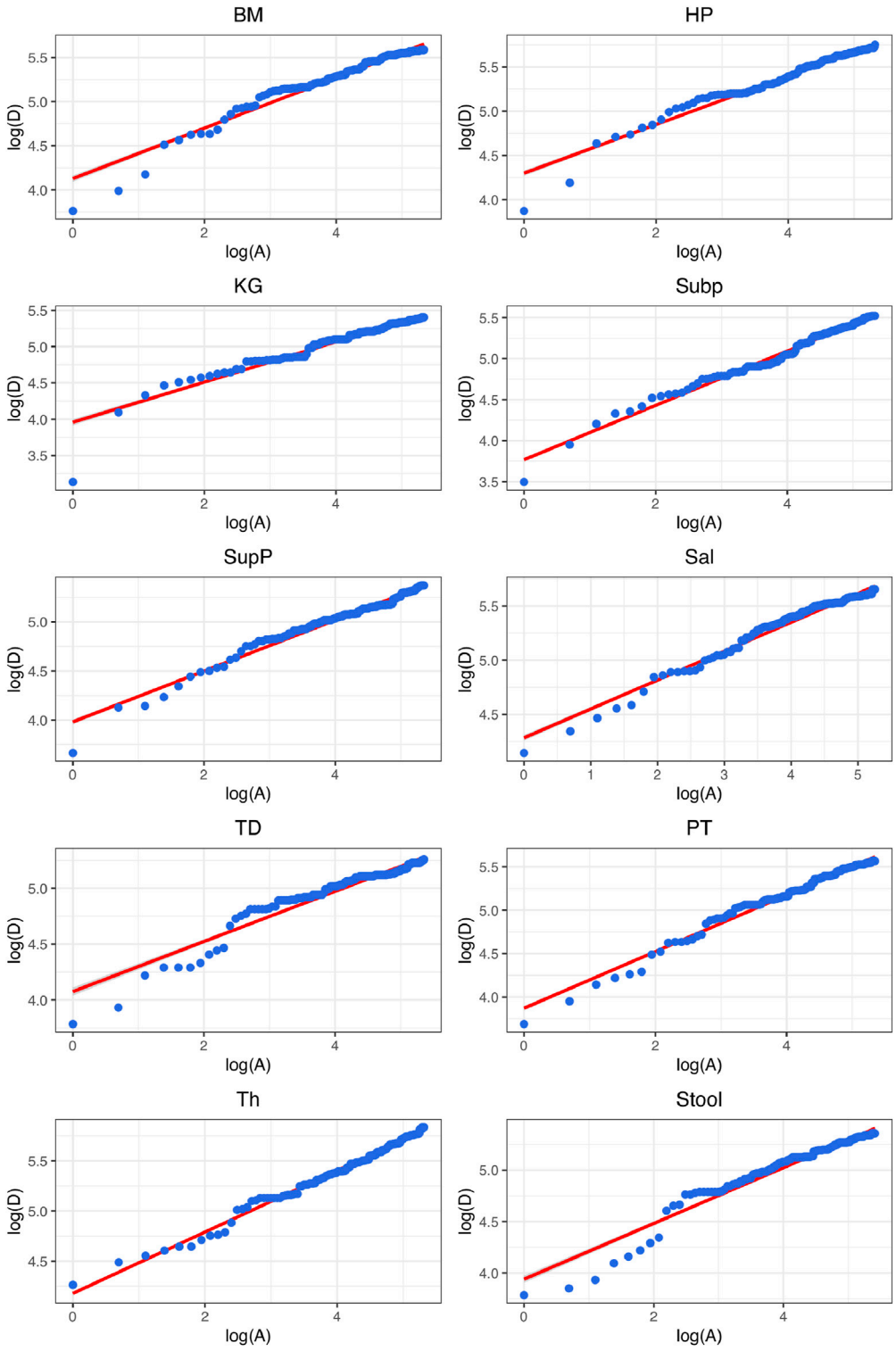


FIGURE 3 | Graphs of fitting the diversity-area relationship (DAR) power-law (PL) model (**Eq. 5**) for the gut microbiome at each of the 10 DT (digestive tract) site, at the genus taxon level, for diversity order $q = 0$ (*i.e.*, species richness).

of their diversity scaling can be more useful practically. In fact, to the best of our knowledge, this study should be the first comprehensive analysis of DT microbiomes at and above genus levels.

Finally, yet another minor limitation of this study is that we could not provide in-depth mechanistic interpretation of the observed patterns as revealed by DAR modeling analysis. On the one hand, DAR as an extension to classic SAR inherited both the merits and limitations of SAR. The SAR was discovered more than a century ago (Watson 1835) largely as an empirical relationship, and some scholars called it "collector's curve"—hinting that when one travels more regions (areas) he or she should be able to collective more species. In the early days of biology, this was indeed the case; botanists and zoologists (collectively known as naturalists, the most famous should be Charles Darwin) were, in the first place, bio-geographers and taxonomists. Arguably one of the most sophisticated mathematical tools they used was the graphing on coordinate papers with pencils, especially on the log-scales. The SAR graphs on coordinate papers would be straight lines (*e.g.*, **Figure 3** and **Eq. 5**), and transforming back to mathematical equation turned out to be a power-function (**Eq. 3**). During the last few decades, many ecologists (*e.g.,* Tjørve and Tjørve 2008; Harte et al., 2009; Sizling et al., 2011; Triantis et al., 2012) have tried to investigate the mechanisms underlying the observed SAR patterns. One of the most influential hypotheses is the self-similarity (scale invariance) hypothesis (*e.g.,* Harte et al., 2009; Sizling et al., 2011), which was also adopted by Ma (2018a) when he extended the classic SAR to general DAR. Interested readers should refer to Harte et al. (2009), Sizling et al. (2011), Ma (2018a) for mechanistic discussion on the SAR and DAR patterns. From a practical perspective, these mechanistic discussions are relatively less relevant, given that SAR has been considered as one of the most important models in conservation biology and biogeography. In our opinion, the DAR, which extends classis SAR from species richness to general diversity metrics, should have equally important applications in microbial biogeography and biogeography in general.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

HC and BY conducted the data analysis; LD and ZM designed the study; HC, LD, and ZM wrote and revised the manuscript; QL and XX prepared the figures. All authors approved the submission.

FUNDING

This research received funding from the following sources: the State Key Laboratory of Genetic Resources and Evolution (#GREKF21-06#, GREKF20-06, GREKF19-07) and Yunnan Province Local University (Part) Basic Research for Youths (No.2019FH001-117 and No. 202001BA070001-100. National Science Foundation of China (Grant #31970116); Cloud-Ridge Industry Technology Leader Grant from Yunnan Development and Reform Commission.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.724661/full#supplementary-material

REFERENCES

Chao, A., Chiu, C.-H., and Hsieh, T. C. (2012). Proposing a Resolution to Debates on Diversity Partitioning. *Ecology* 93 (9), 2037–2051. doi:10.1890/11-1817.1

Chao, A., Chiu, C.-H., and Jost, L. (2014a). Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures through Hill Numbers. *Annu. Rev. Ecol. Evol. Syst.* 45, 297–324. doi:10.1146/annurev-ecolsys-120213-091540

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., et al. (2014b). Rarefaction and Extrapolation with Hill Numbers: A Framework for Sampling and Estimation in Species Diversity Studies. *Ecol. Monogr.* 84, 45–67. doi:10.1890/13-0133.1

Cohen, L. J., Cho, J. H., Gevers, D., and Chu, H. (2019). Genetic Factors and the Intestinal Microbiome Guide Development of Microbe-Based Therapies for Inflammatory Bowel Diseases. *Gastroenterology* 156 (8), 2174–2189. doi:10.1053/j.gastro.2019.03.017

Collingridge, D. S. (2013). A Primer on Quantitized Data Analysis and Permutation Testing. *J. Mixed Methods Res.* 7 (1), 81–97. doi:10.1177/1558689812454457

Connor, E. F., and McCoy, E. D. (1979). The Statistics and Biology of the Species-Area Relationship. *The Am. Naturalist* 113, 791–833. doi:10.1086/283438

Drakare, S., Lennon, J. J., and Hillebrand, H. (2006). The Imprint of the Geographical, Evolutionary and Ecological Context on Species-Area Relationships. *Ecol. Lett.* 9, 215–227. doi:10.1111/j.1461-0248.2005.00848.x

Harte, J., Smith, A. B., and Storch, D. (2009). Biodiversity Scales from Plots to Biomes with a Universal Species-Area Curve. *Ecol. Lett.* 12, 789–797. doi:10.1111/j.1461-0248.2009.01328.x

He, F., and Hubbell, S. P. (2011). Species-area Relationships Always Overestimate Extinction Rates from Habitat Loss. *Nature* 473, 368–371. doi:10.1038/nature09985

Helmus, M. R., Mahler, D. L., and Losos, J. B. (2014). Island Biogeography of the Anthropocene. *Nature* 513, 543–546. doi:10.1038/nature13739

Hill, M. O. (1973). Diversity and Evenness: A Unifying Notation and its Consequences. *Ecology* 54, 427–432. doi:10.2307/1934352

HMP Consortium (Human Microbiome Project Consortium) (2012a). A Framework for Human Microbiome Research. *Nature* 486, 215–221. doi:10.1038/nature11209

HMP Consortium (Human Microbiome Project Consortium) (2012b). Structure, Function and Diversity of the Healthy Human Microbiome. *Nature* 486, 207–214. doi:10.1038/nature11234

Ishikawa, D., Sasaki, T., Osada, T., Kuwahara-Arai, K., Haga, K., Shibuya, T., et al. (2017). Changes in Intestinal Microbiota Following Combination Therapy with Fecal Microbial Transplantation and Antibiotics for Ulcerative Colitis. *Inflamm. Bowel Dis.* 23 (1), 116–125. doi:10.1097/MIB.0000000000000975

Ishikawa, D., Sasaki, T., Takahashi, M., Kuwahara-Arai, K., Haga, K., Ito, S., et al. (2018). The Microbial Composition of Bacteroidetes Species in Ulcerative Colitis Is Effectively Improved by Combination Therapy with Fecal Microbiota Transplantation and Antibiotics. *Inflamm. Bowel Dis.* 24 (12), 2590–2598. doi:10.1093/ibd/izy266

Kump, P. K., Gröchenig, H.-P., Lackner, S., Trajanoski, S., Reicht, G., Hoffmann, K. M., et al. (2013). Alteration of Intestinal Dysbiosis by Fecal Microbiota Transplantation Does Not Induce Remission in Patients with Chronic Active Ulcerative Colitis. *Inflamm. Bowel Dis.* 19 (10), 2155–2165. doi:10.1097/MIB.0b013e31829ea325

Lomolino, M. V. (2000). Ecology's Most General, yet Protean 1 Pattern: the Species-Area Relationship. *J. Biogeogr.* 27, 17–26. doi:10.1046/j.1365-2699.2000.00377.x

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., and Knight, R. (2012). Diversity, Stability and Resilience of the Human Gut Microbiota. *Nature* 489, 220–230. doi:10.1038/nature11550

Ma, Z., and Li, L. (2019). Semen Microbiome Biogeography: An Analysis Based on a Chinese Population Study. *Front. Microbiol.* 9, 3333. doi:10.3389/fmicb.2018.03333

Ma, Z. S. (2018a). DAR (Diversity-Area Relationship): Extending Classic SAR (Species-Area Relationship) for Biodiversity and Biogeography Analyses. *Ecol. Evol.* 8 (20), 10023–10038. doi:10.1002/ece3.4425

Ma, Z. (2018b). Sketching the Human Microbiome Biogeography with DAR (Diversity-Area Relationship) Profiles. *Microb. Ecol.* 77 (3), 821–838. doi:10.1007/s00248-018-1245-6

Ma, Z. (2019). A New DTAR (Diversity-Time-Area Relationship) Model Demonstrated with the Indoor Microbiome. *J. Biogeogr.* 46 (1), 2024–2041. doi:10.1111/jbi.13636

Preston, F. W. (1960). Time and Space and the Variation of Species. *Ecology* 41, 611–627. doi:10.2307/1931793

Rosenzweig, M. L. (1995). *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.

Segata, N., Haake, S., Mannon, P., Lemon, K. P., Waldron, L., Gevers, D., et al. (2012). Composition of the Adult Digestive Tract Bacterial Microbiome Based on Seven Mouth Surfaces, Tonsils, Throat and Stool Samples. *Genome Biol.* 13 (6), R42. doi:10.1186/gb-2012-13-6-r42

Sizling, A. L., Kunin, W. E., Sizlingová, E., Reif, J., and Storch, D. (2011). Between Geometry and Biology: the Problem of Universality of the Species-Area Relationship. *Am. Nat.* 178 (5), 602–611. doi:10.1086/662176

Storch, D., Keil, P., and Jetz, W. (2012). Universal Species-Area and Endemics-Area Relationships at continental Scales. *Nature* 488, 78–81. doi:10.1038/nature11226

Tjørve, E. (2009). Shapes and Functions of Species–Area Curves (II): A Review of New Models and Parameterizations. *J. Biogeogr.* 36, 1435–1445. doi:10.1111/j.1365-2699.2009.02101.x

Tjørve, E., and Tjørve, K. M. C. (2008). The Species–Area Relationship, Self-Similarity, and the True Meaning of the Z-Value. *Ecology* 89, 3528–3533. doi:10.1890/07-1685.1

Triantis, K. A., Guilhaumon, F., and Whittaker, R. J. (2012). The Island Species-Area Relationship: Biology and Statistics. *J. Biogeogr.* 39, 215–231. doi:10.1111/j.1365-2699.2011.02652.x

Watson, H. C. (1835). *Remarks on Geographic Distribution of British Plants*. London: Longman, Rees, Orme, Brown, Green and Longman.

Wei, Y., Zhu, W., Gong, J., Guo, D., Gu, L., Li, N., et al. (2015). Fecal Microbiota Transplantation Improves the Quality of Life in Patients with Inflammatory Bowel Disease. *Gastroenterol. Res. Pract.* 2015, 517597. doi:10.1155/2015/517597

Weingarden, A. R., and Vaughn, B. P. (2017). Intestinal Microbiota, Fecal Microbiota Transplantation, and Inflammatory Bowel Disease. *Gut Microbes* 8 (3), 238–252. doi:10.1080/19490976.2017.1290757

Whittaker, R. J., and Triantis, K. A. (2012). The Species-Area Relationship: An Exploration of that 'most General, yet Protean Pattern'1. *J. Biogeogr.* 39, 623–626. doi:10.1111/j.1365-2699.2012.02692.x

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.

frontiers in Microbiology

PERSPECTIVE
published: 04 October 2021
doi: 10.3389/fmicb.2021.718082

The Macroevolutionary Consequences of Niche Construction in Microbial Metabolism

Djordje Bajić[1,2], María Rebolleda-Gómez[1,2,3], Martha M. Muñoz[1] and Álvaro Sánchez[1,2]**

[1]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, United States, [2]Microbial Sciences Institute, Yale University, West Haven, CT, United States, [3]Department of Ecology and Evolutionary Biology, University of California Irvine, Irvine, CA, United States

OPEN ACCESS

Edited by:
Alejandro Couce, Polytechnic University of Madrid, Spain

Reviewed by:
Jonathan Friedman, Hebrew University of Jerusalem, Israel
Angus Buckling, University of Exeter, United Kingdom
Laila Pamela Partida-Martinez, Centro de Investigaciones y Estudios Avanzados, Instituto Politécnico Nacional de México (CINVESTAV), Mexico

***Correspondence:**
Djordje Bajić
dordje.bajic@yale.edu
Álvaro Sánchez
alvaro.sanchez@yale.edu

Specialty section:
This article was submitted to Evolutionary and Genomic Microbiology, a section of the journal Frontiers in Microbiology

Received: 31 May 2021
Accepted: 20 August 2021
Published: 04 October 2021

Citation:
Bajić D, Rebolleda-Gómez M, Muñoz MM and Sánchez Á (2021) The Macroevolutionary Consequences of Niche Construction in Microbial Metabolism. Front. Microbiol. 12:718082. doi: 10.3389/fmicb.2021.718082

Microorganisms display a stunning metabolic diversity. Understanding the origin of this diversity requires understanding how macroevolutionary processes such as innovation and diversification play out in the microbial world. Metabolic networks, which govern microbial resource use, can evolve through different mechanisms, e.g., horizontal gene transfer or *de novo* evolution of enzymes and pathways. This process is governed by a combination of environmental factors, selective pressures, and the constraints imposed by the genetic architecture of metabolic networks. In addition, many independent results hint that the process of niche construction, by which organisms actively modify their own and each other's niches and selective pressures, could play a major role in microbial innovation and diversification. Yet, the general principles by which niche construction shapes microbial macroevolutionary patterns remain largely unexplored. Here, we discuss several new hypotheses and directions, and suggest metabolic modeling methods that could allow us to explore large-scale empirical genotype-phenotype-(G-P)-environment spaces in order to study the macroevolutionary effects of niche construction. We hope that this short piece will further stimulate a systematic and quantitative characterization of macroevolutionary patterns and processes in microbial metabolism.

Keywords: macroevolution, metabolism, innovation, diversification, niche construction, genotype-phenotype (G-P) map, genotype-by-environment (G × E) interaction, non-commutative epistasis

INTRODUCTION

Prokaryotes exhibit by far the most diverse collection of metabolisms on earth. Disentangling the mechanisms by which such diversity arises is paramount for understanding both the emergence of complex life and the structure and function of modern microbial ecosystems.

Our knowledge about the history of life on earth contains numerous examples suggesting that the process of niche construction might play a central role in diversification. An obvious one is the early appearance of autotrophic metabolism, which profoundly transformed the biosphere by generating complex, energy-rich carbon molecules and releasing oxygen to the atmosphere, creating new ecological opportunities (Schirrmeister et al., 2013; Chen et al., 2020). Many additional examples exist in nature, where niche construction has been observed to play important roles, from the early examples described by Darwin (1881) in his work on earthworms to diatoms or beavers (Odling-Smee et al., 2013). Although, theoretical work has anticipated

numerous ways, in which niche construction might impact evolutionary outcomes (Laland et al., 1999; Silver and Di Paolo, 2006; Krakauer et al., 2009), most of these predictions remain empirically untested. From a general principles standpoint, many open questions remain: How does niche construction itself evolve, and what are the principles and mechanisms that govern it? How does it depend on the external environment, on the architecture of metabolic genotypes or on metabolic strategies? How does niche construction affect macroevolutionary processes such as innovation and diversification?

The short-term, microevolutionary consequences of niche construction have been extensively characterized, both theoretically and empirically (Laland et al., 1999; Odling-Smee et al., 2013). The most paradigmatic of such effects are eco-evolutionary dynamics, which arise because constructed environments (and their effects on selective pressures) depend, within certain limits, on the abundance of the organisms constructing them. This leads to a dynamic coupling between frequency and density-dependent selection, which occur in similar timescales. Because the nature of built environments in microbes is often determined by single genes or mutations, niche construction can link the fate of specific alleles to the current, instantaneous composition of a population (Sanchez and Gore, 2013; Chen et al., 2014). When several species are involved, niche construction can also combine with other ecological interactions to generate more complex phenomena, such as the coexistence of three or more species through intransitive interactions (e.g., rock-paper-scissors; Kerr et al., 2002).

In contrast to microevolution, the macroevolutionary consequences of niche construction, such as microbial innovation and diversification, have been less explored. A rare experimental example of a potentially macroevolutionary event is the appearance of aerobic citrate utilization in *Escherichia coli* in the Long Term Evolution Experiment (LTEE). Recent findings have shown that the two main mutations leading to this innovation, the aerobic expression of dctA and citT, were intimately linked to an eco-evolutionary interaction mediated by the release of metabolites to the environment (Bajić et al., 2018; de Visser et al., 2018). Furthermore, one of the main potentiating mutations that helped "prepare" the genetic background for the evolution of citrate use (*gltA*), likely achieved fixation because of its beneficial effect on acetate, a constructed niche (Quandt et al., 2015). These observations suggest that niche construction might play a key role in microbial metabolic diversification. More broadly, they showcase the potential of microbial experiments to illuminate the mechanisms and the genetic basis underlying macroevolutionary patterns.

At the same time, experiments also have important limitations. In the LTEE, only one out of 12 *E. coli* evolution lines gained the ability to use citrate, and did so only after ~30,000 generations (~20 years of experiment). This illustrates that innovation and exploration of untapped ecological opportunities still depends on historically contingent, and thus rare, combinations of mutations (Blount et al., 2008). Correspondingly, "blind" evolutionary explorations of genotype space still require timescales approaching the limits of what is experimentally feasible, even for organisms with some of the shortest generation times on Earth.

A promising alternative is provided by genome-scale metabolic models, which offer us the possibility to rapidly explore large regions of metabolic genotype-environment space. Using genome-inferred metabolic networks, these models are able to quite accurately simulate the growth of real organisms *in silico*, providing us mechanistic insight into the function of biologically realistic genotype-phenotype-(G-P)-fitness maps. Genome-scale metabolic models have been already successfully applied, for instance, to gain insight into long-term phenotypic evolution in microbes (Plata et al., 2015), study the genomic basis of metabolic innovations (Barve and Wagner, 2013; Hosseini et al., 2015) and explore intriguing origin-of-life scenarios (Goldford et al., 2017). Beyond purely computational studies, genome-scale metabolic models have also proven a powerful tool for experiment design. An astonishing example was the recent obtention of an *E. coli* strain capable of autotrophic metabolism (Gleizer et al., 2019). This achievement used metabolic modeling to predict what new reactions might be needed by *E. coli* to acquire carbon fixation capabilities. Once these reactions were included, experimental evolution took care of integrating them in the regulatory network, allowing *E. coli* to start fixing CO_2 and become autotrophic in a relatively short time. The potential niche construction consequences of such metabolic innovation are self-evident.

We thus believe that, in combination with experiments, genome-scale metabolic models can be an invaluable tool to explore macroevolutionary patterns in microbes. In this short piece, we lay out several future directions, focusing in particular on the effects of niche construction or, more broadly, the two-way interaction between genotype and environment.

NICHE CONSTRUCTION AND THE PARALLEL EXPLORATION OF FITNESS LANDSCAPES

In their landmark work almost 40 years ago, Levins and Lewontin (1985) noted that evolution does not only proceed as a mere adaptation of organisms to the external environment. In addition, as they adapt, organisms also modify the environment, potentially affecting their own selective pressures (Levins and Lewontin, 1985). Evolution becomes then better described as a "dialectic" process, in which genotype and environment perpetually modify each other. This logic, however, challenges the predictions of many established theories that only consider the adaptation of organisms to the "external" environment. For instance, fitness landscapes have been widely used as both a conceptual device and as a tool for predicting evolution (de Visser et al., 2018; Gorter et al., 2018). But if fitness landscapes constantly "deform" during evolution, their utility would be severely compromised (Doebeli et al., 2017).

To what extent, then, is niche construction an ubiquitous process, and to what extent is it able to influence evolutionary

patterns and outcomes? Answering these key questions inevitably requires turning our attention to empirical systems. In recent work, we used constraint-based metabolic modeling to systematically map the diversity of constructed niches on a metabolic genotype space, and their evolutionary consequences (Bajić et al., 2018; de Visser et al., 2018). We found that when a newly constructed niche becomes available, as a result of a mutation, multiple subsequent mutations (often epistatic to each other) are typically needed to take advantage of this new niche. This led to a surprising conclusion: while in the shorter term "static" fitness landscapes are typically predictive, the deformations gain importance as changes in both the environment and the population genotypes accumulate. In addition, recent studies have identified that large numbers of metabolites can be secreted by microbes, often at no cost (de Visser et al., 2018; Magdalena and Wagner, 2018; Pacheco et al., 2019). These results point to the possibility that niche construction might play a preeminent role in evolutionary processes that typically occur over longer timescales, possibly including macroevolutionary patterns such as phenotypic divergence and diversification.

A particularly interesting possibility is that, by enabling the parallel exploration of different fitness landscapes, niche construction could facilitate bridging fitness valleys (Steinberg and Ostermeier, 2016), including those leading to innovations. An illuminating hint of how this might happen comes from a recent work showing that the emergence of complex innovations can be facilitated by stepwise metabolic niche expansion (Szappanos et al., 2016). In order to reach complex innovations requiring two or more mutations, organisms capitalize on more accessible "stepping stone" innovations, allowing them to navigate genotype-space by switching between environments. It is easy to imagine how such "stepping-stones to innovation" could be provided through niche construction (**Figure 1**). In this way, niche construction could blur the lines between ecological and mutation-order speciation (Schluter, 2009), making them contingent on each other. Exploring to what extent can constructed niches open evolutionary paths toward otherwise inaccessible ecological opportunities could provide a mechanistic explanation to the hypothesis that "diversity beget diversity" (Whittaker, 1972), which has been recently shown to apply in microbiomes in some conditions (Madi et al., 2020).

LINKING NICHE CONSTRUCTION TO ADAPTIVE RADIATION

If niche construction facilitates innovation and diversification, this could have profound consequences for our understanding of adaptive radiations, a process considered integral to ecological and phenotypic diversity (Simpson, 1953; Schluter, 2000). Adaptive radiation is the process of phenotypic diversification of organisms into forms that fill different available ecological niches. Phenotypic novelties can facilitate adaptive radiation by allowing organisms to interact with their environments in new ways, in turn generating novel ecological opportunities upon which natural selection can act and prompting adaptive evolution (Simpson, 1953; Stroud and Losos, 2016; Erwin, 2017). The connection between innovation and adaptive radiation has been extensively documented both within the fossil record and using diverse empirical and experimental systems (Schluter, 2000; Losos, 2010; Yoder et al., 2010). Notable examples include the evolution of adhesive silk in spiders, which enhanced prey capture (Bond and Opell, 1998), alternate photosynthetic pathways in desert plants that reduce water loss (Silvestro et al., 2014), and adhesive toe pads that unlock access to arboreal niches in anole lizards (Burress and Muñoz, 2021). However, the study of ecological opportunity has been largely structured around the evolution of features that facilitate access to novel peaks in the adaptive landscape. Comparatively less focus has been given to the role that organisms serve as arbiters of available niches, for example, by constructing new ones or bridging access to new peaks (Emerson and Kolm, 2005; Erwin, 2005; Braakman et al., 2017). In the microbial realm, adaptive radiations of phenotypically diverse lineages can be obtained through experimental evolution, and often involve niche construction (Rainey and Travisano, 1998; Friesen et al., 2004; Le Gac et al., 2008; Kassen, 2009; Schick and Kassen, 2018). However, the scope of these studies in terms of environmental complexity, genetic diversity, and timescales is rather limited compared to plausible scenarios in nature (Consuegra et al., 2017).

A key question is to what extent are novelty, diversification and adaptive radiation in microbes constrained and directed by the spectrum of available genetic variation as opposed to just ecological opportunity (Schluter, 2009; Kassen, 2019). As shown by Schluter (1996), adaptive radiation often proceeds along "genetic lines of least resistance," meaning that evolution occurs in the direction, where most genetic variation is available. If niche construction significantly alters microbial fitness landscapes, it could quantitatively bias the observed patterns of adaptive radiation by changing the distribution of fitness effects, and consequently altering the location of those lines of least resistance. Furthermore, complex genetic architecture (e.g., epistasis) also imposes strong constraints on adaptation (Weinreich et al., 2005, 2006), but this question has been scarcely explored in the context of adaptive radiation, and more generally, macroevolutionary patterns. In the case of bacterial metabolism, the availability of detailed empirical genotype-phenotype maps and the possibility of their prediction has already been very useful to show how the epistatic architecture of metabolic networks can fundamentally constrain innovation. For example, the partial overlapping between the pathways used for different nutrients results in many innovations being readily accessible as a byproduct to the adaptation to a given nutrient (Barve and Wagner, 2013). Exploring in detail how epistasis shapes the response of populations to constructed niches, and how epistatic interactions themselves change with the environment, represents an interesting opportunity for future research.

Furthermore, if niche construction provides "stepping stones" toward other innovations, it could also contribute qualitatively to some of the most iconic patterns of adaptive radiation such as rapid diversification. If we consider that each new adaptation

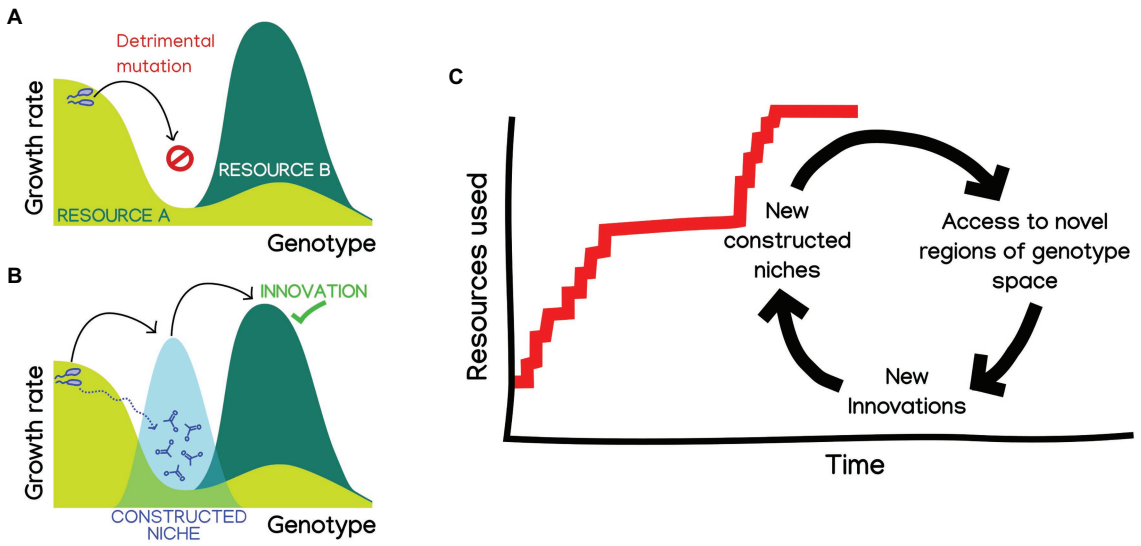


FIGURE 1 | Niche construction could facilitate metabolic innovation and adaptive radiation in microbes. **(A)** Complex innovations, requiring two or more mutations, are hard to reach through adaptive evolution, needing the fixation of neutral or deleterious mutations. **(B)** Niche construction could provide "bridge" environments, which are known to facilitate the navigability of genotype space by adaptive evolution, making innovations readily accessible (Steinberg and Ostermeier, 2016; Szappanos et al., 2016; Pál and Papp, 2017). **(C)** By extension, we can imagine a hypothetical evolution of the number of resources that a microbial population can use, as a succession of innovation events. Changes in constructed niches could make new innovations accessible, which will spur adaptation and thus also building new niches. A hypothetical result of this cyclic process could be rapid bursts of innovation, a punctuated pattern characteristic of adaptive radiation.

bears the potential to transform the environment, we could imagine a "cycle" in which innovation results in new constructed environments, which in turn open up adaptive paths to further downstream innovations, potentially leading to rapid diversification patterns. Scenarios similar to the "stepping stones" model in adaptive radiation have been hinted at by simulations (Sneppen et al., 1995), but they have so far remained in the theoretical realm. Finally, an intriguing question is how niche construction plays out when considering more realistic genotype-phenotype spaces, where the genetic accessibility of phenotypes can be organized in asymmetric and nonrandom topologies (more precisely, "pre-topologies"; Fontana and Schuster, 1998; Stadler et al., 2001; Erwin, 2017). Together, exploring to what extent could niche construction facilitate innovation and release adaptive radiation from the yoke of genetic constraint is a fascinating future direction.

DISCUSSION

In this piece, we argued that the combination of genome-scale metabolic modeling with experiments presents a great opportunity to tackle the role of different evolutionary forces, and niche construction in particular, in microbial macroevolution. Recently, platforms such as "Computation of Microbial Ecosystems in Time and Space" (COMETS; Dukovski et al., 2020) are extending the range of possibilities of genome-scale metabolic models by enabling us to simulate evolution in the context of multispecies ecosystems. COMETS combines population dynamics with a realistic, empirically calibrated genotype-phenotype map that is also environment-sensitive, where mutations can randomly appear as either metabolic reaction deletions or additions (e.g., through horizontal gene transfer) or by random changes in the maximum fluxes through each reaction. Importantly, niche construction emerges naturally in COMETS, as it predicts phenotypes such as secretion of metabolites. This offers an unique opportunity to explore evolution (including macroevolution) with mechanistic insight, allowing us to understand biological processes at lower levels of organization without isolating them from the eco-evolutionary processes in which they are embedded (Bergelson et al., 2021; van Tatenhove-Pel et al., 2021). Furthermore, COMETS also offers sophisticated spatial capabilities. These could be key in understanding the effects of metabolic niche construction (Maynard et al., 2017; van Tatenhove-Pel et al., 2021) as well as help understand observations in natural environments, as well as in emerging experimental platforms such as ecoFABs (Sasse et al., 2019).

One of the hurdles in this path is our current lack of understanding of the relationship between the genotype and the organisms' effects on the environment, particularly through secretion of metabolites. The release of some compounds, such as fermentation byproducts, is well understood (and predicted by metabolic models; Basan et al., 2015; Mori et al., 2016), suggesting that we could generalize this finding and link genotype to build niches in a systematic way. However, exometabolomic analyses typically identify complex metabolite mixtures (Paczia et al., 2012), whose origin is still poorly understood. Understanding the determinants of metabolic secretions represents one of the main current limitations for building a predictive theory of microbial

ecology and evolution, including macroevolutionary processes of qualitative change.

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.

AUTHOR CONTRIBUTIONS

DB, MR-G, MM, and AS have contributed original ideas and perspectives presented in the manuscript. DB wrote the first draft. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a young investigator award from the Human Frontier Science Program (RGY0077/2016), by a Packard Fellowship from the David and Lucile Packard foundation, and by the National Institutes of Health through grant 1R35 GM133467-01 to AS, and by a Templeton Foundation Grant (61866) to MM. MR-G acknowledges support from the Donnelly Fellowship from Yale University.

REFERENCES

Bajić, D., Vila, J. C. C., Blount, Z. D., and Sánchez, A. (2018). On the deformability of an empirical fitness landscape by microbial evolution. *Proc. Natl. Acad. Sci. U. S. A.* 115, 11286–11291. doi: 10.1073/pnas.1808485115

Barve, A., and Wagner, A. (2013). A latent capacity for evolutionary innovation through exaptation in metabolic systems. *Nature* 500, 203–206. doi: 10.1038/nature12301

Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. R., et al. (2015). Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* 528, 99–104. doi: 10.1038/nature15765

Bergelson, J., Kreitman, M., Petrov, D. A., Sanchez, A., and Tikhonov, M. (2021). Functional biology in its natural context: A search for emergent simplicity. *Elife 10*. doi: 10.7554/eLife.67646

Blount, Z. D., Borland, C. Z., and Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 7899–7906. doi: 10.1073/pnas.0803151105

Bond, J. E., and Opell, B. D. (1998). Testing adaptive radiation and key innovation hypotheses in spiders. *Evolution* 52, 403–414. doi: 10.1111/j.1558-5646.1998.tb01641.x

Braakman, R., Follows, M. J., and Chisholm, S. W. (2017). Metabolic evolution and the self-organization of ecosystems. *Proc. Natl. Acad. Sci. U. S. A.* 114, E3091–E3100. doi: 10.1073/pnas.1619573114

Burress, E. D., and Muñoz, M.M. (2021). Ecological opportunity from innovation, not islands, drove the anole lizard adaptive radiation. *Syst. Biol*. doi: 10.1093/sysbio/syab031 [Epub ahead of print].

Chen, A., Sanchez, A., Dai, L., and Gore, J. (2014). Dynamics of a producer-freeloader ecosystem on the brink of collapse. *Nat. Commun.* 5:3713. doi: 10.1038/ncomms4713

Chen, S.-C., Sun, G.-X., Yan, Y., Konstantinidis, K. T., Zhang, S.-Y., Deng, Y., et al. (2020). The great oxidation event expanded the genetic repertoire of arsenic metabolism and cycling. *Proc. Natl. Acad. Sci. U. S. A.* 117, 10414–10421. doi: 10.1073/pnas.2001063117

Consuegra, J., Plucain, J., Gaffé, J., Hindré, T., and Schneider, D. (2017). Genetic basis of exploiting ecological opportunity during the long-term diversification of a bacterial population. *J. Mol. Evol.* 85, 26–36. doi: 10.1007/s00239-017-9802-z

de Visser, J. A. G. M., Elena, S. F., Fragata, I., and Matuszewski, S. (2018). The utility of fitness landscapes and big data for predicting evolution. *Heredity* 121, 401–405. doi: 10.1038/s41437-018-0128-4

Darwin, C. (1881). *The Formation of Vegetable Mould through the Action of Worms, with Observations on their Habits.* London: John Murray.

Doebeli, M., Ispolatov, Y., and Simon, B. (2017). Towards a mechanistic foundation of evolutionary theory. *Elife* 6:e23804. doi: 10.7554/eLife.23804

Dukovski, I., Bajić, D., Chacón, J.M., Quintin, M., Vila, J.C.C., Sulheim, S., et al. (2020). A metabolic modeling platform for the computation of microbial ecosystems in time and space (COMETS). *Nature Protocols (Preprint)*. doi: 10.1038/s41596-021-00593-3

Emerson, B. C., and Kolm, N. (2005). Species diversity can drive speciation. *Nature* 434, 1015–1017. doi: 10.1038/nature03450

Erwin, D. H. (2005). Macroevolution. Seeds of diversity. *Science* 308, 1752–1753. doi: 10.1126/science.1113416

Erwin, D. H. (2017). The topology of evolutionary novelty and innovation in macroevolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372:20160422. doi: 10.1098/rstb.2016.0422

Fontana, W., and Schuster, P. (1998). Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. *J. Theor. Biol.* 194, 491–515. doi: 10.1006/jtbi.1998.0771

Friesen, M. L., Saxer, G., Travisano, M., and Doebeli, M. (2004). Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in *Escherichia coli*. *Evolution* 58, 245–260. doi: 10.1111/j.0014-3820.2004.tb01642.x

Gleizer, S., Ben-Nissan, R., Bar-On, Y. M., Antonovsky, N., Noor, E., Zohar, Y., et al. (2019). Conversion of *Escherichia coli* to generate all biomass carbon from CO_2. *Cell* 179, 1255.e12–63.e12. doi: 10.1016/j.cell.2019.11.009

Goldford, J. E., Hartman, H., Smith, T. F., and Segrè, D. (2017). Remnants of an ancient metabolism without phosphate. *Cell* 168, 1126.e9–34.e9. doi: 10.1016/j.cell.2017.02.001

Gorter, F. A., Aarts, M. G. M., Zwaan, B. J., and de Visser, A. J. G. M. (2018). Local fitness landscapes predict yeast evolutionary dynamics in directionally changing environments. *Genetics* 208, 307–322. doi: 10.1534/genetics.117.300519

Hosseini, S.-R., Barve, A., and Wagner, A. (2015). Exhaustive analysis of a genotype space comprising 10(15) central carbon metabolisms reveals an organization conducive to metabolic innovation. *PLoS Comput. Biol.* 11:e1004329. doi: 10.1371/journal.pcbi.1004329

Kassen, R. (2009). Toward a general theory of adaptive radiation: insights from microbial experimental evolution. *Ann. N. Y. Acad. Sci.* 1168, 3–22. doi: 10.1111/j.1749-6632.2009.04574.x

Kassen, R. (2019). Experimental evolution of innovation and novelty. *Trends Ecol. Evol.* 34, 712–722. doi: 10.1016/j.tree.2019.03.008

Kerr, B., Riley, M. A., Feldman, M. W., and Bohannan, B. J. M. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 418, 171–174. doi: 10.1038/nature00823

Krakauer, D. C., Page, K. M., and Erwin, D. H. (2009). Diversity, dilemmas, and monopolies of niche construction. *Am. Nat.* 173, 26–40. doi: 10.1086/593707

Laland, K. N., Odling-Smee, F. J., and Feldman, M. W. (1999). Evolutionary consequences of niche construction and their implications for ecology. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10242–10247. doi: 10.1073/pnas.96.18.10242

Le Gac, M., Brazas, M. D., Bertrand, M., Tyerman, J. G., Spencer, C. C., Hancock, R. E. W., et al. (2008). Metabolic changes associated with adaptive diversification in *Escherichia coli*. *Genetics* 178, 1049–1060. doi: 10.1534/genetics.107.082040

Levins, R., and Lewontin, R. (1985). *The Dialectical Biologist.* Cambridge, Massachusetts: Harvard University Press.

Losos, J. B. (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* 175, 623–639. doi: 10.1086/652433

Madi, N., Vos, M., Murall, C. L., Legendre, P., and Shapiro, B. J. (2020). Does diversity beget diversity in microbiomes?. *Elife* 9:e58999. doi: 10.7554/eLife.58999

Magdalena, S. R., and Wagner, A. (2018). An enormous potential for niche construction through bacterial cross-feeding in a homogeneous environment. *PLoS Comput. Biol.* 14:e1006340. doi: 10.1371/journal.pcbi.1006340

Maynard, D. S., Bradford, M. A., Lindner, D. L., van Diepen, L. T. A., Frey, S. D., Glaeser, J. A., et al. (2017). Diversity begets diversity in competition for space. *Nat. Ecol. Evol.* 1, 1–8. doi: 10.1038/s41559-017-0156

Mori, M., Hwa, T., Martin, O. C., De Martino, A., and Marinari, E. (2016). Constrained allocation flux balance analysis. *PLoS Comput. Biol.* 12:e1004913. doi: 10.1371/journal.pcbi.1004913

Odling-Smee, F. J., Laland, K. N., and Feldman, M. W. (2013). *Niche Construction: The Neglected Process in Evolution (MPB-37).* Princeton, New Jersey: Princeton University Press.

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

Paczia, N., Nilgen, A., Lehmann, T., Gätgens, J., Wiechert, W., and Noack, S. (2012). Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms. *Microb. Cell Fact.* 11:122. doi: 10.1186/1475-2859-11-122

Pál, C., and Papp, B. (2017). Evolution of complex adaptations in molecular systems. *Nat. Ecol. Evol.* 1, 1084–1092. doi: 10.1038/s41559-017-0228-1

Plata, G., Henry, C. S., and Vitkup, D. (2015). Long-term phenotypic evolution of bacteria. *Nature* 517, 369–372. doi: 10.1038/nature13827

Quandt, E. M., Gollihar, J., Blount, Z. D., Ellington, A. D., Georgiou, G., and Barrick, J. E. (2015). Fine-tuning citrate synthase flux potentiates and refines metabolic innovation in the lenski evolution experiment. *Elife* 4:e09696. doi: 10.7554/eLife.09696

Rainey, P. B., and Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature* 394, 69–72. doi: 10.1038/27900

Sanchez, A., and Gore, J. (2013). Feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS Biol.* 11:e1001547. doi: 10.1371/journal.pbio.1001547

Sasse, J., Kant, E., Cole, B. J., Klein, A. P., Arsova, L., Schlaepfer, P., et al. (2010). Multilab EcoFAB study shows highly reproducible physiology and depletion of soil metabolites by a model grass. *New Phytol.* 222, 1149–1160. doi: 10.1111/nph.15662

Schick, A., and Kassen, R. (2018). Rapid diversification of pseudomonas aeruginosa in cystic fibrosis lung-like conditions. *Proc. Natl. Acad. Sci. U. S. A.* 115, 10714–10719. doi: 10.1073/pnas.1721270115

Schirrmeister, B. E., de Vos, J. M., Antonelli, A., and Bagheri, H. C. (2013). Evolution of multicellularity coincided with increased diversification of cyanobacteria and the great oxidation event. *Proc. Natl. Acad. Sci. U. S. A.* 110, 1791–1796. doi: 10.1073/pnas.1209927110

Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution* 50, 1766–1774. doi: 10.1111/j.1558-5646.1996.tb03563.x

Schluter, D. (2000). *The Ecology of Adaptive Radiation.* Oxford: Oxford University Press.

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science* 323, 737–741. doi: 10.1126/science.1160006

Silver, M., and Di Paolo, E. (2006). Spatial effects favour the evolution of niche construction. Theor. *Theor. Popul. Biol.* 70, 387–400. doi: 10.1016/j.tpb.2006.08.003

Silvestro, D., Zizka, G., and Schulte, K. (2014). Disentangling the effects of key innovations on the diversification of Bromelioideae (bromeliaceae). *Evolution* 68, 163–175. doi: 10.1111/evo.12236

Simpson, G. G. (1953). *The Major Features of Evolution*. New York Chichester, West Sussex: Columbia University Press.

Sneppen, K., Bak, P., Flyvbjerg, H., and Jensen, M. H. (1995). Evolution as a self-organized critical phenomenon. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5209–5213. doi: 10.1073/pnas.92.11.5209

Stadler, B. M., Stadler, P. F., Wagner, G. P., and Fontana, W. (2001). The topology of the possible: formal spaces underlying patterns of evolutionary change. *J. Theor. Biol.* 213, 241–274. doi: 10.1006/jtbi.2001.2423

Steinberg, B., and Ostermeier, M. (2016). Environmental changes bridge evolutionary valleys. *Sci. Adv.* 2:e1500921. doi: 10.1126/sciadv.1500921

Stroud, J. T., and Losos, J. B. (2016). Ecological opportunity and adaptive radiation. *Annu. Rev. Ecol. Evol. Syst.* 47, 507–532. doi: 10.1146/annurev-ecolsys-121415-032254

Szappanos, B., Fritzemeier, J., Csörgő, B., Lázár, V., Lu, X., Fekete, G., et al. (2016). Adaptive evolution of complex innovations through stepwise metabolic niche expansion. *Nat. Commun.* 7:11607. doi: 10.1038/ncomms11607

van Tatenhove-Pel, R. J., Rijavec, T., Lapanje, A., Swam, I. V., Zwering, E., Hernandez-Valdes, J. A., et al. (2021). Microbial competition reduces metabolic interaction distances to the low µm-range. *ISME J.* 15, 688–701. doi: 10.1038/s41396-020-00806-9

Weinreich, D. M., Delaney, N. F., Depristo, M. A., and Hartl, D. L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312, 111–114. doi: 10.1126/science.1123539

Weinreich, D. M., Watson, R. A., and Chao, L. (2005). Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59, 1165–1174. doi: 10.1111/j.1420-9101.2010.02029.x

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon* 21, 213–251.

Yoder, J. B., Clancey, E., Des Roches, S., Eastman, J. M., Gentry, L., Godsoe, W., et al. (2010). Ecological opportunity and the origin of adaptive radiations. *J. Evol. Biol.* 23, 1581–1596. doi: 10.1111/j.1420-9101.2010.02029.x

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 01 November 2021
doi: 10.3389/fmicb.2021.719725

Salinity Drives Functional and Taxonomic Diversities in Global Water Metagenomes

Huaihai Chen[1], Kayan Ma[1], Yu Huang[1], Yuchun Yang[1]*, Zilong Ma[1] and Chengjin Chu[2]*

[1] State Key Laboratory of Biocontrol, School of Ecology, Sun Yat-sen University, Shenzhen, China, [2] State Key Laboratory of Biocontrol, School of Life Sciences, School of Ecology, Sun Yat-sen University, Guangzhou, China

A tight association between microbial function and taxonomy is the basis of functional prediction based on taxonomy, but such associations have been controversial in water biomes largely due to the probable prevalence of functional redundancy. However, previous studies on this topic used a relatively coarse resolution of ecosystem functioning, potentially inflating the estimated functional redundancy. Thus, a comprehensive evaluation of the association between high-resolution functional traits and taxonomic diversity obtained from fresh and saline water metagenomic data is urgently needed. Here, we examined 938 functionally and taxonomically annotated water metagenomes obtained worldwide to scrutinize the connection between function and taxonomy, and to identify the key driver of water metagenomes function or taxonomic composition at a global scale. We found that pairwise similarity of function was significantly associated with taxonomy, though taxonomy had higher global dissimilarity than function. Classification into six water biomes resulted in greater variation in taxonomic compositions than functional profiles, as the key regulating factor was salinity. Fresh water microbes harbored distinct functional and taxonomic structures from microbes in saline water biomes, despite that taxonomy was more susceptible to gradient of geography and climate than function. In summary, our results find a significant relationship between taxonomic diversity and microbial functioning in global water metagenomes, although microbial taxonomic compositions vary to a larger extent than functional profiles in aquatic ecosystems, suggesting the possibility and necessity for functional prediction of microorganisms based on taxonomy in global aquatic ecosystems.

Keywords: metagenomics, microbes, taxonomy, function, salinity, freshwater

OPEN ACCESS

Edited by:
Maria Elena Martino, University of Padua, Italy

Reviewed by:
Valeria Souza, National Autonomous University of Mexico, Mexico
Alexander Eiler, Uppsala University, Sweden

****Correspondence:***
Huaihai Chen
chenhh68@mail.sysu.edu.cn
Yuchun Yang
yangych55@mail.sysu.edu.cn

Specialty section:
This article was submitted to Aquatic Microbiology, a section of the journal Frontiers in Microbiology

Received: *03 June 2021*
Accepted: *05 October 2021*
Published: *01 November 2021*

Citation:
Chen H, Ma K, Huang Y, Yang Y, Ma Z and Chu C (2021) Salinity Drives Functional and Taxonomic Diversities in Global Water Metagenomes. Front. Microbiol. 12:719725. doi: 10.3389/fmicb.2021.719725

INTRODUCTION

Microbial communities are major regulators of biogeochemical processes and ecosystem functions (Hall et al., 2018). While global decline in biodiversity will negatively impact ecosystem functions and services in both aquatic and terrestrial ecosystems (Jousset et al., 2016; Schmidt et al., 2017), understanding the relationship between microbial functional profile and taxonomic composition is essential for predicting ecosystem functioning based on microbial diversity under various environmental disturbances (Torsvik and Øvreås, 2002; Wellington et al., 2003; McGill et al., 2006). It is often presumed that though microbial communities are sensitive to disturbance, their overall ecosystem functioning remains relatively stable, as many microbes are probably functionally

redundant (Allison and Martiny, 2008). The association between microbial diversity and ecosystem functioning may be obscure due to functional redundancy. Yet, the extent to which such functional redundancy could affect our potential to evaluate the global consequences of shifting microbial diversity on ecosystem functioning remains largely unknown.

In contrast to the theory of functional redundancy, increasing evidence has shown that species richness can be positively related to multiple aspects of ecosystem functioning, including resource usage, nutrient cycling, and biomass accumulation (Covich et al., 2004; Downing, 2005; Balvanera et al., 2006). Compared to richness, species identity in community composition may be more suitable for predicting the ecosystem functioning. By analyzing community composition, it has been found that a decreasing gradient of microbial diversity in freshwater metagenomes could impact both broad and specialized functions (Peter et al., 2011; Delgado-Baquerizo et al., 2016a). Thus, any shift in bacterial composition could cause at least a proportional depletion of microbial capability to support ecosystem functions, implying minor functional redundancy in freshwater metagenomes. However, the characterization of ecosystem functioning of these studies was generally of low resolution, mostly in terms of certain biogeochemical processes only. A comprehensive understanding of the correlation between microbial functional and taxonomic diversities will require a thorough evaluation of microbial functional composition, which has not been analyzed and thus leaving a wide gap of knowledge that needs to be addressed.

In an analysis of global marine metagenomes, decoupling of function and taxonomy has been suggested, as the taxonomy is highly variable within specific functional groups, and environmental conditions strongly influence the distribution of functional groups while only weakly affect taxonomic composition within individual functional groups (Louca et al., 2016b). However, a lower deviation in functional profiles than taxonomic composition alone, which suggests a certain extent of functional redundancy, cannot refute the dependency of microbial functional profiles on taxonomic compositions if there exists a significant association between the two. More importantly, it has been found that salinity is the major factor regulating global aquatic microbial community, even more influential than sampling types, extreme temperature, and pH (Lozupone and Knight, 2007). Since aquatic microbes are strongly affected by physical and chemical properties of aquatic ecosystems, solely focusing on marine microbiomes may underestimate the significant effects of salinity on microbial distribution and relevant function in water biomes. Thus, it is of great importance to combine both fresh and saline water biomes in quantitative analysis to determine the relationship between functional and taxonomic diversities so as to understand the impact of future global change on diversity loss and ecosystem functioning in water biomes.

The importance of multifunctionality (Hector and Bagchi, 2007), which is the assessment of multiple functions performed by different species at the same time, has been highlighted to avoid overestimating functional redundancy (Gamfeldt et al., 2008). The microbial multifunctionality index composed by multiple assays needs to be quantified to better represent functional traits corresponding to taxonomic diversity (Bastida et al., 2016; Delgado-Baquerizo et al., 2016b). In recent decades, metagenomics has been increasingly used as a promising comparative tool (Tringe et al., 2005) to assess microbial functional diversities (Fierer et al., 2012a,b, 2013), since the abundance of each gene can be specific to a particular environmental process and numerous ecosystem functions can be examined all together in one environmental sample (Allison and Martiny, 2008). To date, open-source web servers, such as Metagenomic Rapid Annotations using Subsystems Technology (MG-RAST) (Meyer et al., 2008), are publicly available for meta-analyses (Nelson et al., 2016; Ramírez-Flandes et al., 2019) based on simultaneous metagenomic annotation against functional and taxonomic databases, allowing direct comparisons between functional profiles and taxonomic compositions. Yet, a comprehensive metagenomic analysis to elucidate the extent to which microbial functional profiles respond to different scales of variations in taxonomic compositions is still lacking for global water metagenomes.

Here, we constructed a dataset of 938 water metagenomes, functionally and taxonomically annotated, acquired from peer-reviewed publications and MG-RAST database. Based on this dataset of global water metagenomes, we specifically tested the hypotheses that (1) across the globe, microbial functional profiles are associated with taxonomic compositions, though taxonomy may harbor less similarity and larger variation than function; (2) salinity is the potential key driver for significant shifts in both functional and taxonomic diversities across the globe.

MATERIALS AND METHODS

Data Collection

Instead of directly obtaining available shotgun metagenomic data from a public server as in previous global metagenomic studies (Nelson et al., 2016; Ramírez-Flandes et al., 2019), we strategically selected water metagenomes published in peer-reviewed journals to ensure the quality and completeness of the data. Specifically, we searched peer-reviewed publications from 2006 to 2019 from the Web of Science database using the terms such as "water metagenome," "shotgun sequencing," and "MG-RAST." We included studies directly deposited or analyzed water metagenomes using shotgun sequencing without amplification in the MG-RAST database, and water metagenomes publicly available in the MG-RAST database. Based on the Study ID and/or MG-RAST ID reported in the publications, we extracted a data matrix of water metagenomes from the MG-RAST database. Specifically, in the "Analysis" function of the MG-RAST server, we typically loaded both SEED Subsystems (Overbeek et al., 2013) (Function, level 3, 2, and 1) as functional profiles and RefSeq (Tatusova et al., 2013) databases (genus, family, order, class, and phylum levels) as taxonomic compositions.

The analyses were performed using default settings (maximum *e*-value cutoff was $1e^{-5}$, minimum identity cutoff was 60%, and minimum alignment length was 50) (Meyer et al., 2008). We further merged the data matrix of each function

extracted from different studies together to build up new datasets of microbial functional profiles annotated in the Subsystems database and taxonomic compositions annotated in RefSeq database. We chose Subsystems database for functional grouping rather than KEGG Orthology (KO; Kanehisa et al., 2015), Clusters of Orthologous Groups (COG; Galperin et al., 2014), and Non-supervised Orthologous Groups (NOG; Huerta-Cepas et al., 2015) databases because Subsystems showed diverse classification at level 1, allowing us to conduct more detailed comparison of functions among water biomes. We chose RefSeq database rather than traditional ribosomal RNA databases, such as Ribosomal Database Project (RDP; Cole et al., 2008), Greengenes (DeSantis et al., 2006), or Silva LSU/SSU (Pruesse et al., 2007) databases because taxonomic hits in RefSeq database were over 1000-fold higher than rRNA databases, which was comparable to functional hits for comparison. To expand our datasets, water metagenomes with/without assembly were both included.

In total, this study included 938 water metagenomes around the world extracted from 55 MG-RAST studies published in 55 peer-reviewed papers (**Figure 1**). Detailed information of each water metagenome extracted from publications and the MG-RAST database was given (**Supplementary Table 1**). Study ID, MG-RAST ID, sample name, base pair (bp), and reads were obtained from the metadata of each water metagenome in the MG-RAST database. The geographic coordinates of latitudes (LAT) and longitudes (LONG) of each water metagenome were directly obtained from publications or MG-RAST studies. All water metagenomes were grouped into six categories, including three freshwater biomes (groundwater, surface freshwater, and urban wastewater), and three saline water biomes (coastal seawater, marine seawater, and surface saline water).

Statistical Analyses

Hits of each functional or taxonomic category in the data matrix were standardized to relative abundance through dividing by total hits to remove bias in difference in sequencing depths and read lengths among different studies. To calculate the pairwise similarity of taxonomy, based on the relative taxonomic abundance at the genus level, we calculated the Bray-Curtis similarity following log transformation of the compositional taxonomic data by constructing a pairwise Bray-Curtis similarity matrix between each pair of two samples, which were further transformed to lists of pairwise Bray-Curtis similarities ordered by sample pair names in PRIMER 7 (Plymouth Routines in Multivariate Ecological Research Statistical Software, v7.0.13, PRIMER-E Ltd., United Kingdom). To calculate the pairwise similarity of function, based on the functional abundance at the function gene level, we calculated the Bray-Curtis similarity following log transformation of the compositional functional data by constructing pairwise Bray-Curtis similarity matrix between each pair of two samples, which were further transformed to lists of pairwise Bray-Curtis similarities ordered by sample pair names in PRIMER 7. To examine the relationship between functional and taxonomic diversities, Pearson's correlations were constructed between the transformed lists of pairwise Bray-Curtis similarity of metagenomes annotated using the Subsystems database at Level 3 (Function) and the RefSeq database at the genus level (Taxonomy). These data processing approaches for the analyses followed the requirement of processing relative abundance of compositional data (Gloor et al., 2017). The box plots were constructed based on the pairwise similarity of function and taxonomy to compare similarity ranges of functional and taxonomic compositions related to the aquatic metagenomes. The triangular pairwise Bray-Curtis similarity matrix was used to analyze functional and taxonomic composition structures of water metagenomes by principal coordinates analysis (PCoA) and permutational multivariate analysis of variance (PERMANOVA) among biomes in PRIMER 7. Linear and non-linear regression models were constructed to examine whether functional diversity depends on taxonomic diversity. In addition, the RELATE analysis was performed to evaluate the relatedness between functional and taxonomic diversities by calculating a Spearman's Rho correlation coefficient in PRIMER 7. Difference in the relative abundance of dominant microbial functional and taxonomic compositions for fresh vs. saline and among six water biomes was analyzed by the linear discriminant analysis effect size (LEfSe) method[1] (Segata et al., 2011). Heatmaps were constructed using HeatMapper (Babicki et al., 2016). The significance level was set at $\alpha = 0.05$ unless otherwise stated.

Co-occurrence network analysis was performed using the Molecular Ecological Network Analyses Pipeline[2] (Zhou et al., 2011; Deng et al., 2012). The data matrix of standardized relative abundance multiplied by 10^6 that meets the requirements of the pipeline was uploaded to construct the network with default settings, including (1) only keeping the species present in more than a half of all samples; (2) only filling with 0.01 in blanks with paired valid values; (3) taking logarithm with recommended similarity matrix of Pearson's correlation coefficient; (4) calculation order to decrease the cutoff from the top using regress Poisson distribution only. A default cutoff value (similarity threshold, S_t) for the similarity matrix was generated to assign a link between the pair of species. Then, the global network properties, the individual nodes' centrality, and the module separation and modularity calculations were run based on default settings using greedy modularity optimization. Network files were exported and visualized using Cytoscape software (Shannon et al., 2003).

RESULTS AND DISCUSSION

Correlation Between Function and Taxonomy

To examine the connection between function and taxonomy, a total of 938 water metagenomes was used to create 439,453 pairwise comparisons of Bray-Curtis similarity between functional and taxonomic diversities to find out whether the correlation of function and taxonomy was significant in the global aquatic metagenomes (**Figure 2**). The correlations

[1] http://huttenhower.sph.harvard.edu/lefse/

[2] http://ieg4.rccc.ou.edu/MENA/

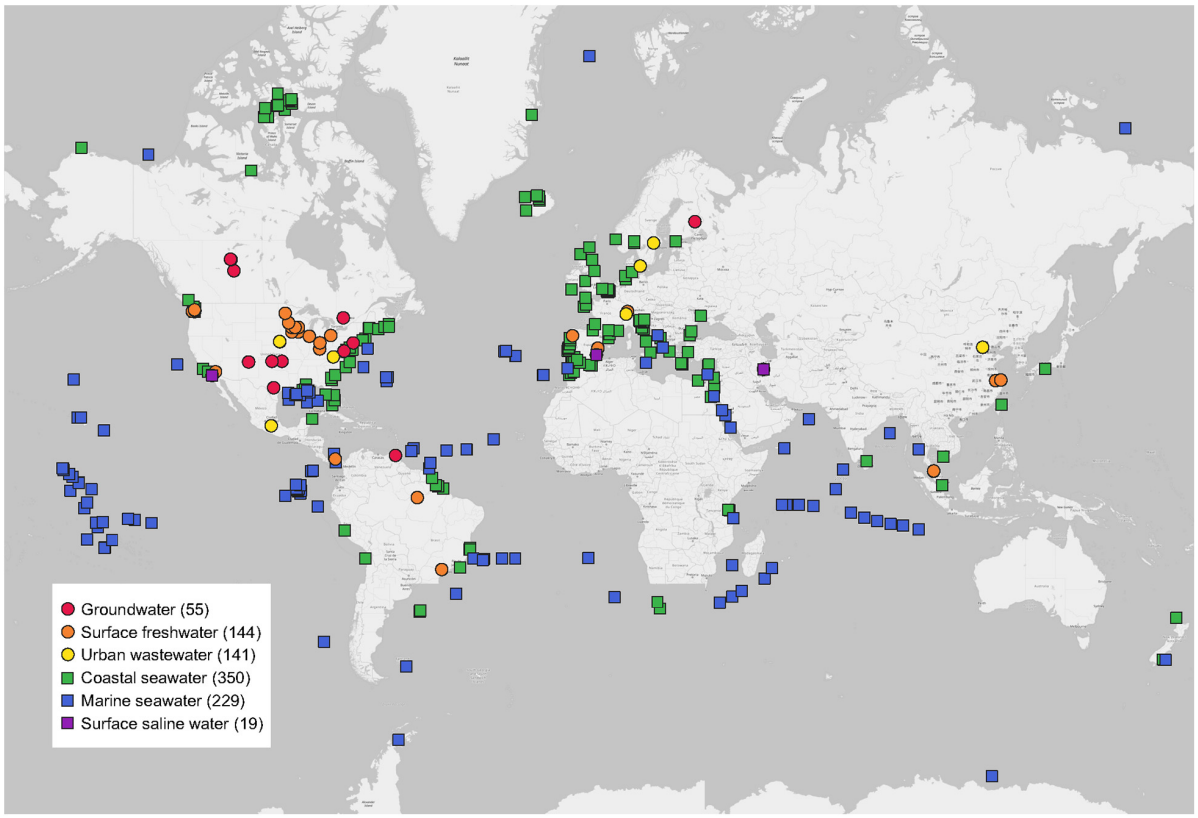


FIGURE 1 | Global distribution of water metagenomes. Locations are grouped into six water biomes from 55 publications used in this study. Sample sizes of each are given in parentheses.

showed that pairwise similarity of function was positively correlated to that of taxonomy as indicated by the logarithmic regression (r^2 = 0.3344, P < 0.0001) (**Figure 2A**), which had higher coefficient than the linear regression (r^2 = 0.3209, P < 0.0001) (**Supplementary Figure 1**). This suggested that in water microbiomes, the variation of functional traits responds positively to taxonomic shifts, so the dependency of microbial function on taxonomy could be expected across fresh and saline water biomes. The box plots were constructed based on the pairwise similarity of function and taxonomy to compare similarity ranges of functional and taxonomic compositions related to the aquatic metagenomes. For the functional compositions at specific function gene levels, the average pairwise similarity of function at three levels of SEED Subsystems was greater than taxonomy at five phylogenetic levels (**Figure 2B**), suggesting that the variation of function and taxonomy had different sensitivity to environmental selection.

Since functional composition of higher resolution can be yielded from metagenomic analyses, we were able to better elucidate its relationship with taxonomic diversity in microbiomes than previous studies. Cross-biome soil metagenomics have shown that alpha diversity levels of microbial community were significantly associated with functional attributes (Fierer et al., 2012b). Thus, community richness can be potentially useful for inferring functional diversity. Based on regional metagenomics of soil microbial communities, a strong correlation has been observed between functional and taxonomic community composition (Fierer et al., 2013), suggesting a potential connection between beta diversities of taxonomy and function across global microbiomes. Our results further strengthened this argument by demonstrating significant associations between shifts in microbial taxonomy and functional variation in the water microbial community across the globe. This is also strong evidence for an insignificant, if any, functional redundancy in microbes at this resolution.

Functional and Taxonomic Diversities Across Water Biomes

Based on pairwise Bray-Curtis similarity, Principal Coordinates Analysis (PCoA) was performed to examine how microbial functional profiles and taxonomic compositions differ globally among water biomes (**Figure 3**). The PERMANOVA showed that both function and taxonomy of water microbiomes among the six water biomes were significantly distinct (P < 0.0001). Compared to taxonomy (Pseudo-F = 70.41, Sq. root = 25.96), function had lower variation (Pseudo-F = 48.16, Sq. root = 9.29) across global water biomes. Greater between-sample dissimilarity in taxonomy than function suggests that the taxonomic structure of microbial communities was more sensitive to environmental

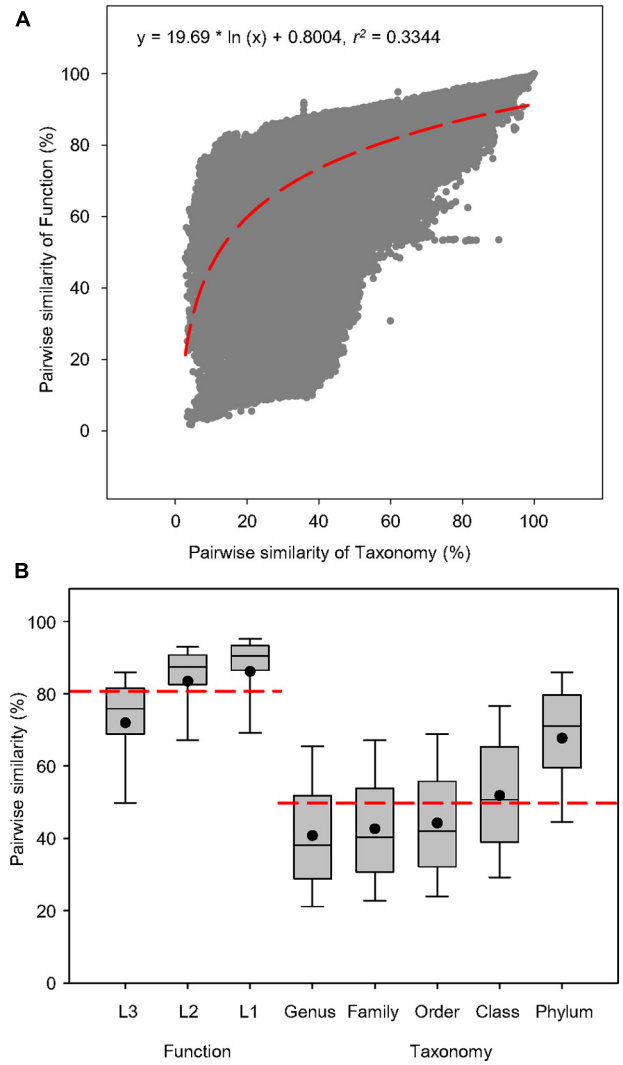


FIGURE 2 | Pairwise similarity of functional and taxonomic diversities. **(A)** Logarithmic regression between pairwise Bray-Curtis similarity of water metagenomes annotated in the Subsystems database at level 3 (Function) and RefSeq database at the genus level (Taxonomy). Regression equations and coefficients are given. The logarithmic regression line was shown as the red dashed line. **(B)** Box plots and mean values of pairwise Bray-Curtis similarity of water metagenomes annotated in the Subsystems database at levels 3, 2, and 1 (Function) and RefSeq database at genus, family, order, class, and phylum levels (Taxonomy). The mean values of pairwise Bray-Curtis similarity was shown as the red dashed line.

selection, though environmental conditions also influence the functional traits in microbial communities. Specifically, for functional distribution, the three saline water biomes (coastal seawater, marine seawater, and surface saline water) were grouped together, whereas among the three freshwater biomes groundwater and urban wastewater were distributed near each other with surface freshwater in the middle (**Figure 3A**). The taxonomic compositions showed that the three freshwater biomes clustered together and were distinct from the three saline water biomes (**Figure 3A**), suggesting a key role of salinity in driving taxonomy and function of water microbiomes.

The relative abundances of functions were similar across global water microbiomes, while the taxonomy of microbial community was significantly more variable across the six water biomes (**Figure 3B**). Similar functional structures but high taxonomic variability have also been found in microbial communities from replicates of "miniature" aquatic ecosystems (Louca et al., 2016a). It is often assumed that genome streamlining (Morris et al., 2012) and horizontal gene transfer (David and Alm, 2011), common in prokaryotic populations, have contributed to similar functions performed by distinct taxa. In addition, it has been suggested that the taxonomic composition may be influenced by environmental factors distinct from those affecting the functional structure of microbial communities (Louca et al., 2018).

In functional profiles, the most significant difference was observed in genes involved in virulence, disease, and defense (VDD) as indicated by the LEfSe analysis (**Figure 3B**). These genes were markedly more abundant in groundwater and urban wastewater than the others, possibly because these waters were closer to human activity and hence more susceptible to pathogen and disease, leading to the proliferation of genes associated with VDD functions, such as antibiotic resistance genes (Czekalski et al., 2014). Coastal and marine seawater had greater proportion of genes associated with functions related to phages, prophages, transposable elements, and plasmids. Because the oceans are the largest reservoir of genetic diversity of virus (Suttle, 2007), the importance and focus of aquatic virology are perhaps most evident and mainly on the viromes of marine environments (Jover et al., 2014).

In terms of taxonomic compositions, surface freshwater and urban wastewater harbored the highest proportion of Actinobacteria among the biomes, while Euryarchaeota was the most prevelent in groundwater and surface saline water (**Figure 3B**). In addition, the two ocean biomes of coastal seawater and marine seawater were dominated by α-proteobacteria and γ-proteobacteria. The percentage of β-proteobacteria was greater in the three freshwater microbiomes than the others. Groundwater had the highest proportion of δ-proteobacteria. In agreement with our results, it has been found that α-proteobacteria and γ-proteobacteria were significantly enriched in marine ecosystems, while β-proteobacteria favored fresh environment (Hu et al., 2014; Morrissey and Franklin, 2015; Pavloudi et al., 2017). Many halotolerant bacteria with the capability to grow in a wide range of salinities are found to belong to the class of α-proteobacteria and γ-proteobacteria (Etesami and Beattie, 2018). Our extensive lines of evidence suggest that the global response of Proteobacteria was class-specific, mainly due to salinity effects. In conclusion, these striking differences in the relative abundance of dominant groups confirmed that the taxonomy of water microbiomes was more variable than function across global water biomes.

Salinity Is the Key Regulator

To test if salinity is a key driver for microbial function and taxonomy, microbiomes from a freshwater environment were compared with those from saline water (**Figure 4**). The PERMANOVA revealed that fresh and saline water microbiomes had distinct functional and taxonomic structures ($P < 0.0001$)

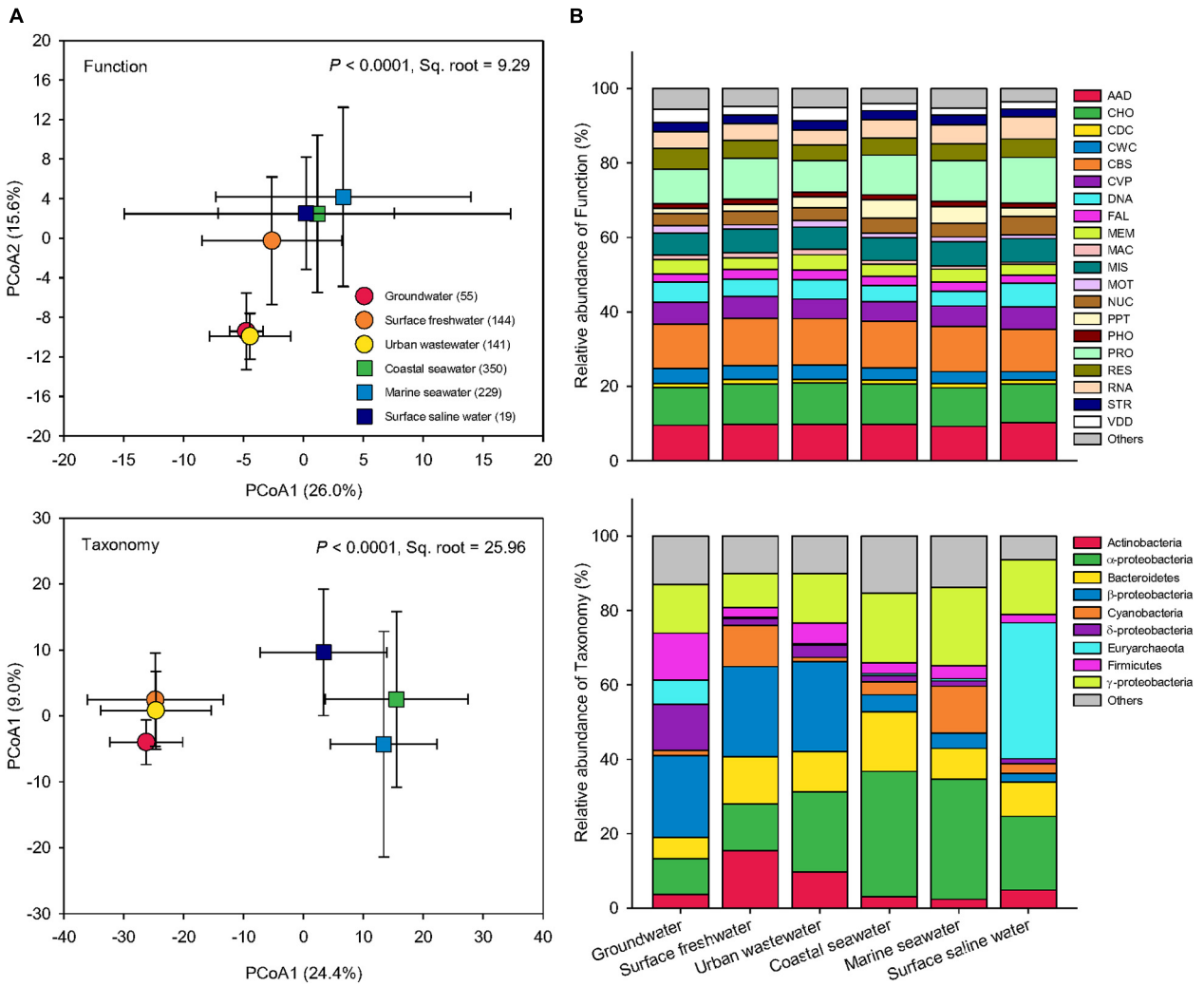


FIGURE 3 | Functional and taxonomic diversities across six water biomes. **(A)** Principal coordinates analysis (PCoA) showing beta-diversity of water metagenomes annotated in the Subsystems database at level 3 (Function) and RefSeq database at the genus level (Taxonomy) among six biomes. The error bars represent the standard deviation of data ranges. Sample sizes of each biome are given in parentheses. Variations explained by two principal coordinate dimensions are given in parentheses by percentage. *P* values and Sq. root of PERMANOVA are also given. **(B)** Relative abundance of dominant water metagenomes (mean >1%) annotated in the Subsystems database at level 1 (Function), and the RefSeq database at the phylum/class levels (Taxonomy) among six water biomes. AAD, amino acids and derivatives; CHO, carbohydrates; CDC, cell division and cell cycle; CWC, cell wall and capsule; CBS, clustering-based subsystems; CVP, cofactors, vitamins, prosthetic groups, pigments; DNA, DNA metabolism; FAL, fatty acids, lipids, and isoprenoids; MEM, membrane transport; MAC, metabolism of aromatic compounds; MIS, miscellaneous; MOT, motility and chemotaxis; NUC, nucleosides and nucleotides; PPT, phages, prophages, transposable elements, plasmids; PHO, phosphorus metabolism; PRO, protein metabolism; RES, respiration; RNA, RNA metabolism; STR, stress response; VDD, virulence, disease and defense.

(**Figure 4A**). The differences in function (Pseudo-F = 112.37, Sq. root = 8.52) and taxonomy (Pseudo-F = 221.30, Sq. root = 27.92) between fresh vs. saline water biomes were comparable to the variation caused by grouping the microbiomes according to the six water biomes, showing that salinity can be the key factor regulating the variation of function and taxonomy in water microbiomes. However, it should be noted that determinations of salinity in this study are qualitatively based on the descriptions in the publications or MG-RAST metadata files rather than on direct measurements of salinity levels. The community variation in both function and taxonomy in saline water biomes was larger than fresh water microbiomes, probably due to much wider distribution of sampling locations of coastal and marine seawater.

To find out whether latitude gradient also drives taxonomic and functional diversities besides salinity, Pearson's correlations were performed between the absolute latitude of sampling locations and PCoA1 scores of functional and taxonomic beta-diversity structures in fresh and saline water microbiomes, respectively (**Figure 4B**). Functional profiles showed no significant latitudinal gradient in either fresh or saline water microbiomes ($P > 0.05$). On the contrary, taxonomic compositions in both fresh and saline environments had significant associations with the absolute latitude ($P < 0.001$).

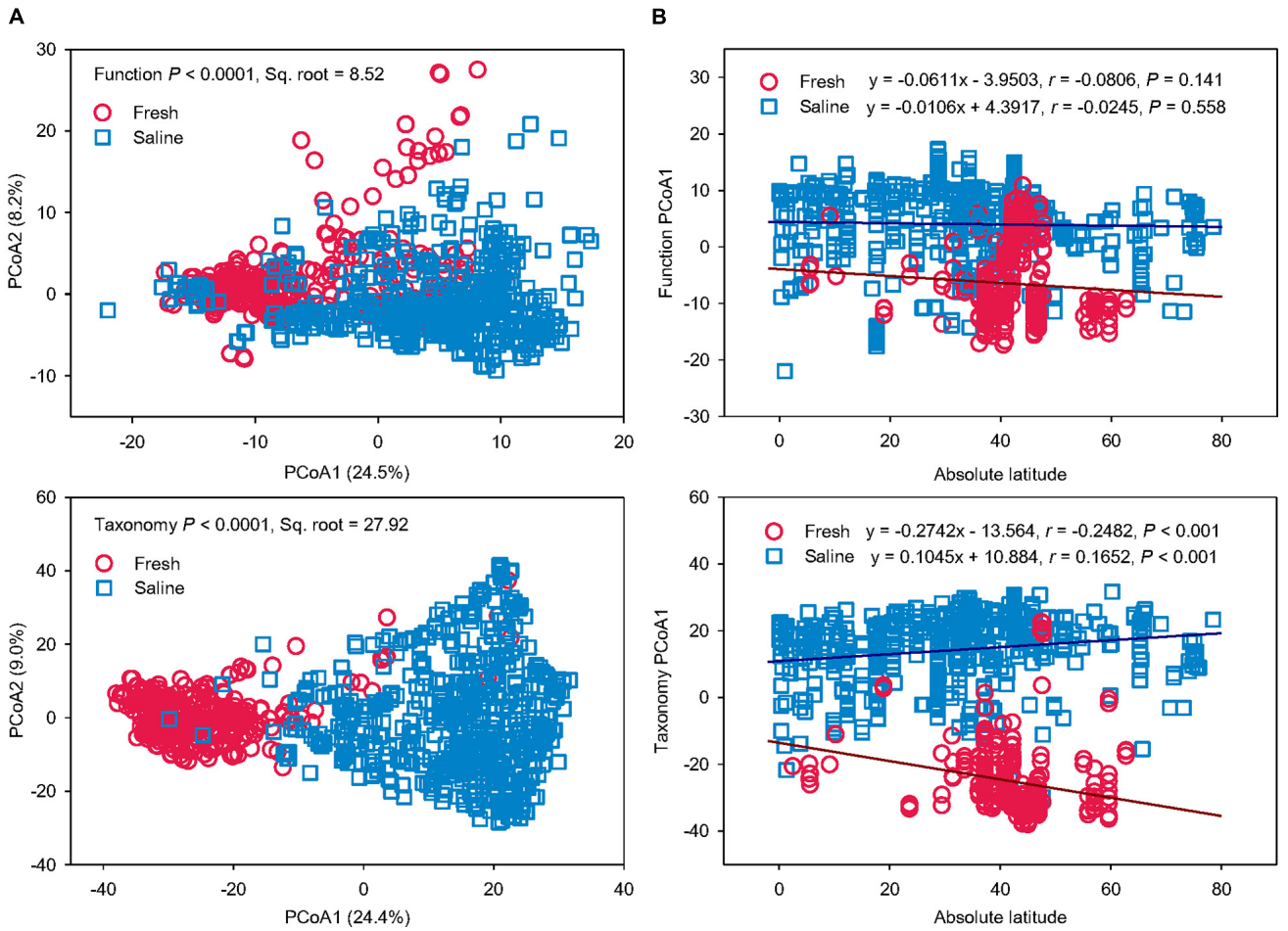


FIGURE 4 | Salinity drives microbial functional and taxonomic diversities. **(A)** Principal coordinates analysis (PCoA) showing beta-diversity of microbial functional and taxonomic diversities between fresh vs. saline water metagenomes annotated in Subsystems L3 (Function) and RefSeq at the genus level (Taxonomy). The error bars represent the standard deviation of data ranges. Variations explained by two principal coordinate dimensions are given in parentheses by percentage. *P* values and Sq. root of PERMANOVA are also given. **(B)** Pearson's correlations of the absolute latitude of water metagenome locations with principal coordinate (PCoA) 1 scores of Bray-Curtis similarity of fresh vs. saline water metagenomes annotated in Subsystems L3 (Function) and RefSeq at the genus level (Taxonomy). Regression equations, coefficients (*r*), and *P* values are given. The linear regression lines of fresh and saline water metagenomes were shown as the brown and blue solid lines, respectively.

With increasing latitude, microbial taxonomy in freshwater biomes became more different from that in the saline water biomes. Latitudinal diversity gradient, a decline of biodiversity with latitude (Hillebrand, 2004), has been found in alpha-diversity of microbes in the terrestrial (Zhou et al., 2016) and aquatic (Fuhrman et al., 2008; Huerta-Cepas et al., 2015) environments, probably attributable to temperature gradient. For the first time, we further showed this latitudinal correlation of beta-diversity in microbial taxonomic compositions rather than functional profiles in the marine ecosystems. These results suggest that global environmental drivers, such as geography and climate factors, are impacting microbial diversity in aquatic ecosystems, but the key regulator for functional traits was solely salinity in water biomes.

Significant differences in function and taxonomy between fresh and saline water microbiomes were revealed by LEfSe analysis (**Figure 5**). Generally, functional profiles of freshwater microbiomes were dominated by metabolisms of aromatic compounds, iron, and nitrogen, membrane transport and cell wall and capsule, and virulence, disease, and defense, focusing more on functions associated with nutrient cycling and defense mechanisms (**Figure 5**). By comparison, in the saline microbiomes' functional profiles, functions related to primary metabolisms of nucleotides, amino acids, proteins, and carbohydrates were more abundant (**Figure 5**). In taxonomy, freshwater microbiomes were dominated by Actinobacteria, Clostridia, and Bacilli of Firmicutes, Burkholderiales, and Rhodocyclales belonging to β-proteobacteria and δ-proteobacteria (**Figure 5**). On the contrary, saline water biomes had higher abundance of Archaea, Eukaryota, and viruses. Among bacteria, only γ-proteobacteria had consistently higher abundance at class or lower levels in the saline environment, though Bacteroidetes, Cyanobacteria, and α-proteobacteria were also dominant at the phylum levels in saline water microbiomes

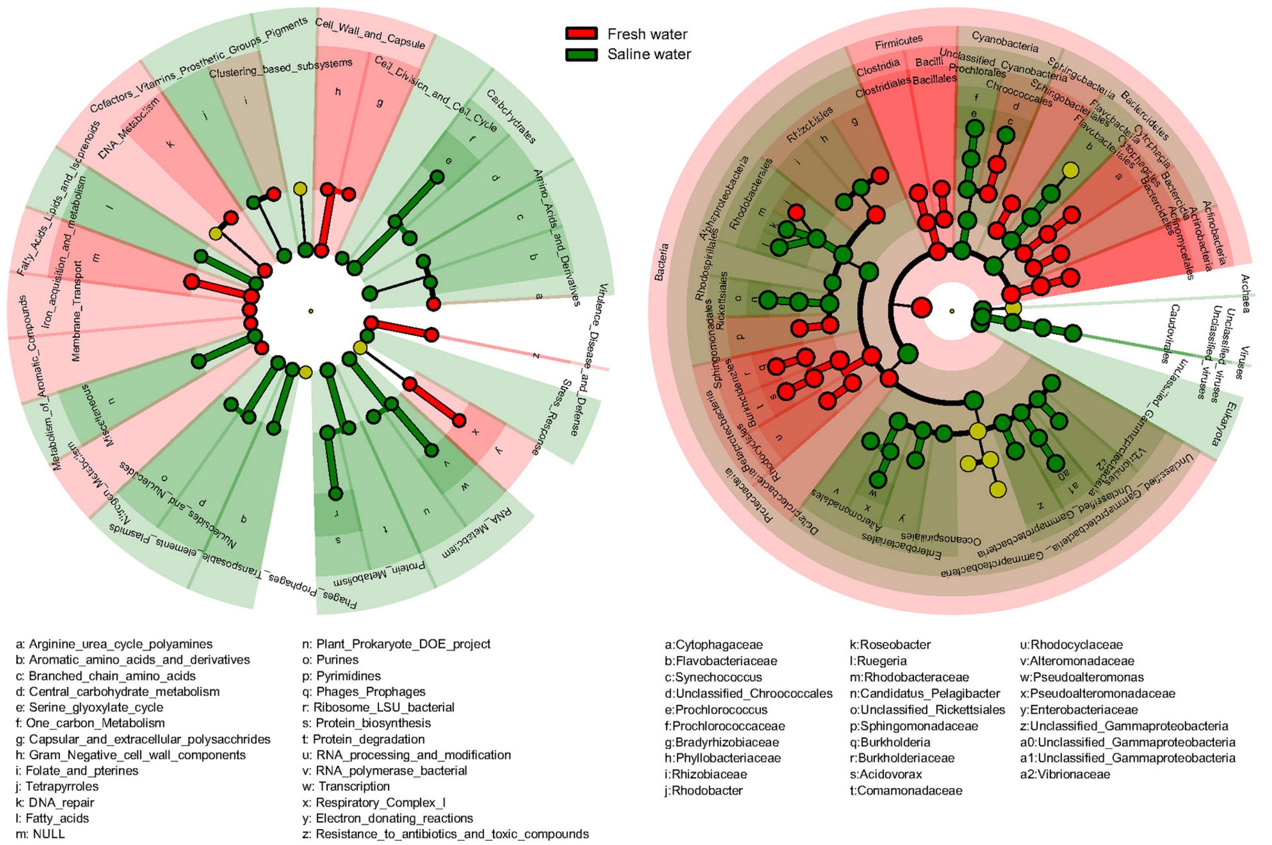
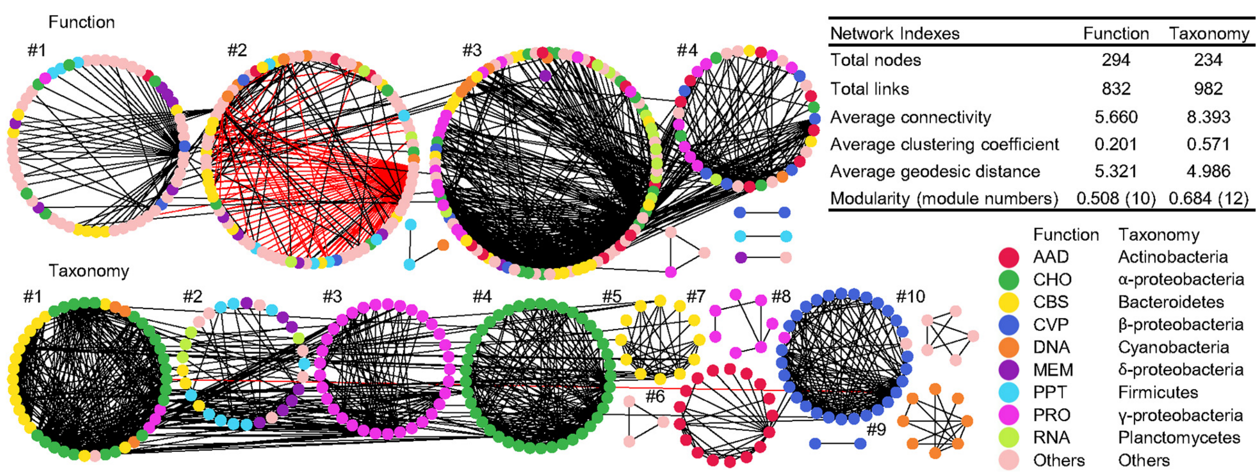


FIGURE 5 | Difference of functional and taxonomic compositions between fresh and saline water. Linear discriminant analysis effect size (LEfSe) method showing the significant differences in the relative abundance of dominant microbial functional and taxonomic compositions (mean >1%) between fresh vs. saline water metagenomes annotated in Subsystems (Function) and RefSeq (Taxonomy). From the center outward, in the left panel, each circle represents level 1, 2, and 3 of functional annotation, and in the right panel, each circle represents the level of domain, phylum, class, order, family, and genus from taxonomic annotation, respectively. The functional and taxonomic compositions with significant differences are labeled by colors.



Network Indexes	Function	Taxonomy
Total nodes	294	234
Total links	832	982
Average connectivity	5.660	8.393
Average clustering coefficient	0.201	0.571
Average geodesic distance	5.321	4.986
Modularity (module numbers)	0.508 (10)	0.684 (12)

FIGURE 6 | Co-occurrence networks of functional and taxonomic compositions. Network graphs with submodule structures of water metagenomes annotated in Subsystems L3 (Function) and RefSeq at the genus level (Taxonomy). Node color represents classification of Subsystems L1 (Function) or RefSeq at the phylum/class levels (Taxonomy). A black edge indicates a positive interaction between two nodes, while a red edge indicates a negative interaction. Summary of key network indexes is given in the table.

(**Figure 5**). Our results suggested that microorganisms in water environments are apparently under strong selective pressures exerted by salinity, so the general biogeochemical properties, such as salinity, still primarily determine the linkage between function and taxonomy. It should be noted that the salinity is just one factor that could be considered in this study, as not all the metagenomes included had their metadata of all/certain environmental variables reported. Thus, it became really difficult to summarize and analyze a full list of environmental variables on the diversity of taxonomy and function of metagenomes, which may not rule out the effects of co-varying factors. Future work should target a complete evaluation of environmental variables to pinpoint specific environmental factors to drive both taxonomic and functional diversities in aquatic microbes.

Functional and Taxonomic Co-occurrence Networks

Co-occurrence networks of function and taxonomy were generated to examine the potential interaction of functional and taxonomic compositions in global water biomes (**Figure 6**). Network graphs with submodule structures and key network indexes indicated similar network topology and complexity between functional and taxonomic interaction patterns. The functional network had 738 positive interactions and 94 negative links while taxonomy only had positive interactions (**Figure 6**). Within each module, functional nodes were classified as diverse functional categories. However, submodule structures of taxonomy showed that each module was comprised of genera that were mainly from the same phylum or class, such as γ-proteobacteria in module #3 and 7, α-proteobacteria in module #4, Bacteroidetes in module #5, γ-actinobacteria in module #6, β-proteobacteria in module #8, and Cyanobacteria in module #9 (**Figure 6**).

CONCLUSION

Our findings provide what is, to our knowledge, the first direct global metagenomic evidence to support that microbial functions are significantly correlated to taxonomy in aquatic ecosystems, mainly because salinity drives microbial functional and taxonomic diversities at the global level. However, the taxonomic diversity was more significantly associated with latitudinal gradient than the functional diversity. Our study has also brought some new insights regarding the functional redundancy of water microbiomes: while functional redundancy is probably inevitable, zooming into tens of thousands of specific functional genes like this study could significantly increase the dependency of function on taxonomy. Thus, these results have significant implications for developing microbial-ecosystem models for functional prediction based on microbial taxonomy.

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**.

AUTHOR CONTRIBUTIONS

HC conceived the study, performed the data analysis, interpreted the results, and drafted the manuscript. YY, ZM, and CC secured the research funding. KM, YH, YY, and ZM critically assessed and interpreted the findings. All authors discussed results, commented on, edited, revised, and approved the manuscript.

FUNDING

This study was funded by the National Natural Science Foundation of China (Nos. 3210010286 and 3210130173) and Basic and Applied Basic Research Foundation of Guangdong Province (Nos. 2020A1515111033 and 2021A1515011195).

ACKNOWLEDGMENTS

We thank the authors of the publications included in our study, without which this global metagenomics would not be possible. We also thank two reviewers for their insightful and constructive comments for manuscript improvement.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.719725/full#supplementary-material

REFERENCES

Allison, S. D., and Martiny, J. B. (2008). Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11512–11519. doi: 10.1073/pnas.0801925105

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., et al. (2016). Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* 44, W147–W153. doi: 10.1093/nar/gkw419

Balvanera, P., Pfisterer, A. B., Buchmann, N., He, J.-S., Nakashizuka, T., Raffaelli, D., et al. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* 9, 1146–1156. doi: 10.1111/j.1461-0248.2006.00963.x

Bastida, F., Torres, I. F., Moreno, J. L., Baldrian, P., Ondoño, S., Ruiz-Navarro, A., et al. (2016). The active microbial diversity drives ecosystem multifunctionality and is physiologically related to carbon availability in Mediterranean semi-arid soils. *Mol. Ecol.* 25, 4660–4673. doi: 10.1111/mec.13783

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., et al. (2008). The ribosomal database project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37, D141–D145. doi: 10.1093/nar/gkn879

Covich, A. P., Austen, M. C., Bärlocher, F., Chauvet, E., Cardinale, B. J., Biles, C. L., et al. (2004). The role of biodiversity in the functioning of freshwater and marine benthic ecosystems. *BioScience* 54, 767–775. doi: 10.1641/0006-3568(2004)054[0767:TROBIT]2.0.CO;2

Czekalski, N., Gascón Díez, E., and Bürgmann, H. (2014). Wastewater as a point source of antibiotic-resistance genes in the sediment of a freshwater lake. *ISME J.* 8, 1381–1390. doi: 10.1038/ismej.2014.8

David, L. A., and Alm, E. J. (2011). Rapid evolutionary innovation during an Archaean genetic expansion. *Nature* 469, 93–96. doi: 10.1038/nature09649

Delgado-Baquerizo, M., Giaramida, L., Reich, P. B., Khachane, A. N., Hamonts, K., Edwards, C., et al. (2016a). Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. *J. Ecol.* 104, 936–946. doi: 10.1111/1365-2745.12585

Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Jeffries, T. C., Gaitan, J. J., Encinar, D., et al. (2016b). Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat. Commun.* 7:10541. doi: 10.1038/ncomms10541

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *BMC Bioinformatics* 13:113. doi: 10.1186/1471-2105-13-113

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072. doi: 10.1128/AEM.03006-05

Downing, A. L. (2005). Relative effects of species composition and richness on ecosystem properties in ponds. *Ecology* 86, 701–715. doi: 10.1890/03-0239

Etesami, H., and Beattie, G. A. (2018). Mining halophytes for plant growth-promoting halotolerant bacteria to enhance the salinity tolerance of non-halophytic crops. *Front. Microbiol.* 9:148. doi: 10.3389/fmicb.2018.00148

Fierer, N., Ladau, J., Clemente, J. C., Leff, J. W., Owens, S. M., Pollard, K. S., et al. (2013). Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. *Science* 342, 621–624. doi: 10.1126/science.1243768

Fierer, N., Lauber, C. L., Ramirez, K. S., Zaneveld, J., Bradford, M. A., and Knight, R. (2012a). Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J.* 6:1007. doi: 10.1038/ismej.2011.159

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., et al. (2012b). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21390–21395. doi: 10.1073/pnas.1215210110

Fuhrman, J. A., Steele, J. A., Hewson, I., Schwalbach, M. S., Brown, M. V., Green, J. L., et al. (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7774–7778. doi: 10.1073/pnas.0803070105

Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2014). Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* 43, D261–D269. doi: 10.1093/nar/gku1223

Gamfeldt, L., Hillebrand, H., and Jonsson, P. R. (2008). Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology* 89, 1223–1231. doi: 10.1890/06-2091.1

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., and Egozcue, J. J. (2017). Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* 8:2224. doi: 10.3389/fmicb.2017.02224

Hall, E. K., Bernhardt, E. S., Bier, R. L., Bradford, M. A., Boot, C. M., Cotner, J. B., et al. (2018). Understanding how microbiomes influence the systems they inhabit. *Nat. Microbiol.* 3, 977–982. doi: 10.1038/s41564-018-0201-z

Hector, A., and Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature* 448:188. doi: 10.1038/nature05947

Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *Am. Nat.* 163, 192–211. doi: 10.1086/381004

Hu, Y., Wang, L., Tang, Y., Li, Y., Chen, J., Xi, X., et al. (2014). Variability in soil microbial community and activity between coastal and riparian wetlands in the Yangtze River estuary–Potential impacts on carbon sequestration. *Soil Biol. Biochem.* 70, 221–228. doi: 10.1016/j.soilbio.2013.12.025

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2015). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi: 10.1093/nar/gkv1248

Jousset, A., Eisenhauer, N., Merker, M., Mouquet, N., and Scheu, S. (2016). High functional diversity stimulates diversification in experimental microbial communities. *Sci. Adv.* 2:e1600124. doi: 10.1126/sciadv.1600124

Jover, L. F., Effler, T. C., Buchan, A., Wilhelm, S. W., and Weitz, J. S. (2014). The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nat. Rev. Microbiol.* 12, 519–528. doi: 10.1038/nrmicro3289

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2015). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462. doi: 10.1093/nar/gkv1070

Louca, S., Parfrey, L. W., and Doebeli, M. (2016b). Decoupling function and taxonomy in the global ocean microbiome. *Science* 353, 1272–1277. doi: 10.1126/science.aaf4507

Louca, S., Jacques, S. M. S., Pires, A. P. F., Leal, J. S., Srivastava, D. S., Parfrey, L. W., et al. (2016a). High taxonomic variability despite stable functional structure across microbial communities. *Nat. Ecol. Evol.* 1:0015. doi: 10.1038/s41559-016-0015

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'connor, M. I., et al. (2018). Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* 2, 936–943. doi: 10.1038/s41559-018-0519-1

Lozupone, C. A., and Knight, R. (2007). Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11436–11440. doi: 10.1073/pnas.0611525104

McGill, B. J., Enquist, B. J., Weiher, E., and Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends Ecol. Evol.* 21, 178–185. doi: 10.1016/j.tree.2006.02.002

Meyer, F., Paarmann, D., D'souza, M., Olson, R., Glass, E. M., Kubal, M., et al. (2008). The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386. doi: 10.1186/1471-2105-9-386

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

Morrissey, E. M., and Franklin, R. B. (2015). Evolutionary history influences the salinity preference of bacterial taxa in wetland soils. *Front. Microbiol.* 6:1013. doi: 10.3389/fmicb.2015.01013

Nelson, M. B., Martiny, A. C., and Martiny, J. B. (2016). Global biogeography of microbial nitrogen-cycling traits in soil. *Proc. Natl. Acad. Sci. U.S.A.* 113, 8033–8040. doi: 10.1073/pnas.1601070113

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2013). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226

Pavloudi, C., Kristoffersen, J. B., Oulas, A., De Troch, M., and Arvanitidis, C. (2017). Sediment microbial taxonomic and functional diversity in a natural salinity gradient challenge Remane's "species minimum" concept. *PeerJ* 5:e3687. doi: 10.7717/peerj.3687

Peter, H., Beier, S., Bertilsson, S., Lindström, E. S., Langenheder, S., and Tranvik, L. J. (2011). Function-specific response to depletion of microbial diversity. *ISME J.* 5:351. doi: 10.1038/ismej.2010.119

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864

Ramírez-Flandes, S., González, B., and Ulloa, O. (2019). Redox traits characterize the organization of global microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 116, 3630–3635. doi: 10.1073/pnas.1817554116

Schmidt, D. J., Pouyat, R., Szlavecz, K., Setälä, H., Kotze, D. J., Yesilonis, I., et al. (2017). Urbanization erodes ectomycorrhizal fungal diversity and may cause microbial communities to converge. *Nat. Ecol. Evol.* 1:0123. doi: 10.1038/s41559-017-0123

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12:R60. doi: 10.1186/gb-2011-12-6-r60

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Suttle, C. A. (2007). Marine viruses – major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750

Tatusova, T., Ciufo, S., Fedorov, B., O'neill, K., and Tolstoy, I. (2013). RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res.* 42, D553–D559. doi: 10.1093/nar/gkt1274

Torsvik, V., and Øvreås, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol.* 5, 240–245. doi: 10.1016/S1369-5274(02)00324-7

Tringe, S. G., Von Mering, C., Kobayashi, A., Salamov, A. A., Chen, K., Chang, H. W., et al. (2005). Comparative metagenomics of microbial communities. *Science* 308, 554–557. doi: 10.1126/science.1107851

Wellington, E. M., Berry, A., and Krsek, M. (2003). Resolving functional diversity in relation to microbial community structure in soil: exploiting genomics and stable isotope probing. *Curr. Opin. Microbiol.* 6, 295–301. doi: 10.1016/S1369-5274(03)00066-3

Zhou, J., Deng, Y., Luo, F., He, Z., and Yang, Y. (2011). Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO2. *mBio* 2:e00122–11. doi: 10.1128/mBio.00122-11

Zhou, J., Deng, Y., Shen, L., Wen, C., Yan, Q., Ning, D., et al. (2016). Temperature mediates continental-scale diversity of microbes in forest soils. *Nat. Commun.* 7:12083. doi: 10.1038/ncomms12083

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.

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