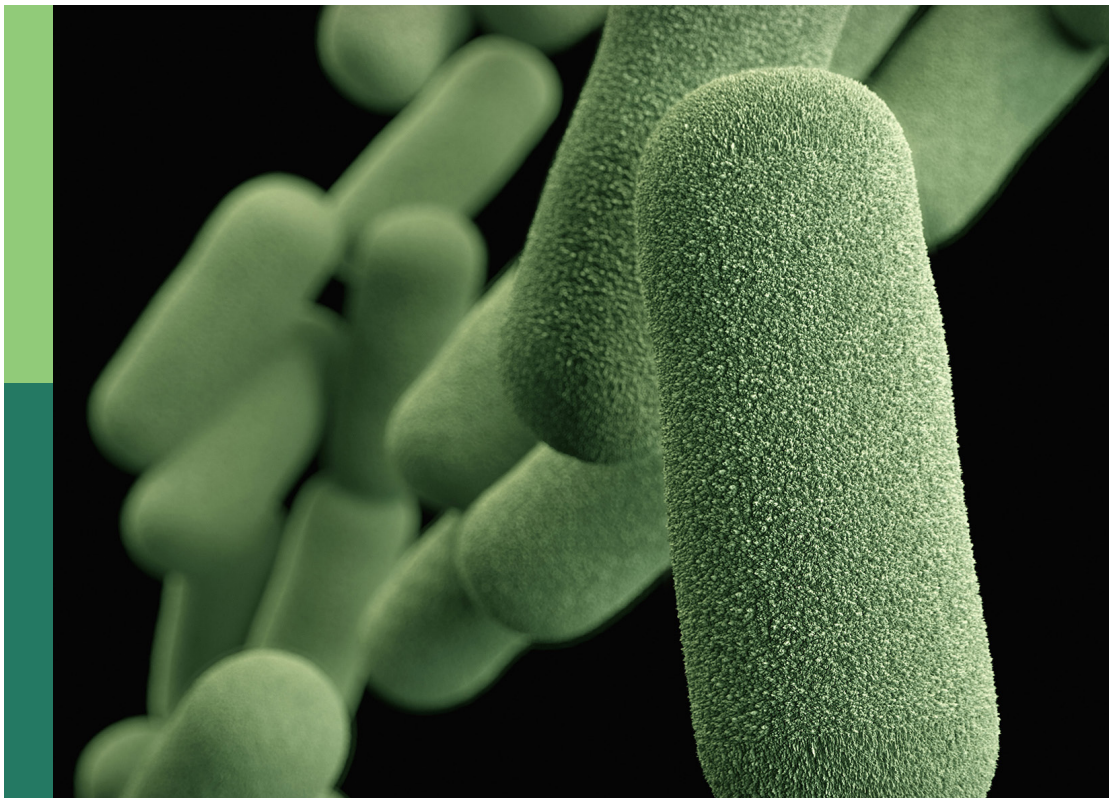
frontiers | Research Topics

Women in environmental microbiomes

Edited by
Rebecca C. Mueller and Catherine Gehring

Published in
Frontiers in Microbiomes



ISSN 1664-8714
ISBN 978-2-8325-5976-5
DOI 10.3389/978-2-8325-5976-5

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Women in environmental microbiomes

Topic editors

Rebecca C. Mueller — Invasive Species and Pollinator Health Research Unit, Western Regional Research Center, Agricultural Research Service (USDA), United States

Catherine Gehring — Northern Arizona University, United States

Citation

Mueller, R. C., Gehring, C., eds. (2025). *Women in environmental microbiomes*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-5976-5

Table of contents

frontiers | Frontiers in Microbiomes

TYPE Editorial
PUBLISHED 24 January 2025
DOI 10.3389/frmbi.2025.1537069

OPEN ACCESS

EDITED AND REVIEWED BY
Helen L. Hayden,
The University of Melbourne, Australia

*CORRESPONDENCE
Rebecca C. Mueller
rebecca.mueller@usda.gov
Catherine A. Gehring
catherine.gehring@nau.edu

RECEIVED 29 November 2024
ACCEPTED 08 January 2025
PUBLISHED 24 January 2025

CITATION
Mueller RC and Gehring CA (2025) Editorial: Women in environmental microbiomes. *Front. Microbiomes* 4:1537069. doi: 10.3389/frmbi.2025.1537069

Editorial: Women in environmental microbiomes

Rebecca C. Mueller[1]* and Catherine A. Gehring[2]*

[1]Invasive Species and Pollinator Health Unit, Western Regional Research Center, United States Department of Agriculture (USDA) Agricultural Research Service, Albany, CA, United States, [2]Department of Biological Sciences and Center for Adaptable Western Landscapes, Northern Arizona University, Flagstaff, AZ, United States

KEYWORDS

industrial waste, climate change, human health, host traits, agriculture, bioremediation

Editorial on the Research Topic
Women in environmental microbiomes

Diversity and productivity scale across multiple systems, from ecosystems to workplaces (Saxena, 2014, https://futurumcareers.com/why-is-diversity-important-for-productivity). The productivity, novelty, and impact of gender-diverse research groups exceed those with less equity (Sarabi and Smith, 2023), indicating that gender biases that exclude women from STEM careers are limiting the scope of scientific breakthroughs from some of the best potential scientists. Multiple lines of evidence suggest that female scientists are subjected to discrimination during their careers. Despite progress in earning advanced degrees in STEM, women remain under-represented in higher level career positions (Ross et al., 2022; Llorens et al., 2021), particularly tenure-track positions in academia (Casad et al., 2021). Women are less likely to be credited in scientific papers for research done in teams, particularly in high impact research (Ross et al., 2022) and receive more negative reviews both when applying for intramural grants (Witteman et al., 2019) and submitting manuscripts for publication (Hagan et al., 2020). Likely due to societal expectations, female scientists show less willingness to self-promote compared to male scientists. Relative to men, women are less likely to cite their prior work in subsequent publications (King et al., 2017) and undersell their abilities in self-evaluations when questions are focused on tasks generally perceived as male-type (Exley and Kessler, 2022). Biases against female scientists are often perpetuated by women. Female faculty showed similar bias against women as male faculty during the hiring process (Moss-Racusin et al., 2012), emphasizing the need for mentorship activities that reduce biases based on gender to encourage women's participation in STEM at all career stages (Stout et al., 2011). Despite these limitations, signs of progress in gender equity in STEM are emerging; in the last decades, representation for female authors in scientific publications (Huang et al., 2020) and prestigious honorary societies such as the National Academy of Sciences (Card et al., 2023) have increased, along with increasing willingness to hire female faculty (Williams and Ceci, 2015).

This Research Topic, titled "Women in Environmental Microbiomes", is part of a series of Research Topics launched by Frontiers to celebrate International Women's Day (March 8 each year) and includes nine publications with a female researcher as the first author or the corresponding author; more than 40 female authors contributed to these studies. The featured papers illustrate how women scientists, from students to leaders of well-established research labs, are furthering the study of microbiomes. The goal of *Frontiers in Microbiomes* is to "advance our understanding of how microbiomes generate positive or

negative outcomes for their hosts and environments". This goal reflects increasing recognition that a diverse suite of microbes occurs in every environment where they influence processes at every level of biological organization. The term "microbiome" is now often recognized by the public because of pioneering research on humans that began with microbial community characterization, expanded to the mechanisms underpinning host-microbiome interactions, identifying linkages between a diversity of human conditions and microbial communities that led to novel treatments (Rackaityte and Lynch, 2020). Some of the earliest work investigating connections between consortia of microbes and human health was led by a woman, Dr. Abigail Salyers, who successfully demonstrated the importance of microbes other than model taxa that were the focus of previous research (Whitaker and Barton, 2018). The rapidly growing study of the plant microbiome including its metabolite production also is leading to new understanding and applications, particularly in agriculture (Compant et al., 2025; Saikkonen et al., 2020). Expanded development and deployment of 'omics tools and world-wide collaboration has begun to provide a synthetic understanding of the forces that shape the microbiome of the planet (Thompson et al., 2017).

Our view of the elements of the microbiome and the techniques available to study it also have expanded over the past two decades, changes that are illustrated in this Research Topic. While early microbiome research focused on bacteria, the field now investigates and acknowledges the importance of consortia of viruses, archaea, fungi and protists, often in combination. For example, two studies in this special feature explore how microbiome analysis at different temporal or spatial scales within a common system yielded insights into broader relationships. Sacharow et al. used amplicon sequencing to explore patterns of Cercozoan protist diversity in spring barley (*Hordeum vulgare*) comparing leaf, root and soil compartments at two different developmental stages. Protists are arguably the least studied members of the microbiome though their complex roles in the rhizosophere and plant health are increasingly documented (Xiong et al., 2020; Bahroun et al., 2021). Sacharow et al. found strong differences in protist communities among leaves, roots and soils but few differences related to plant developmental stage. Both sampling location and time were important to the bacterial and viral communities of the River Erpe examined by Rodriguez-Ramos et al. Repeated sampling and metagenomic analysis of surface and porewater compartments showed distinct communities associated with each of the compartments and identified highly consistent viral communities over time in surface waters but not in porewater (Rodriguez-Ramos et al.). These studies illustrate the value of complex sampling designs that explore microbial diversity across multiple gradients simultaneously.

The effect of host genetics and host traits on microbiome composition and function has been an important area of study, particularly in human health and agriculture. Two papers in this Research Topic highlight the value of these types of studies while also exploring relevant societal issues - our poor understanding of women's health issues and the importance of assessing changes to the soil microbiome following the planting of transgenic crops. Zheng et al. used shotgun sequencing of the bacteria and virus/phage communities of the vaginal microbiome of women with and without Polycystic Ovarian Syndrome (PCOS) and obesity to show that vaginal dysbiosis was associated with a decrease in phages alongside increased bacterial diversity (Zheng et al.). In contrast, the vaginal microbiomes of women without PCOS and who were not obese were similar to one another and less diverse. This study advances our understanding of an important health concern for many women as PCOS is a leading cause of infertility globally, yet remains poorly diagnosed and understood (WHO - https://www.who.int/news-room/fact-sheets/detail/polycystic-ovary-syndrome). In fact, many diseases that affect women are poorly understood potentially due to underfunding of research on those diseases relative to the death and disability they cause (Smith, 2023). Similarly, generating sufficient food to feed a growing global population requires an understanding of the potential consequences of leveraging emerging technologies to boost production. As we increasingly use genetic tools to alter the traits of important agricultural crops, assessment of changes in microbiome composition and function are critical to ensuring the safety of these transgenic crops for large scale planting. As of July 2024, > 90% of U.S. corn, upland cotton, and soybeans were produced using genetically engineered varieties, with genetic engineering primarily focused on herbicide tolerance and insect resistance (https://www.ers.usda.gov/data-products/adoption-of-genetically-engineered-crops-in-the-united-states/recent-trends-in-ge-adoption/). One of several concerns of this widespread use of transgenic crops is an inadvertent and potentially detrimental impact on non-target species that could alter soil biodiversity and ecosystem function. Yang et al. examined the effect of a different type of genetic modification of cotton, the insertion of *ScALDH21* from a desiccation-tolerant moss, on soil bacterial and fungal communities. They observed that the diversity and dominant fungal and bacterial communities at both the phylum and genus level were similar between transgenic and non-transgenic genotypes. The differences they observed in microbial communities were associated with soil properties rather than plant genotype. Although these two studies arrive at different answers regarding host traits, they both carefully document a range of possible influences on microbiome composition to strengthen their conclusions.

The importance of the microbiome in accurately predicting how ecosystems respond to both biotic and abiotic perturbations is particularly important now as rapid environmental changes are occurring in many areas of the world. For example, in 2024, South America had the second driest October on record while North America had the 12th driest. Dry conditions were accompanied by record high temperatures that exacerbated dry conditions resulting in severe drought (NOAA Global Drought Monitor https://gdis-noaa.hub.arcgis.com/pages/drought-monitoring). Goemann et al. combined amplicon sequencing of archaea, bacteria and fungi with metabolomics of root exudates to examine the impact of differing levels of drought stress on the rhizobiome of a widespread prairie grass, *Bouteloua gracilis*, identifying drought-responsive clades in all microbial groups, along with metabolites that were more abundant in drought-stressed plants. Natural disturbances such as volcanic eruptions also can alter microbial

communities to varying degrees depending on the intensity of disturbance. Maltz et al. used amplicon sequencing to document the responses of fungi, arbuscular mycorrhizal fungi, archaea and bacteria to different landscapes affected by the eruption of Mount St. Helens, finding that old growth and clearcut forests differed in fungal community composition, but not diversity while areas of the pumice plain colonized by lupine had higher diversity of particular guilds of arbuscular mycorrhizal fungi in plots with historic access to pocket gophers. Studies on the soil microbiome also focus increasingly on restoration of microbes that have been reduced or lost due to disturbance. Markovchick et al. used field experiments to examine the legacy effects of an invasive plant, mediated through soil modifications, to subsequent success of native species. They found that active restoration of altered microbial communities substantially improved native plant survival. Understanding microbiome responses and how they feed back to affect native vegetation is critically important to predicting future plant distributions and to implementing restoration following disturbance (Coban et al., 2022).

Many industrial applications depend on microbial metabolism, including wastewater treatment and production of biofuels, but these communities often have been treated as a "black box", with little understanding of the organisms or functional pathways that drive these processes. Harnessing the metabolic potential of the microbial community would be best applied by identifying the organisms best suited for these services. Pharmaceutical residues are now found worldwide in aquatic and terrestrial ecosystems (Patel et al., 2019). Due to their harmful effects on flora and fauna, identifying efficient mechanisms for removal from wastewater is critical. Bodle et al. examined the role of microbial activity in the degradation of pharmaceuticals commonly released into wastewater streams, and found initial degradation declined along with the abundance of key bacterial and archaeal families. Similarly, fiber byproducts from agriculture and forestry are waste products that can be used to generate biofuels, chemicals and materials. Wong et al. used metagenomic sequencing of communities from an industrial anaerobic digester to characterize carbohydrate active enzymes and found that a small subset of the community participated in lignocellulose degradation, including poorly characterized lineages. Their results suggest there is still unlocked potential of microbes that would improve the transformation efficiency of these byproducts. A better understanding of the taxonomy and function of communities in biotechnology applications will likely improve efficiencies and outcomes, demonstrating the value of a mechanistic understanding of microbiomes to refine their services for human well-being.

We thank the authors, reviewers and subject editors who contributed to this Research Topic and helped highlight the important contributions of women across the globe to the field of microbiome research.

Author contributions

RM: Writing – original draft, Writing – review & editing. CG: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. CG was supported by NSF DEB-2017877 and the Lucking Family Professorship.

Acknowledgments

We thank the authors who submitted their work; it was a pleasure to read contributions by women across such diverse topics. We also thank the reviewers for their professional and helpful assessments of the submitted articles and the staff of *Frontiers* for their assistance.

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

References

Bahroun, A., Jousset, A., Mrabet, M., Mhamdi, R., and Mhadhbi, H. (2021). Protists modulate *Fusarium* root rot suppression by beneficial bacteria. *Appl. Soil Ecol.* 168, 104158. doi: 10.1016/j.apsoil.2021.10415

Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2023). Gender gaps at the academies. *Proc. Natl. Acad. Sci.* 120, e2212421120. doi: 10.1073/pnas.2212421120

Casad, B. J., Franks, J. E., Garasky, C. E., Kittleman, M. M., Roesler, A. C., Hall, D. Y., et al. (2021). Gender inequality in academia: Problems and solutions for women faculty in STEM. *J. Neurosci. Res.* 99, 13–23. doi: 10.1002/jnr.24631

Coban, O., De Deyn, G., and van der Ploeg, M. (2022). Soil microbiota as game-changers in restoration of degraded lands. *Science* 375, abe0725. doi: 10.1126/science.abe0725

Compant, S., Cassan, F., Kostić, T., Johnson, L., Brader, G., Trognitz, F., et al. (2025). Harnessing the plant microbiome for sustainable crop production. *Nat. Rev. Microbiol.* 23, 9-23. doi: 10.1038/s41579-024-01079-1

Exley, C. L., and Kessler, J. B. (2022). The gender gap in self-promotion. *Q. J. Economics* 137, 1345–1381. doi: 10.1093/qje/qjac003

Hagan, A. K., Topçuoğlu, B. D., Gregory, M. E., Barton, H. A., and Schloss, P. D. (2020). Women are underrepresented and receive differential outcomes at ASM journals: a six-year retrospective analysis. *MBio* 11, 10–128. doi: 10.1128/mBio.01680-20

Huang, J., Gates, A. J., Sinatra, R., and Barabási, A. L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proc. Natl. Acad. Sci.* 117, 4609–4616. doi: 10.1073/pnas.1914221117

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., and West, J. D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius* 3, 2378023117738903. doi: 10.1177/2378023117738903

Llorens, A., Tzovara, A., Bellier, L., Bhaya-Grossman, I., Bidet-Caulet, A., Chang, W. K., et al. (2021). Gender bias in academia: A lifetime problem that needs solutions. *Neuron* 109, 2047–2074. doi: 10.1016/j.neuron.2021.06.002

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci.* 109, 16474–16479. doi: 10.1073/pnas.1211286109

Patel, M., Kumar, R., Kishor, K., Mlsna, T., Pittman, C. U. Jr., and Mohan, D. (2019). Pharmaceuticals of emerging concern in aquatic systems: chemistry, occurrence, effects, and removal methods. *Chem. Rev.* 119, 3510–3673. doi: 10.1021/acs.chemrev.8b00299

Rackaityte, E., and Lynch, S. V. (2020). The human microbiome in the 21st century. *Nat. Commun.* 11, 5256. doi: 10.1038/s41467-020-18983-8

Ross, M. B., Glennon, B. M., Murciano-Goroff, R., Berkes, E. G., Weinberg, B. A., and Lane, J. I. (2022). Women are credited less in science than men. *Nature* 608, 135–145. doi: 10.1038/s41586-022-04966-w

Saikkonen, K., Nissinen, R., and Helander, M. (2020). Toward comprehensive plant microbiome research. *Front. Ecol. Evol.* 8. doi: 10.3389/fevo.2020.00061

Sarabi, Y., and Smith, M. (2023). Gender diversity and publication activity—an analysis of STEM in the UK. *Res. Eval.* 32, 321–331. doi: 10.1093/reseval/rvad008

Saxena, A. (2014). Workforce diversity: A key to improve productivity. *Proc. Economics Finance* 11, 76–85. doi: 10.1016/S2212-5671(14)00178-6

Smith, K. (2023). Women's health research lacks funding - in a series of charts. *Nature* 617, 28–29. doi: 10.1038/d41586-023-01475-2

Stout, J. G., Dasgupta, N., Hunsinger, M., and McManus, M. A. (2011). STEMing the tide: using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *J. Pers. Soc. Psychol.* 100, 255. doi: 10.1037/a0021385

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Jansson, J. K., Gilbert, J. A., et al. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551, 457–463. doi: 10.1038/nature24621

Whitaker, R. J., and Barton, H. A. (2018). "Abigail Salyers: an almost unbeatable force," in *Women in Microbiology*, *1st ed* (American Society for Microbiology Press), 243–251.

Williams, W. M., and Ceci, S. J. (2015). National hiring experiments reveal 2: 1 faculty preference for women on STEM tenure track. *Proc. Natl. Acad. Sci.* 112, 5360–5365. doi: 10.1073/pnas.1418878112

Witteman, H. O., Hendricks, M., Straus, S., and Tannenbaum, C. (2019). Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. *Lancet* 393, 531–540. doi: 10.1016/S0140-6736(18)32611-4

Xiong, W., Song, Y., Yang, K., Gu, Y., Wei, Z., Kowalchuk, G. A., et al. (2020). Rhizosphere protists are key determinants of plant health. *Microbiome* 8, 27. doi: 10.1186/s40168-020-00799-9

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 09 August 2023
DOI 10.3389/frmbi.2023.1199766

OPEN ACCESS

EDITED BY
Jun Zhao,
Nanjing Normal University, China

REVIEWED BY
Neslihan Taş,
Berkeley Lab (DOE), United States
Minglei Ren,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Kelly C. Wrighton
kelly.wrighton@colostate.edu

RECEIVED 04 April 2023
ACCEPTED 14 July 2023
PUBLISHED 09 August 2023

CITATION
Rodríguez-Ramos J, Oliverio A, Borton MA, Danczak R, Mueller BM, Schulz H, Ellenbogen J, Flynn RM, Daly RA, Schopflin L, Shaffer M, Goldman A, Lewandowski J, Stegen JC and Wrighton KC (2023) Spatial and temporal metagenomics of river compartments reveals viral community dynamics in an urban impacted stream.
Front. Microbiomes 2:1199766.
doi: 10.3389/frmbi.2023.1199766

Spatial and temporal metagenomics of river compartments reveals viral community dynamics in an urban impacted stream

Josué Rodríguez-Ramos[1], Angela Oliverio[1,2], Mikayla A. Borton[1,3], Robert Danczak[3], Birgit M. Mueller[4], Hanna Schulz[4,5], Jared Ellenbogen[1], Rory M. Flynn[1], Rebecca A. Daly[1], LeAundra Schopflin[1], Michael Shaffer[1], Amy Goldman[3], Joerg Lewandowski[4,5], James C. Stegen[3] and Kelly C. Wrighton[1]*

[1]Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO, United States, [2]Department of Biology, Syracuse University, Syracuse, NY, United States, [3]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, United States, [4]Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany, [5]Humboldt University, Berlin, Germany

Although river ecosystems constitute a small fraction of Earth's total area, they are critical modulators of microbially and virally orchestrated global biogeochemical cycles. However, most studies either use data that is not spatially resolved or is collected at timepoints that do not reflect the short life cycles of microorganisms. To address this gap, we assessed how viral and microbial communities change over a 48-hour period by sampling surface water and pore water compartments of the wastewater-impacted River Erpe in Germany. We sampled every 3 hours resulting in 32 samples for which we obtained metagenomes along with geochemical and metabolite measurements. From our metagenomes, we identified 6,500 viral and 1,033 microbial metagenome assembled genomes (MAGs) and found distinct community membership and abundance associated with each river compartment (e.g., *Competibacteraceae* in surfacewater and *Sulfurimonadaceae* in pore water). We show that 17% of our viral MAGs clustered to viruses from other ecosystems like wastewater treatment plants and rivers. Our results also indicated that 70% of the viral community was persistent in surface waters, whereas only 13% were persistent in the pore waters taken from the hyporheic zone. Finally, we predicted linkages between 73 viral genomes and 38 microbial genomes. These putatively linked hosts included members of the *Competibacteraceae*, which we suggest are potential contributors to river carbon and nitrogen cycling via denitrification and nitrogen fixation. Together, these findings demonstrate that members of the surface water microbiome from this urban river are stable over multiple diurnal cycles. These temporal insights raise important considerations for ecosystem models attempting to constrain dynamics of river biogeochemical cycles.

KEYWORDS

phage, time-series, auxiliary metabolic genes, hyporheic zone, genome, biogeochemistry, stability, biogeography

Introduction

Rivers are key modulators of global biogeochemical cycles and provide a dynamic, moving passageway between terrestrial and aquatic ecosystems (Allen and Pavelsky, 2018). Corresponding to ~7% of global CO_2 and ~5% of global CH_4 emissions per year, rivers contribute up to 2,508 Tg $yr^{-1}$of carbon dioxide (CO_2), and ~30.5 Tg yr^{-1} of methane (CH_4) (Villa et al., 2020; Rosentreter et al., 2021; Friedlingstein et al., 2022; Liu et al., 2022). Microbial communities are key orchestrators of carbon and nitrogen transformations in rivers, where they contribute between 40-90% of hyporheic zone respiration (Pusch and Schwoerbel, 1994; Naegeli and Uehlinger, 1997; Rodríguez-Ramos et al., 2022). Despite a general understanding of the importance of microbial metabolism, river viral communities and their impacts on microbial communities remain poorly described.

Viruses are the most abundant organism on the planet, with estimates of up to 10^{31} viral particles worldwide (Hendrix et al., 1999; Munn, 2006; Bar-On et al., 2018; Mushegian, 2020). These viral predators are mostly studied in marine ecosystems, where viruses can lyse 20-40% of bacteria daily (Weinbauer, 2004; Weinbauer and Rassoulzadegan, 2004; Suttle, 2007; Chow and Suttle, 2015; Guidi et al., 2016) and play key roles reprogramming their bacterial hosts with ecosystem-wide consequences (Sullivan et al., 2006; Anantharaman et al., 2014; Hurwitz and U'Ren, 2016). Although research has mostly focused on marine ecosystems, recent efforts have been made to expand our knowledge of natural viral communities in freshwater aquatic environments like lakes (Roux et al., 2017; Berg et al., 2021) and estuaries (Hewson et al., 2001; Cissoko et al., 2008). Early studies in these systems have shown viral like particle (VLP) abundances and viral productivity (i.e., the number of viruses produced per hour) in rivers can be equivalent, or higher, than those in marine systems (Peduzzi and Luef, 2008; Corinaldesi et al., 2010; Rowe et al., 2012; Peduzzi, 2016). Additionally, early river studies found that up to 80% of bacterial isolate strains from sediments had virulent phage that could be isolated (Lammers, 1992). Together, these foundational works highlight the importance of viral predation in regulating microbial dynamics in river ecosystems.

There are two key reasons why it remains difficult to link viral communities to river ecosystem function. First, river microbiome studies are rarely genome-resolved, both from a bacterial and viral perspective. While there is still much to explore, most information on aquatic virus dynamics pertains to oceanic studies (Vincent and Vardi, 2023), and rivers are described as one of the most underexplored aquatic ecosystem with metagenomics, second only to glacier microbiomes (Chu et al., 2020). Although the taxonomic composition of microbial communities in rivers has been well-described by 16S rRNA gene amplicon surveys (Hou et al., 2017; Nelson et al., 2019), it remains unclear how microbial membership relates to relevant ecosystem processes. Likewise, our ability to link the viral community to their respective microbial hosts, and subsequently to ecosystem biogeochemistry, remains hindered by a lack of genome-resolved studies. Second, river studies are often not temporally constrained. Although significant changes in river chemistry and hydrology are observed at seasonal periods (Tomalski et al., 2021), they are also known to change at sub-daily scales (Lundquist and Cayan, 2002; Alonso et al., 2017), particularly in human-impacted rivers affected by wastewater treatment plant effluent and reservoirs (Luo et al., 2020; Wang et al., 2021; Lu et al., 2022). This is particularly important when considering the microbial component of river systems, as microbial generations are on the scale of minutes to hours, and microbiomes can shift metabolically in hours (Wang et al., 2015; Erbilgin et al., 2017; Gibson et al., 2018). Nonetheless, river microbiome time-series are often resolved at seasonal scales (Kaevska et al., 2016; Malki et al., 2021), meaning our understanding of viral and microbial community dynamics across relevant temporal gradients (i.e. hours) remains poorly understood.

To address these knowledge gaps, we collected a finely resolved metagenomic time-series at the River Erpe near Berlin, Germany, a lowland river receiving treated wastewater. Our sampling campaign included biogeochemical measurements every 3 hours for 48 hours across both surface water (SW) and pore water (PW) compartments that were paired to metagenomics and metabolomics (Figures 1A–D). This study design provided a metagenomically resolved dataset which enabled us to interrogate how viral and microbial communities are structured across river compartments, and how this metabolic potential could modulate biogeochemical processes. Additionally, the temporal resolution of our dataset allowed us to analyze both the persistence of viral and microbial communities across compartments, as well as the individual genome stability throughout the 48 hours of sampling. Finally, by using genome-resolved metagenomics, we show that viruses can be linked to hosts in river ecosystems, and that these linkages can reveal putative interactions that may be relevant to understanding the temporal dynamics of ecosystem biogeochemistry.

Methods

Sample collection, DNA isolation, and chemical characterization

The River Erpe is highly influenced by diurnally fluctuating effluent volumes of the Münchehofe wastewater treatment plant and consists of up to 80% treated wastewater (Mueller et al., 2021). Our sampling site is in a side channel with a mean discharge of 25 l/s (Lewandowski et al., 2011; Mueller et al., 2021) (Figure 1A). For sample collection, a sampling station was set up ~1m from the shoreline of the River Erpe side channel "Rechter Randgraben" (52.476416, 13.625710), 1.6km from the wastewater treatment plant outlet leading the same water as in the main channel as previously described (Mueller et al., 2021), and in accordance to the Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDRS) protocol (Stegen and Goldman, 2018). Samples were collected on September 25, 2018. More information on the River Erpe sampling methods can be found in another publication from our team, as well as the original public data repository (Wells et al., 2019; Mueller et al., 2021). Briefly, for surface water (SW), 60ml at a time of SW were collected manually with a syringe and tubing fixed in the water column and

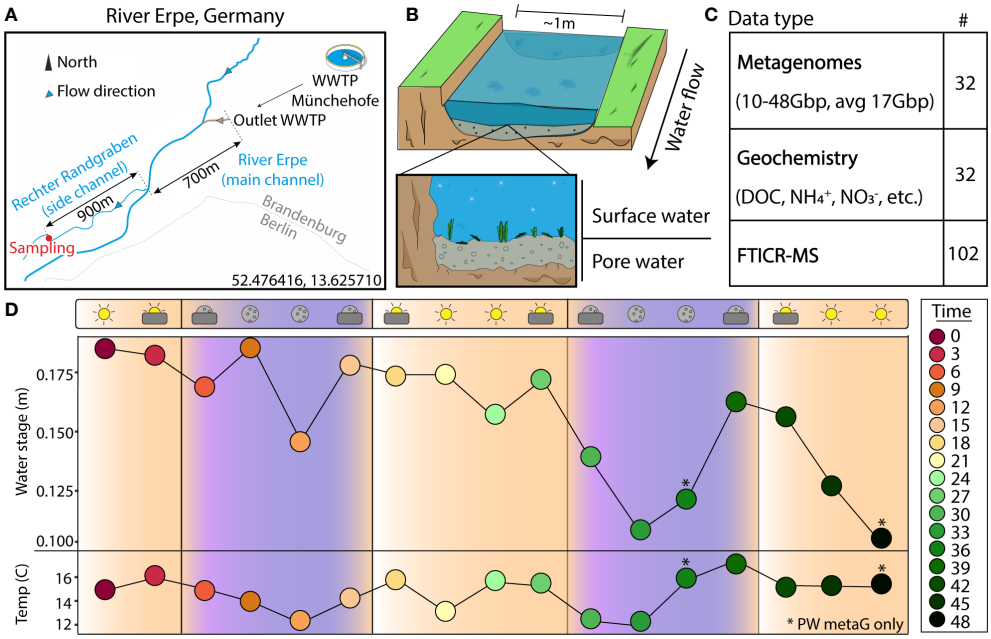


FIGURE 1
Experimental design enables a genome- and time-resolved view of microbial communities at a finely scaled resolution. **(A)** River Erpe sampling site that is located near Berlin, Germany. **(B)** Conceptual schematic of the surface and pore water compartments that were sampled as part of this research. **(C)** Table of data types that were collected as part of this sampling effort. **(D)** Sampling schematic over 48-hour period with two ecological variables (water stage, and temperature) shown across the timepoints collected. The colors and icons highlight the hour of the day when samples were collected. Asterisks (*) denote samples where only pore water metagenomes were collected

then passed through a 0.20µm filter until clogged. A cap was then put on the filter, filled with 3ml RNAlater, and refrigerated until extraction. For pore water (PW), 60ml of PW from 25cm sediment depth were collected with a stainless-steel rod in the middle of the channel. The rods were covered with a filter mesh sock over the screened area at the tip, pushed into the sediment, and equipped with a Teflon suction line. Samples were then taken by manually pulling 60ml of PW with syringes attached to the suction line and filtering them through a 0.20µm filter until clogged. The filter was then capped, filled with 3ml RNAlater, and refrigerated until extraction. Each of these processes were repeated every 3 hours over a period of 48hrs in September of 2018, resulting in 15 SW and 17 PW metagenomes. 2 SW samples failed due to lack of biomass. For DNA isolation, filters were cut into ~5mm^2 pieces and added to the bead bashing tubes of Quick-DNA Soil Microbe Microprep Kit (Zymo). The nucleic acids were then extracted according to the manufacturer protocol and sequenced at the Genomics Shared Resource Anschutz Medical Campus, Colorado. Accession numbers, total metagenomic reads, and sample sizes can be found on Supplemental Table 1 and the original data repository (Wells et al., 2019).

Chemical characterization was performed as previously described (Mueller et al., 2021). Water samples were filtered with 0.2µm polyethersulfone Sterivex for Fourier transform ion cyclotron resonance mass spectrometer (FTICR-MS) analysis or regenerated cellulose for all other analytes, then acidified to a pH of 2 with 2M HCl and stored at -18°C until analysis. Samples were analyzed at the Leibniz Institute of Freshwater Ecology and Inland Fisheries for nitrate and sulfate (ion chromatography, Metrohm 930 Compact IC Flex), ammonium and soluble reactive phosphorous (SRP) (segmented flow analyzer Skalar SAN, Skalar Analytical B.V., Netherlands), and manganese and iron (inductively coupled plasma optical emission spectrometry (ICP-OES), (ICP iCAP 6000 series, Thermo Fisher Scientific Inc.). Dissolved organic carbon (DOC) concentrations were analyzed via infrared gas analyzer (NDIR) after combustion (TOC/TN Analyzer, Shimadzu). Dissolved organic matter (DOM) data is part of the WHONDRS dataset (Wells et al., 2019) and was analyzed using a 12T Bruker SolariX FTICR-MS (Bruker, SolariX, Billerica, MA, USA) at the Environmental Molecular Sciences Laboratory in Richland, WA. Once peaks were picked using the Bruker data analysis software and formulas were assigned using Formularity (Tolić et al., 2017), DOM was classified into seven compound classes based upon hydrogen to carbon ratio (H:C), and oxygen to carbon (O:C) ratios (Kim et al., 2003). FTICR-MS analysis does not allow for a quantitative approach, therefore compound class data was analyzed qualitatively, and DOM composition was evaluated using the number of molecular formulas in every compound class as described in the original publication (Mueller et al., 2021). The biogeochemical measurements for this study can all be found on Supplemental Table 1.

Metagenome data processing and assembly

Each set of metagenomic reads were trimmed using Sickle v1.33 with default settings (Joshi NA, 2011), and assessed using FastQC (v0.11.2) (Andrews, n.d.). Trimmed reads were then assembled with either 1) metaSPAdes BBCMS pipeline (v3.13.0) (Metagenome

Assembly Workflow (v1.0.1) — NMDC Workflows 0.2a documentation, n.d.), 2) Megahit (v1.2.9) (Li et al., 2015), or 3) IDBA UD (v.1.1.0) (Peng et al., 2012). For metaSPAdes pipeline, reads were merged into a single.fa file using fq2fa (Shen et al., 2016). Then, bbcms was run with flags "mincount = 2", and "highcountfraction = 0.6", followed by metaSPAdes using kmers 33, 55, 77, 99, 127, and flag "–meta". For Megahit, reads were assembled with flags "k-min = 31", "k-max = 121", "k-step = 10", and "m = 0.4". For IDBA_UD, samples were rarefied to 25% of reads using BBMAP's reformat.sh (Bushnell, 2014) with flags "samplerate = 0.25" and "sampleseed = 1234". These 25% of subset reads were then merged into a single.fa file using fq2fa (Shen et al., 2016) and then assembled with default parameters. Assembly statistics for each sample can be found in Supplemental Table 1.

Viral identification, taxonomy, and annotations

Viral metagenome assembled genomes (vMAGs) were identified from each set of assemblies using Virsorter2 and CheckV using the established protocols.io methods (Guo et al., 2021a; Guo et al., 2021b). Resulting genomes were then screened based on VirSorter2 and checkV output for viral and host gene counts, VirSorter2 viral scores, and hallmark gene counts (Guo et al., 2021b). Viruses were then annotated with DRAM-v using the "–use_uniref" flag, and further manually curated according to the established protocol (Shaffer et al., 2020; Guo et al., 2021b). The resulting subset of 6,500 viral genomes were clustered at 95% ANI across 85% of shortest contig per MIUViG standards (Roux et al., 2018) resulting in 1,230 viral populations.

Viral taxonomic identification of viral populations was performed using protein clustering methods with vContact2 using default methods (Bin Jang et al., 2019). We supplemented the standard RefSeq v211 database containing 4,533 vMAGs with viral genomes from an additional 303 river and wastewater treatment plant metagenomes that were publicly available from 1) JGI IMG/VR (6,254 vMAGs ≥10kb), 2) two previously unpublished anaerobic digestor metagenomic datasets that were mined in-house (14,436 vMAGs ≥10kb) (https://doi.org/10.5281/zenodo.7709817), 3) a previously published wastewater treatment plant sludge database (7,443 vMAGs ≥10kb) (Shi et al., 2022), 4) a previously available reference database that included freshwater ecosystem viruses (2,032 vMAGs ≥10kb) (Rodríguez-Ramos et al., 2022), and 5) the 43 TARA Oceans Virome datasets (5,476 vMAGs ≥10kb) (Brum et al., 2015). This resulted in an additional 35,641 reference vMAGs in our network. Proteins file for all vMAGs used in the network as well as accession numbers are available on Zenodo (https://doi.org/10.5281/zenodo.7709817). Results from vContact2 can be found in Supplemental Table 2.

Viral population genome representatives were annotated using DRAM-v (Shaffer et al., 2020). To identify putative auxiliary metabolic genes (AMGs), auxiliary scores were assigned by DRAM-v to each annotated gene based on the following previously described ranking system: A gene is given an auxiliary score of 1 if there is at least one hallmark gene on both the left and right flanks, indicating the gene is likely viral. An auxiliary score of 2 is assigned when the gene has a viral hallmark gene on one flank and a viral-like gene on the other flank. An auxiliary score of 3 is assigned to genes that have a viral-like gene on both flanks (Shaffer et al., 2020; Rodríguez-Ramos et al., 2022). Genes identified by DRAM-v as being high-confidence possible AMGs (auxiliary scores 1-3) were subjected to protein modeling using Protein Homology/AnalogY Recognition Engine (PHYRE2) (Kelley et al., 2015), and manually verified. All files for vMAG quality and annotations can be found in Supplemental Table 2.

Bacterial and archaeal metagenomic binning, quality control, annotation, and taxonomy

Bacterial and archaeal genomes were binned from each set of assemblies with MetaBAT v2.12.1 (Kang et al., 2019) as previously described (Rodríguez-Ramos et al., 2022). Briefly, reads were mapped to each respective assembly to get coverage information using BBmap (Bushnell, 2014), and then MetaBAT was run with default settings on each assembly after filtering for scaffolds ≥2,500bp. Quality for each MAG was then assessed using CheckM (v1.1.2) (Parks et al., 2015). To ensure that only quality MAGs were utilized for analyses, we discarded all MAGs that were not medium quality (MQ) to high quality (HQ) according to MIMAG standards (Bowers et al., 2017), resulting in 1,033 MAGs. These MAGs were dereplicated using dRep (Olm et al., 2017) at 95% identity, resulting in 125 MAGs. These 125 MQHQ MAGs were annotated using the DRAM pipeline (Shaffer et al., 2020) as previously described (Rodríguez-Ramos et al., 2022). For taxonomic analyses, MAGs were classified using the Genome Taxonomy Database (GTDB) Toolkit v1.5.0 on November 2021 using the r202 database (Chaumeil et al., 2019). Genome quality, annotations, and taxonomy are reported in Supplemental Table 3.

Virus host linkages

To identify virus-host linkages, we used 1) CRASS (Direct Repeat/Spacer based) v1.0.1 (Skennerton et al., 2013), 2) VirHostMatcher (alignment-free oligonucleotide frequency based) v.1.0.0 (Ahlgren et al., 2017), and 3) PHIST (all-versus-all exact matches based) v.1.0.0 (Zielezinski et al., 2021). CRASS protocol and scripts used are described in detail on GitHub (see Data availability). VirHostMatcher was run with default settings, and the best possible hit for each virus was considered only if it had a d2* dissimilarity score of < 0.2. PHIST was run with flag "-k = 25", and a PHIST hit was considered only if it had a significant adjusted p-value of < 0.05. To be classified as a virus-host linkage, a virus-host pair had to be predicted by the significant consensus of both VirHostMatcher and PHIST or a virus-host pair had to have a CRASS linkage. With this consensus method, CRASS links, which were always considered good hits, agreed across 60% of predictions at the Genus level, 80% of predictions at the Order level, and 87% at

the Class level, suggesting high accuracy of consensus-only, non-CRASS linked virus-host pairs. All virus-host predictions are in Supplemental Table 2.

Genome relative abundance and normalization

To estimate the relative abundance of each vMAG and MAG, metagenomic reads for each sample were mapped to a database of vMAGs or MAGs with Bowtie2 (Langmead and Salzberg, 2012) at an identity of 95%, with minimum contig coverage of 75% and minimum depth coverage of 3x. To normalize abundances for known temporal omics data biases (Coenen et al., 2020), we performed a library size normalization of abundance tables using TMM (Robinson and Oshlack, 2010). Given that PW and SW organism abundances were drastically different in magnitude, and that abundance zeroes across compartments are likely real zeroes, vMAGs and MAGs were considered to be present if detectable in at least 10% of samples in either compartment. Organisms detected in > 10% PW samples were labeled "pore", organisms detected in > 10% SW samples were labeled "surface", organisms > 10% PW and SW samples were labeled "both", and organisms that were in < 10% SW and PW samples were removed. Based on these groups, the TMM abundances file was split into two different files, one for PW samples (n = 17) including "pore" and "both" organisms, and one for SW samples (n = 15) including "surface" and "both" organisms. Abundances for vMAGs and MAGs can be found in Supplemental Table 2, 3, and specific commands can be found on GitHub.

Temporal and statistical analyses

Temporal analyses were all performed in R with the TMM normalized abundances described above. To determine which environmental parameters were significantly driving differences across our compartments, we performed multiple regressions using envfit in the vegan R package (Oksanen et al., 2016) across multiple types of ordinations. Principal Coordinate Analysis (PCA) for biogeochemistry were done with vegan in R. Dissimilarities in community composition were calculated with the Bray-Curtis metric in vegan (Oksanen et al., 2016) for all vMAGs and MAGs that were present in >3 samples per each compartment. Nonmetric multidimensional scaling (NMDS) was then used with k = 2 dimensions for visualization. An analysis of similarity (ANOSIM) was performed using the base R stats package in order to determine community similarity between river compartments. PERMANOVA analyses were done in R using the adonis function from vegan. The NMDS ordinations of the vMAGs and MAGs were compared using the PROCRUSTES function in vegan. To visualize the relative contribution of each biogeochemical variable, we calculated the envfit vector using function ordiArrowMul and plotted them using ggplot. Shannon's H' were done using TMM normalized values with vegan in R. Species accumulation curves were done using the vegan function specaccum in R. All R code and files are available on GitHub.

To determine the relative stability of surface and pore water communities, we first calculated the differences in Bray-Curtis dissimilarity for each sample and its prior timepoint and then ran an unpaired t test to compare the mean differences across compartments with the vegan package in R. For assigning the persistence of the different genomes, we used previously established metrics to assess persistent (present in ≥ 75% of samples), intermittent (present > 25% <75% of samples), or ephemeral (present in ≤ 25% of samples) categories (Chow and Fuhrman, 2012). For establishing the abundance stability, we assessed the total number of samples in which each individual persistent genome fluctuated by ± 25% of the median relative abundance value across all samples. Then, using the established cutoffs by Fuhrman and Chow et al. (Chow and Fuhrman, 2012)., we categorized our genomes as stable (shifting in ≤ 25% of samples), intermediately stable (shifting in > 25% < 75% of samples) and unstable (shifting in ≥ 75% of samples). Fishers exact test for count data was used for assessing the significance of difference in stability metrics using fisher.test from R base stats package. The enrichment analyses for AMGs were performed using a hypergeometric test between the total AMGs in our dataset and the individual groups of AMGs present in either compartment. The code used is available on GitHub. All temporal analyses and results are in Supplemental Table 4.

To reduce the complexity of our microbial data so we could link viral and microbial communities more concretely to ecosystem biogeochemical cycling, we applied a Weighted Gene Correlation Network Analysis (WGCNA) to identify which groups of organisms co-occurred using TMM normalized values in R with package WGCNA (Langfelder and Horvath, 2008; R Core Team, 2018). A signed hybrid network was performed with a combined dataset of MAGs and vMAGs on a per-compartment basis. For SW, we used a minimum power threshold of 14 and a minimum module size of 20. For PW, we used a minimum power threshold of 8 with and a minimum module size of 20. For both networks, a reassign threshold of 0, and a merge cut height of 0.3 were used.

To link the modules to ecosystem biogeochemistry, we performed sparse partial least square regressions (sPLS) on the groups of organisms in each module. sPLS were done using TMM normalized values of co-occurring communities that resulted from WGCNA above in R with package PLS (Chung et al., 2012). Subnetwork membership was related to the overall genome significance for nitrate as described in the WGCNA tutorials document (see GitHub code) using R and the WGCNA package (Langfelder and Horvath, 2008). Full code for WGCNA and SPLS are available on GitHub along with detailed instructions and input files. Visualizations for the AMG and WGCNA figures were made using RawGraphs (Mauri et al., 2017).

Results

Metagenomics uncovers viral novelty and biogeography of River Erpe viruses

We sampled 17 pore water (PW) and 15 surface water (SW) metagenomes collected over a 48-hour period using a Eulerian

sampling scheme (i.e., at a fixed location) and collected 565.5Gbp of paired metagenomic sequencing (10.2-47.9Gbp/sample, 17.7Gbp avg.) (Figure 1 and Supplemental Table 1). Assembly of these samples revealed 6,861 viral metagenome assembled genomes (vMAGs), of which 6,500 vMAGs were ≥10kb in length and were subsequently clustered into 1,230 species-level vMAGs (Supplemental Table 2). The average vMAG genome fragment was 24,164bp (180,216bp max) in the PW, and 19,553bp (153,177bp max) in the SW (Supplemental Table 2). Viral MAG richness was consistently 8 times higher ($p < 0.01$) in the SW (845.0 ± 124.4) compared to the PW (108.3 ± 49.7) and likely drove differences ($p < 0.01$) in Shannon's diversity (H') recorded for the SW (SW = 6.05 ± 0.17, PW = 3.67 ± 0.49) (Supplemental Figure 1). In addition to our vMAGs, we identified 1033 metagenome assembled genomes (MAGs) that were dereplicated at 95% identity into 125 medium and high-quality genome representatives. Similarly, MAG richness was higher ($p < 0.01$) in the SW (SW = 62.6 ± 7.2, PW = 21.8 ± 9.0), and showed significantly different patterns ($p < 0.01$) in terms of Shannon's (H') (SW = 2.9 ± 0.17, PW = 2.6 ± 0.3) (Supplemental Figure 1).

Viruses from freshwater systems are not well sampled in the databases commonly used for taxonomic assignment in viral studies (Elbehery and Deng, 2022). To determine the extent of novel viral diversity recovered, we mined additional set of 21,022 vMAGs from a variety of freshwater, wastewater, and marine samples and added this to the original vContact2 database (Supplemental Table 2, see Materials and Methods). We then performed protein clustering of our unique 1,230 viruses with this modified aquatic database, revealing 3,030 viral clusters (VCs). This network was composed of 19,623 nodes with 679,402 edges, which was simplified to only show protein clusters that contained at least 1 vMAG from this study (Figures 2A–C).

Of our 1,230 vMAGs, 1% clustered to known taxonomic representatives of the Caudovirales Order (8 Podoviridae, 7 Siphoviridae, 3 Myoviridae). Of the remaining vMAGs, 37% clustered only to Erpe viruses, constituting 189 novel genera. An additional 41% did not cluster to any vMAG in our database and were "singletons" or "outliers". Interestingly, 17% of our total vMAGs and nearly half of our novel genera were cosmopolitan in aquatic ecosystems, meaning that while these vMAGs failed to

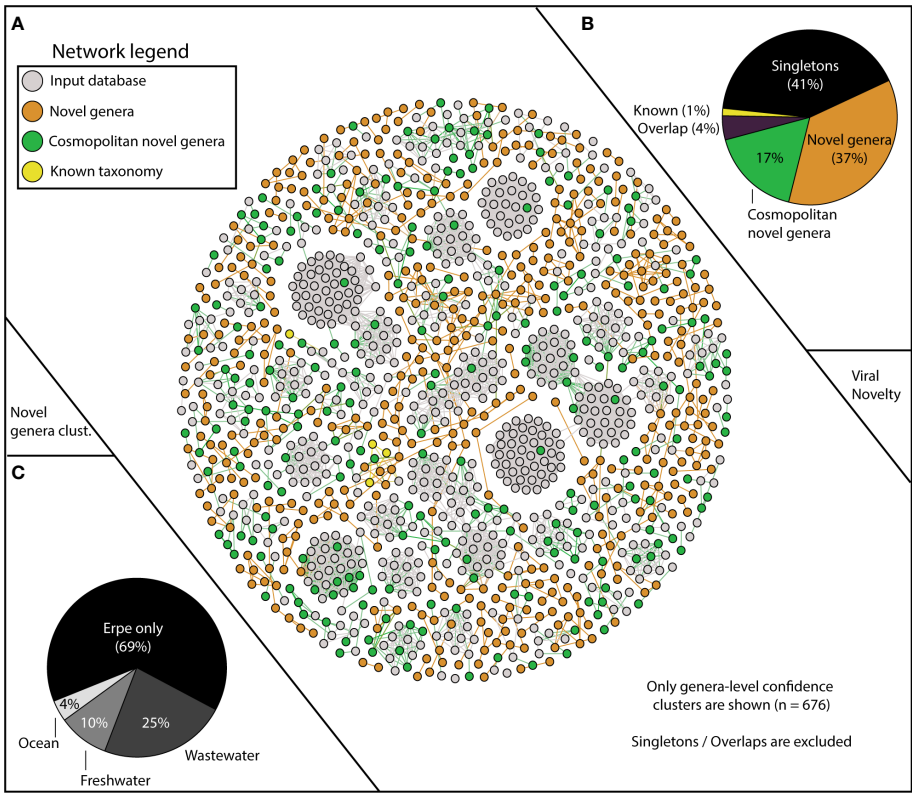


FIGURE 2
vContact2 reveals Erpe vMAG database constitutes mostly novel genera, and a portion of these are cosmopolitan. **(A)** vContact2 protein cluster (PC) similarity network where nodes represent vMAGs and edges show similarity across edges. Only high-confidence genera-level clusters are shown (n=676) with node color representing whether the vMAG pertains to our input databases (gray) or other categories assigned to vMAGs recovered here: orange shows novel genera (clustering only with Erpe genomes), green shows cosmopolitan novel genera (clustering with viruses from additional input database not from RefSeq), and yellow represents vMAGs with known taxonomy (clustering with known RefSeq vMAGs). Singletons (genomes that do not cluster with any other genomes) are excluded from the visualization (n=518). **(B)** Pie chart shows the distribution of the different categories from the vContact2 network of vMAGs recovered. "Overlap" refers to a category where vContact2 assigns a vMAG to more than one cluster but cannot confidently place in either. **(C)** Pie chart shows the proportion of vMAGs from novel genera in this study that were clustering with vMAGs from different environmental input databases.

cluster with taxonomically known strains, they did cluster with vMAGs recovered from other ecosystems (Figure 2B). Specifically, our cosmopolitan novel genera clustered with vMAGs from wastewater treatment plant sludge or effluent (n=168), other rivers surface or sediment samples (n=65), and marine samples of the TARA oceans dataset (n=25) (Figure 2C). Notably, adding these additional viral genomes reduced the total number of River Erpe vMAGs that were categorized as singletons or outliers, resulting in the addition of 49 novel genera.

Viral and microbial River Erpe microbiomes are compartment-specific

The collected biogeochemistry was significantly structured across compartments and explained a large portion of the total variation in our samples (R^2 = 0.79, p < 0.01) (Figure 3A). The surface water compartment was driven mostly by 1) the accumulation of alternative terminal electron acceptors (i.e., nitrate (NO_3^-), and sulfate (SO_4^{2-})), 2) the availability of nitrogen

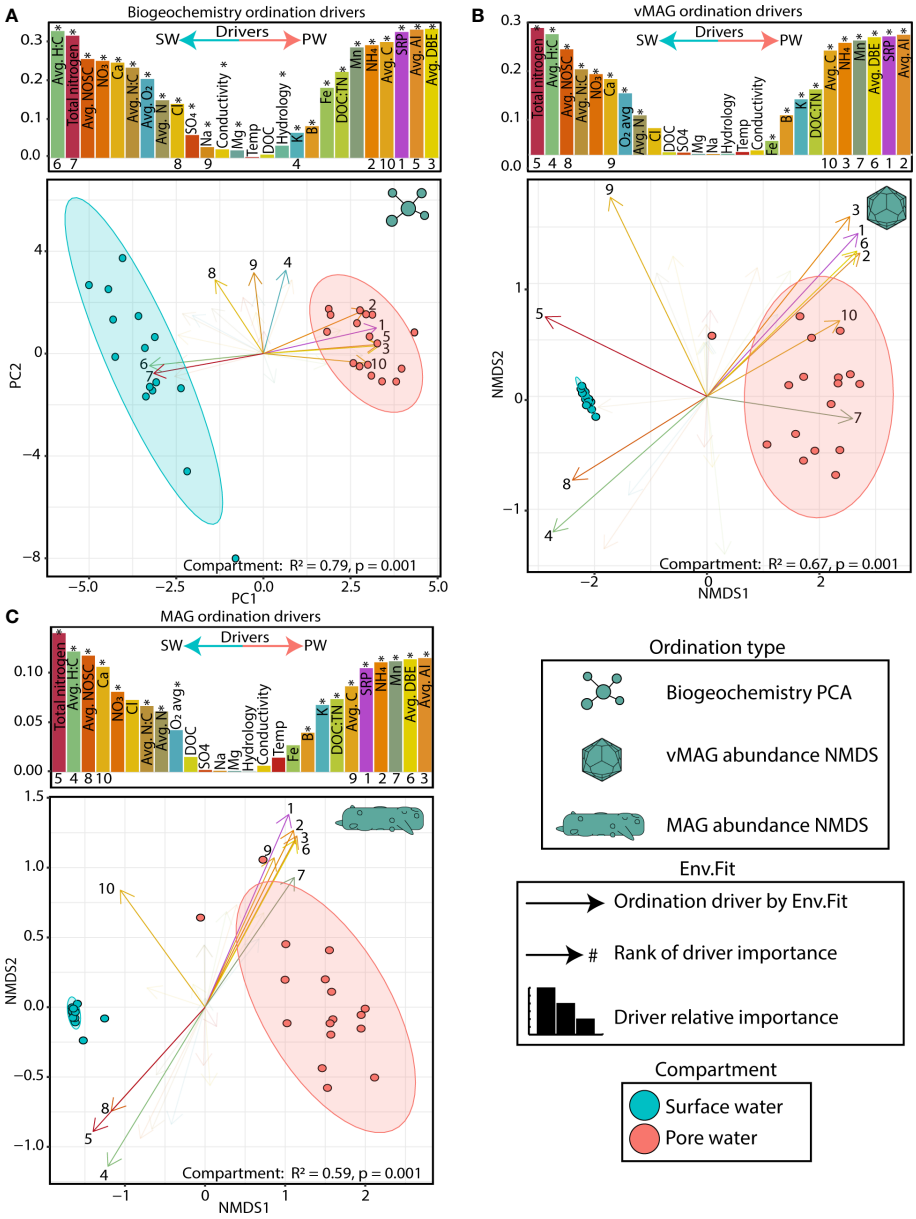


FIGURE 3
Surface and pore water compartments have distinct viral communities and distributions are driven by biogeochemistry **(A)** PCA plot of biogeochemical measurements where loadings and bars show the biogeochemical drivers per compartment. The size of bars represents the distance between the end of a loading arrow and the center of the plot. Within each bar plot, the drivers are labeled, and asterisks denote significant drivers by env.fit. The top 10 most significant drivers are numbered below each bar and are shown with solid, numbered arrows within the ordination below. **(B)** NMDS ordination of river pore water and surface water vMAG abundances with bars and arrows showing the same as in **(A)**. **(C)** NMDS ordination of river pore water and surface water MAG abundances with bars and arrows showing the same as in **(A)**. Non-compound abbreviations are: nominal oxidative state of carbon (NOSC), calcium (Ca), chlorine (Cl), sodium (Na), magnesium (Mg), dissolved organic carbon (DOC), soluble reactive phosphorous (SRP), aromaticity index (AI), and double bond equivalents (DBE). Note: NOSC values are plotted as the absolute value per value per sample (i.e., a higher SW NOSC driver value translates to a more negative NOSC measurement).

compounds (i.e., total nitrogen, avg. N), and 3) a more negative overall nominal oxidative state of carbon (NOSC) and a higher H:C ratio. Conversely, the pore water was characterized by 1) accumulation of NH_4, 2) the availability of soluble reactive phosphorous (SRP), and 3) the overall concentration of carbon (avg. C), its aromaticity index (AI), and the quantity of double bond equivalents per molecule (DBE). In summary our data indicated more oxidative conditions in the SW (Mueller et al., 2021) while the FTICR-MS data showed that SW carbon was likely more labile, accessible, and thermodynamically favorable.

To determine how viral and microbial communities were structured across these biogeochemical gradients, we recruited the time-series metagenomic reads to our viral database of 1,230 dereplicated vMAGs and 125 MAGs and then performed non-metric multidimensional scaling (NMDS) ordinations (Figures 3B, C). Like the geochemical PCA plots, PERMANOVA analyses showed that river compartment explained 67% ($p < 0.01$) and 59% ($p < 0.01$) of the variation in viral and microbial communities, respectively. The drivers of both viral and microbial communities were nearly identical in both magnitude and direction. Similarly, a PROCRUSTES analyses showed that vMAG and MAG ordinations are highly coordinated with each other (sum of squares = 0.027, corr. = 0.99, $p < 0.01$) (Supplemental Figure 2) emphasizing the expected dependencies between our identified viral and microbial communities due to our methods. Further highlighting these compartmental distinctions, the abundances of 85% of vMAGs (n = 1051) and 67% of MAGs (n = 87) were indicators of only one compartment (Supplemental Table 5). Interestingly, across both viral and microbial ordinations as well as our PCA, time only explained an additional 4-5% of the total variation, albeit significantly ($p = 0.03$, $p = 0.02$, and $p < 0.01$, respectively), likely due to long travel times and hydrological separation (Supplemental Table 5).

Temporally resolved metagenomics unveils compartment-level stability and persistence of viral and microbial communities of the River Erpe

SW metagenomic temporal samples for both vMAGs and MAGs were on average 2-fold more similar than PW by Bray-Curtis dissimilarities (BC) (vMAG t = 6.3; MAG t = 6.2, $p < 0.01$) (Figures 4A, B). We next evaluated whether the individual temporal persistence of the viral and microbial genomes shared similar patterns to the BC across compartments, and categorized members using persistence metrics that were previously established (Chow and Fuhrman, 2012). Briefly, if a viral genome was in more than 75% of the samples it was designated as persistent, between 25-75% of samples it was intermittent, and in less than 25% it was ephemeral. Of the 1,035 vMAGs detected in the SW compartment, 70% were categorized as "persistent", with the remainder being 25% intermittent and 5% ephemeral. Contrastingly, of the 374 vMAGs detected in the PW, only 11% were categorized as persistent, with the remainder being 26% intermittent and 63% ephemeral (Figures 4C, D). Similarly, the bacterial and archaeal MAGs shared comparable persistence patterns across the compartments (Figures 4E, F). Combined, these results showed that SW communities were less temporally dynamic in terms of BC and had more persistently sampled genomes than the PW.

We then assessed whether the relative abundance of persistent genomes was also temporally stable. Based on a prior study (Chow and Fuhrman, 2012), we tallied the number of samples in which persistent vMAG and MAG relative abundances exceeded ± 25% of their respective median (Figures 4G, H, Supplementary Table 4). Our results showed that both the relative abundance of vMAGs and MAGs in the SW fluctuate less over time than the PW as shown by Fishers exact t test ($p < 0.01$). Our persistence and temporal stability results supplement the observation that surface water communities in this urban stream change less over the 48-hour period than pore water communities which are more dynamic.

Genome-resolved virus-host analyses demonstrated viruses could infect highly abundant, phylogenetically diverse microbial genomes

We were able to predict hosts for 73 vMAGs, matching 30% (n = 38) of our total microbial genomes to a viral partner (Figure 5). A majority (62%) of vMAGs with host associations were from the SW compartment, with 22% of host-associated vMAGs found in the PW, and around 10% found across both compartments. MAGs that had viruses linked to them were highly abundant, with 54% of our linked vMAGs infecting hosts of the top 25% most abundant MAGs. At the phylum level, 11 of the 20 identified phyla had evidence for a viral host. Notably, all the phyla that could not be assigned a viral link had 2 or less MAG representatives, with the exception of *Desulfobacterota* which had 6 MAGs. Additionally, of the 51 *Patescibacteria* MAGs we recovered in this study, we uncovered 12 possible viral genome links, which to our knowledge is one of the few reports of possible infective agents for members of this phylum (Holmfeldt et al., 2021; Trubl et al., 2021), and is the only one thus far reported in rivers. Ultimately, nearly a third of the genera from our MAG database as defined by GTDB were successfully linked to a vMAG, providing further evidence that viral predation is likely pervasive across these river microbial communities.

To decipher the potential impacts that viral predation could have on biogeochemical cycling across the collected timeseries, we metabolically characterized the 38 viral-linked MAGs from our genome-resolved database and saw a wide array of metabolisms spanning ecosystem chemical gradients (Figure 5). Across both compartments, viruses were inferred to impact hosts that could modulate both aerobic and microaerophilic metabolism (carbon respiration), as well as anaerobic metabolisms (nitrate reduction, fumarate reduction, fermentation, and nitrogen fixation). For example, vMAGs were predicted to infect hosts with metabolisms such as methanogenesis (e.g., *Methanothrix*), and sulfur metabolisms (e.g., *Sulfurimonas*), which were encoded more

10.3389/frmbi.2023.1199766

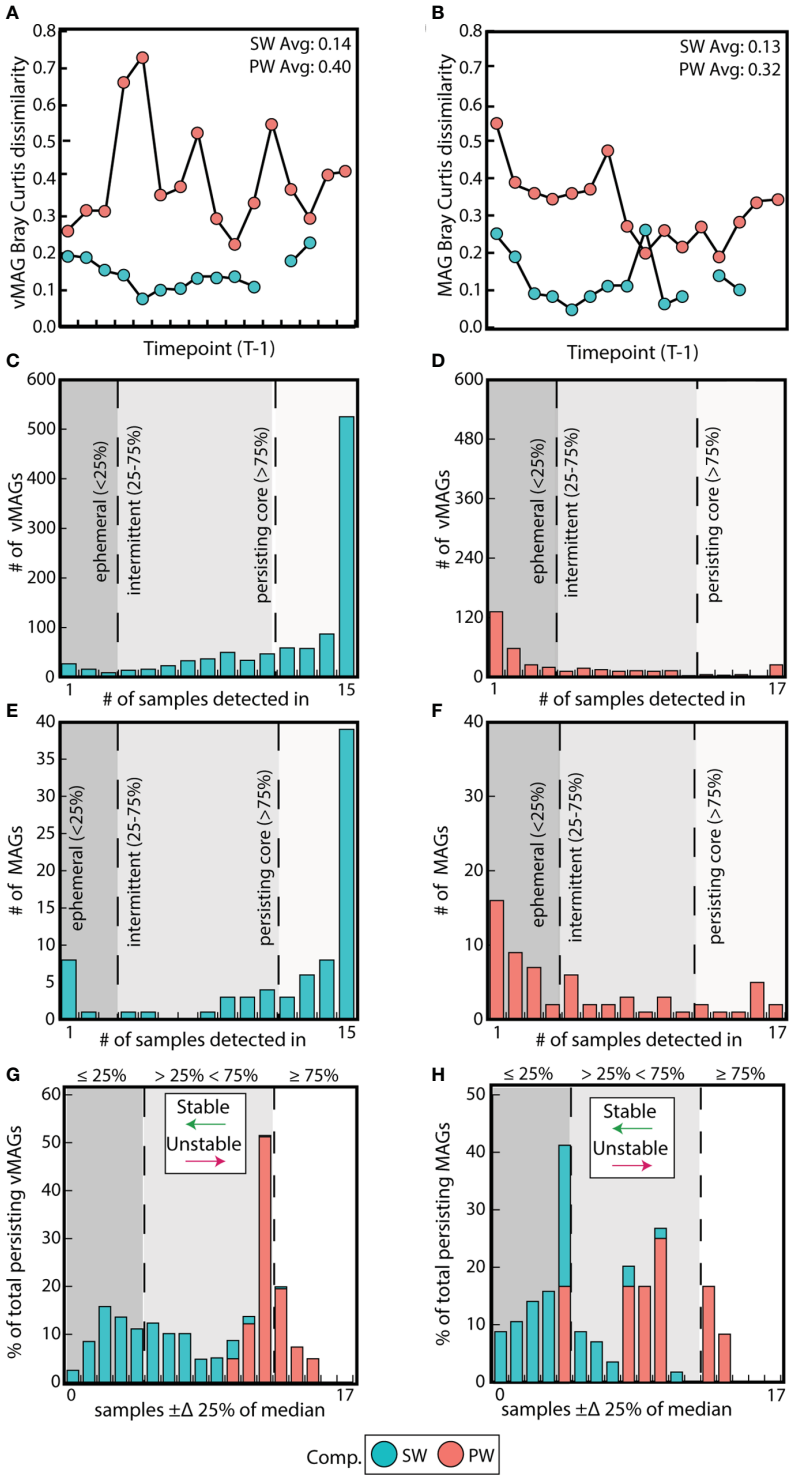


FIGURE 4
Surface water communities are more stable and persistent than pore water communities. **(A)** Difference in Bray-Curtis dissimilarities between each sample and its prior timepoint calculated for vMAGs and **(B)** MAGs per compartment. **(C)** Bar plots show the number of persistent, intermittent, and ephemeral vMAGs in the SW and **(D)** the PW. **(E)** Bar plots show the number of persistent, intermittent, and ephemeral MAGs in the SW and **(F)** the PW. **(G)** Bar plot where the x-axis shows the number of samples where each vMAG that fluctuates above or below 25% of their median values and the y-axis shows the normalized total percentage of persistent genomes per each compartment that are fluctuating. **(H)** Identical bar plots to those in **(G)** but for MAGs.

10.3389/frmbi.2023.1199766

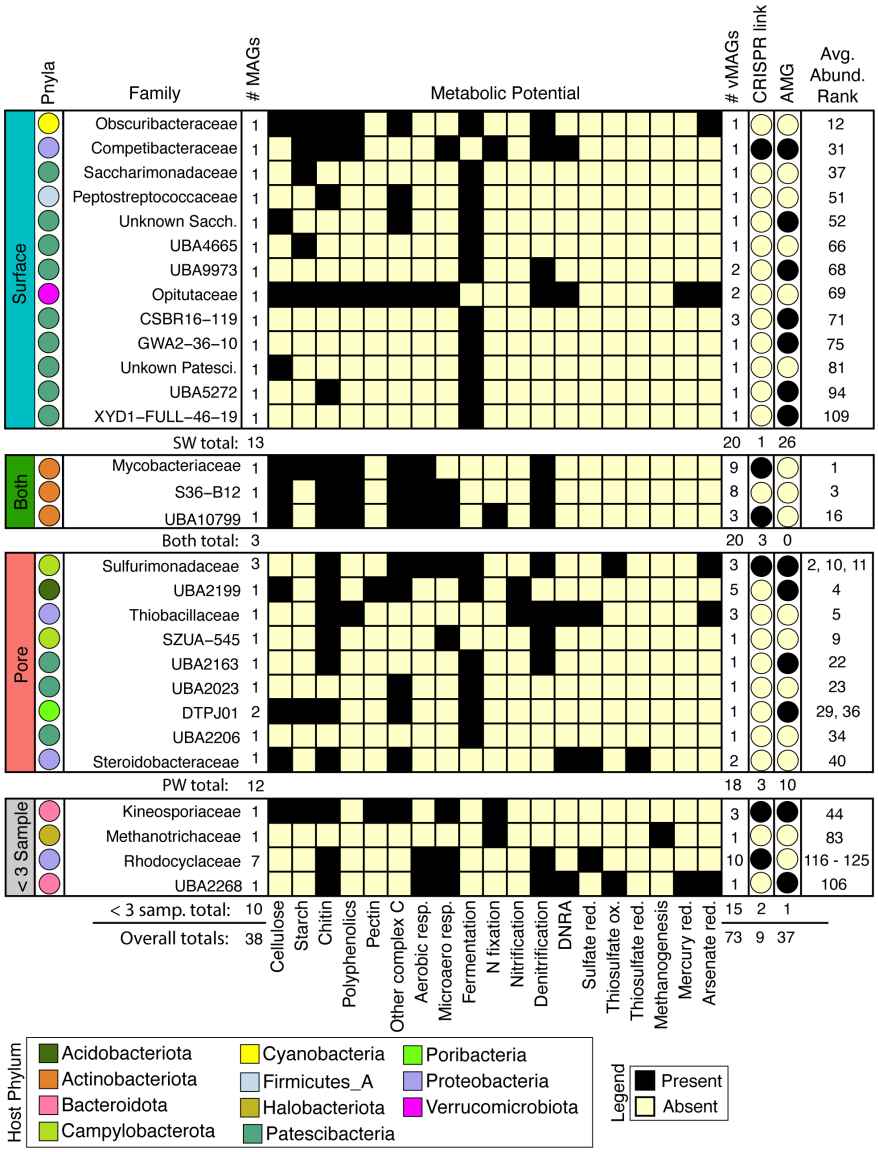


FIGURE 5
Viruses infect abundant microorganisms in rivers which can influence aerobic and anaerobic C, N, and S cycling by predation or auxiliary metabolic genes. MAG families that had a linkage to a virus are shown and split into their compartment-level distributions. From left to right: Colors of each circle on the leftmost side represent the Phyla, and for each family the total number of MAGs are shown. The presence absence heatmap describes the metabolisms of each family. Following the heatmap are the number of vMAGs that are linked in each family, whether the virus-host link is predicted by CRISPR or consensus method, and if at least 1 infecting vMAG with an AMG is reported. Numbers below each bounding box show totals of above criteria. The overall average rank of each MAG within a family is shown in the rightmost column.

predominantly by MAGs in the PW. vMAGs were also predicted to infect members that encoded denitrification pathways which were prevalent in organisms across both compartments (e.g., *Nanopelagicales*). Interestingly, 20% of the vMAGs that infected hosts were cosmopolitan, with representatives identified in other freshwater and wastewater systems (Supplementary Table S2). Together, our genome-resolved database of microbial metabolisms and their putatively infecting viruses gives insight into the underpinnings of River Erpe metabolisms, and show that genome-resolved, river microbiome studies can provide critical perspectives for understanding the impact that viruses can have in river ecosystems.

Virally encoded auxiliary metabolic genes can potentially alter host metabolic machinery in this urban-impacted river

In addition to the impact on microbial communities via predation, viruses can also mediate biogeochemical cycles through enhancing host metabolism with Auxiliary Metabolic Genes (AMGs). We mined our 1,230 vMAGs for putative AMGs and found 165 unique viral AMG candidates after quality filtering, which encompassed 65 unique gene IDs. We failed to see a statistical enrichment for the number of AMGs in either compartment (Fisher's exact $p = 0.77$), suggesting their shared

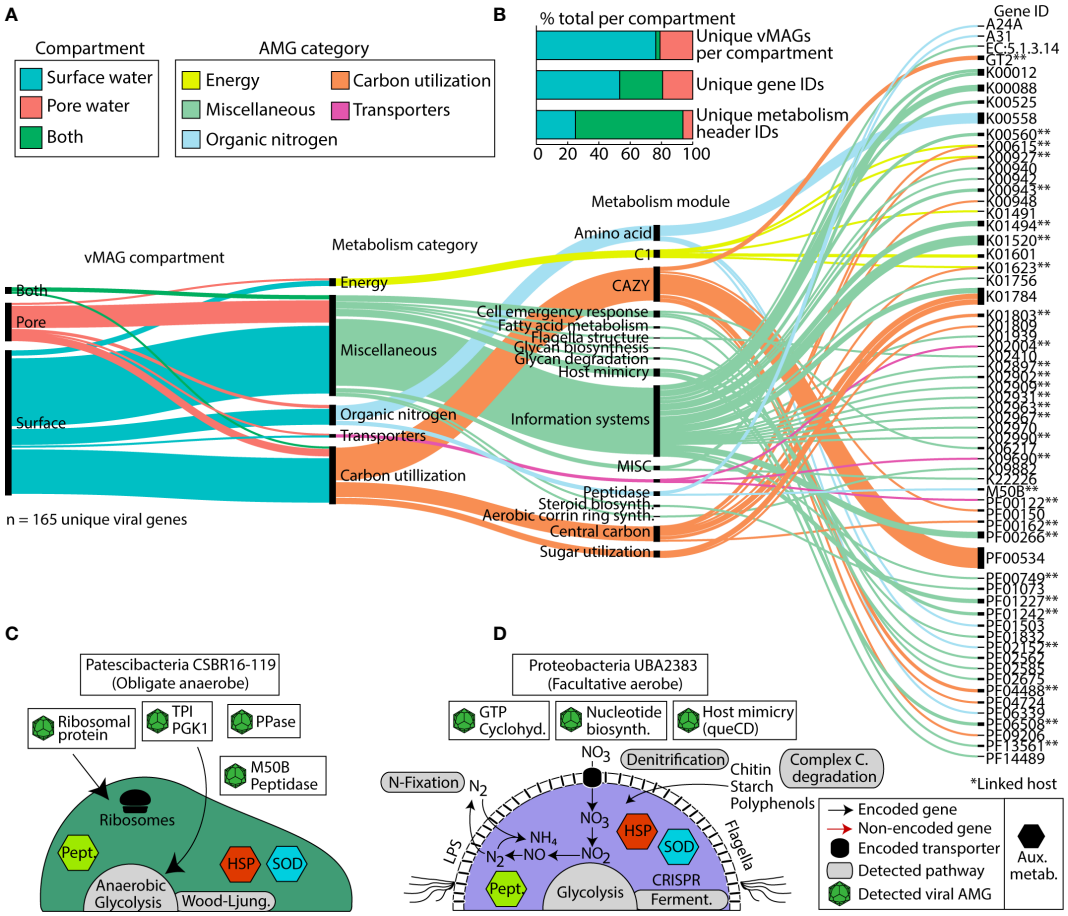


FIGURE 6
Distribution of viral Auxiliary Metabolic Genes (AMGs) and their function reveals key viral interactions that can enhance host metabolism in river ecosystems. **(A)** Alluvial plot shows the subset of AMGs (77%, n=165) that had a metabolic function annotated by DRAM-v and were 1) not at the end of a contig and 2) did not contain a transposon like element. In the first vertical line, colors show the compartments that each vMAG with an AMG was detected in. The second vertical line shows the different DRAM-v metabolic categories for each AMG. The next vertical line shows the specific metabolic module name as categorized by DRAM. The final line contains each of the Gene IDs for the detected AMGs. Genes that can have multiple functions (n = 13) are duplicated and treated as individual genes within each category. **(B)** Stacked bar charts show the proportion of total AMGs encoded in vMAGs from different compartments at the scaffold, gene ID, and metabolism header ID level as shown in **(A)**. **(C, D)** Genome cartoons of two computationally linked bacterial hosts and their respective metabolisms. Detected viral AMGs are shown as viral icons above each genome cartoon. Pept., peptidases; HSP, heat shock proteins; SOD, superoxide dismutase; queCD, 7-cyano-7-deazaguanine synthase; 6-carboxy-5,6,7,8-tetrahydropterin synthase. Asterisks (**) denote AMGs that were encoded within a virus that had a computationally linked host.

importance for the River Erpe. The functionalities of these AMGs at the gene annotation level (e.g., KO number) were mostly conserved across compartments, with only 27% of unique gene IDs present in both compartments. However, at the DRAM-v functional module level (e.g., amino acid metabolism) 69% of metabolisms were present across both our ecological gradients (Figure 6B). Conserved DRAM categories across compartments pertained to carbon utilization (e.g., CAZyme inferred substrates (cellulases), glycolysis), energy generation (e.g., CO_2 fixation (reductive pentose phosphate pathway)), and other reactions (methionine degradation). We note that genes necessary for viral replication like nucleotide biosynthesis, ribosomal proteins, host mimicry, glycan biosynthesis, cofactor and vitamin metabolism, and molecular transporters were conserved between compartments.

There were also some unique AMGs that did show compartment specificity. For example, within the surface water we exclusively detected AMGs for organic nitrogen mineralization and transcriptional regulation (i.e., peptidase M50), sugar metabolism (i.e., fructose and mannose (mannose-6-phosphate isomerase)), and motility (i.e., flagellar motor switch protein *FliG*). We note that among our putative AMGs, we also identified several glycosyltransferases (i.e., GT1, GT2, GT17, and a general sugar binding GT) (Figure 6A). These GT genes are commonly reported as carbohydrate degradation enzymes in other studies, particularly those annotated as glycosyltransferase 2 (GT2) because of the breadth of reactions in their CAZyme families. As such, while we report these in our figure and supplemental information for transparency, we urge caution when inferring these activities in carbon degradation.

We next considered AMGs that either expanded the host metabolism or that were complementary to the host metabolism (i.e., Class I AMGs) (Hurwitz and U'Ren, 2016). Of the 12 *Patescibacteria* MAGs that had possible viral genome links, MAG representative CSBR16-119 had two possible vMAG linkages. A

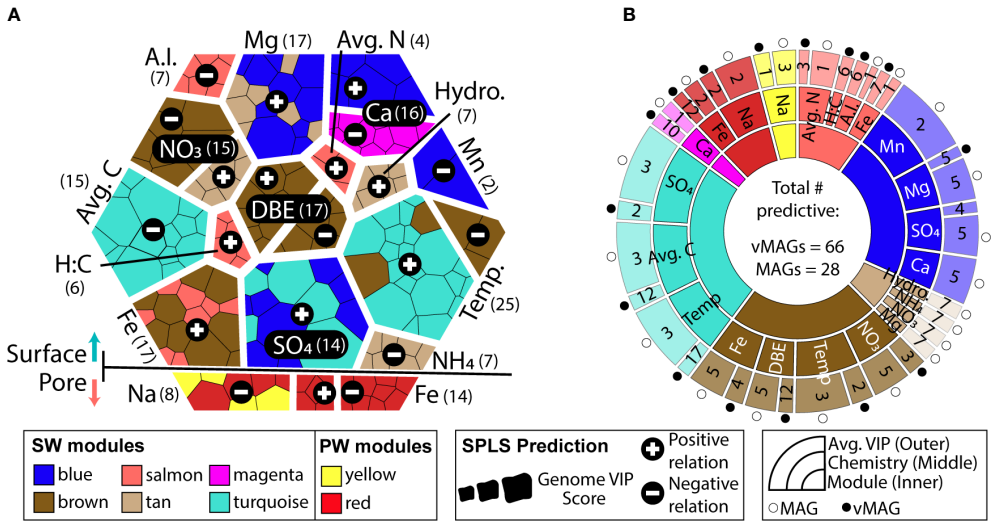


FIGURE 7
WGCNA co-occurrence networks reveal ecologically similar groups that are related to overall ecosystem biogeochemistry. **(A)** Voronoi diagram shows VIP values of predictions for each predictive genome using a hierarchy structure. Each amorphous square within a group represents a single MAG or vMAG. At the first level (i.e., splitting of the large hexagon into upper and lower groups), SW (top) and PW (bottom) predictions are shown. At the second level (i.e., grouping of individual chemical variables predicted across each compartment), individual chemical variables are shown, per each compartment, and how many vMAGs/MAGs were predictive are denoted by numbers next to each variable name. At the third level (i.e., individual amorphous square or genomes), shapes are sized by the VIP score (>1) of genomes that predict that variable and are colored by their respective WGCNA module. **(B)** Sunburst diagram shows the predictive WGCNA modules in the innermost level, followed by what chemical values each module predicts in the middle level. The outer level shows the average variable importance in projection (VIP) score for each genome type: vMAG (black circles) and MAGs (white circles) for that chemical prediction.

comparison of the metabolic capabilities of the host and viral genomes indicated multiple shared genes (Figure 6C, Supplemental Table 2). For example, a peptidase-like protein (M50) that is inferred transcriptional regulator (Rawlings et al., 2018) was present in both the *Patescibacteria* MAG and its infecting vMAG. Across the length of the open reading frame, these bacterial and viral genes shared 77% and 99% nucleotide and amino acid similarity, respectively (Supplemental Table 2). The microbial host genome also had a single copy of ribosome L28 encoded, and two viral genomes putatively infecting this host contained a relevant homolog to L28 (>93% identity, over 90% query coverage) (Supplemental Table 2).

A second putatively infected genome was *Proteobacteria* UBA2383 (a novel unclassified *Competibacteraceae*) which had broad metabolic capabilities and was persistent in our samples (Figure 6D). This MAG was inferred to be a facultative aerobe encoding genes for aerobic respiration and for denitrification. UBA2383 encoded genes supporting a heterotrophic lifestyle including CAZymes necessary for the degradation of complex carbon substrates (e.g., chitin, starch, and polyphenol) and the enzymatic capacity to utilize these substrates for energy (e.g., glycolysis, tricarboxylic acid cycle). This MAG also encoded the ability to fix nitrogen and denitrify. The two vMAGs that were associated with this genome encoded genes to support host metabolism (e.g., GTP cyclohydrolase) which generates important co-factors for bacterial metabolic processes (Supplemental Table 2) (He and Rosazza, 2003). Additional AMGs encoded by infecting viruses could potentially enhance nucleotide biosynthesis (dCTP deaminase, dUTP pyrophosphatase, thymidylate synthase) as well as other viral functions like host mimicry genes (i.e., 7-cyano-7-deazaguanine synthase, 6-carboxy-5,6,7,8-tetrahydropterin synthase) to avoid the CRISPR defense mechanisms encoded within the host *Proteobacteria*. More research using non-homology based methods, as well as expression patterns of these AMGs would help confirm their functionality and activity in this urban-impacted stream.

Co-occurrence networks elucidate ecological patterns that inform ecosystem biogeochemistry

To link viral and microbial communities more concretely to ecosystem biogeochemical cycling, we leveraged our collected temporal samples and applied a Weighted Gene Correlation Network Analysis (WGCNA) to identify organismal groups that co-occurred over the 48-hour sampling time. Highlighting the clear distinctions in SW and PW compartments, WGCNA analyses could not be reasonably performed simultaneously on a combined dataset (scale free topology model fit max = 0.32 at power = 20). As such, using only microbial and viral genomic abundances from either SW or PW separately, we identified 15 and 4 co-occurring modules in the SW and PW, respectively (Supplemental Figure 4). The largest module in both networks (turquoise module) contained 254 genomes in the SW and 71 in the PW. In the SW compartment, the overall modules had an average richness of 66 vMAGs and 5

MAGs, while in the PW they had an average richness of 46 vMAGs and 10 MAGs.

Both surface and pore water communities had modules of co-occurring genomes that were significantly related by sparse partial least square regressions (sPLS) to the collected biogeochemical measurements ($R^2 > 0.3$, $p < 0.05$) (Figure 7A, Supplemental Figure 4). Only total Fe concentrations were related to modules in both the SW (brown, salmon modules) and PW (red module). SW modules were uniquely related to variables pertinent to nitrogen (nitrate, average total nitrogen), carbon (average total carbon, aromaticity index, hydrogen:carbon), as well as physical (temperature, water stage) and geochemical (magnesium, calcium, manganese, ammonium, sulfate) features in these samples. Of the 8 modules that were significantly related to ecosystem hydrobiogeochemical features, viruses had significant variable importance in projection scores (VIP > 1) in 7 of them, and 70% of the most significantly related genomes across all regressions were viral (Figure 7B).

Of the 73 vMAGs and 38 MAGs that were computationally linked (Figures 5, 6), nearly a quarter of those vMAGs and a third of MAGs were grouped into the same co-occurring modules. Interestingly, the SW brown module was related to the total nitrate concentrations in our dataset and contained a co-occurring virus-host link (Figure 8A). The host genome was the *Competibacteraceae* genome in Figure 6D and its putatively infecting a virus, which together could play roles in modulating the nitrogen cycling through both fixation and denitrification. This virus and microbial host pair had significant negative correlations to nitrate concentrations and were the second and fourth most significantly related genomes to nitrate within the brown module. The virus bacterial ratio (VBR) for these two organisms was nearly 1:1 and significantly correlated, which is expected of kill the winner dynamics (Trubl et al., 2021), and ultimately highlighting the possible dependency of an infecting vMAG and its host (Figure 8B). In support of this relationship, the viral genome coverages were on average 10x more than the putative host MAG coverage, suggesting a possible lytic infection lifestyle. Further underlining the importance of these related genomes, both were designated as persistent (i.e., present in >75% of all collected timepoints) and were the 1st (vMAG) and 9th (MAG) most abundant genomes detected in the surface waters.

Discussion

Viral reference databases underrepresent certain habitats, missing cosmopolitan, ecologically relevant lineages

Nearly a quarter of our Erpe viruses formed genus-level clusters with viruses from wastewater and freshwater systems, and of those, 11% encoded a putative AMG with functions for metabolisms such as carbon utilization, organic nitrogen transformations, and housekeeping functions (i.e., transporters and flagellar assembly). While the protein clustering of River Erpe vMAGs to wastewater viruses was not entirely surprising given the sampling location was

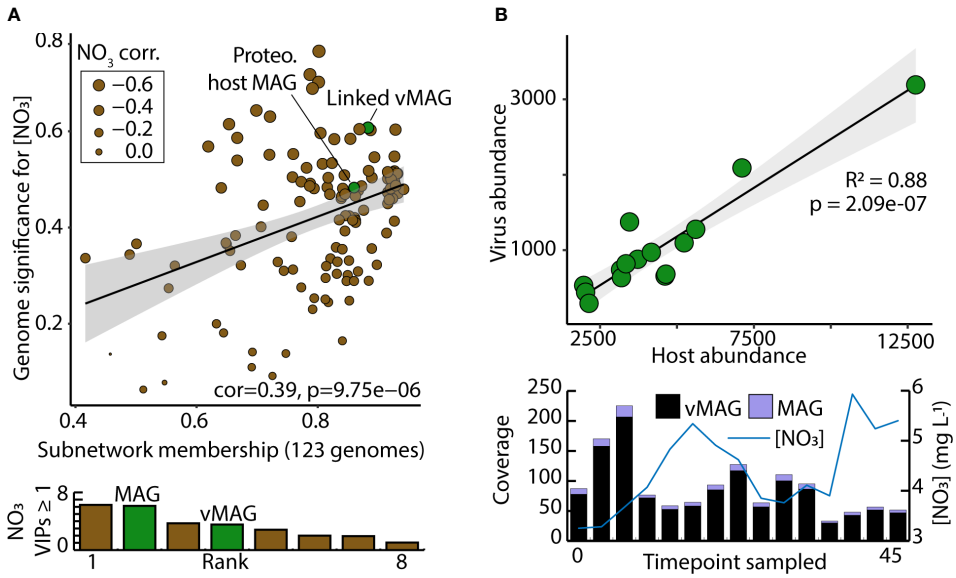


FIGURE 8
Computationally linked vMAG and MAG pair that share co-occurrence patterns demonstrate high significance for nitrate, and display kill-the-winner dynamics. **(A)** Scatterplot depicts the genomic significance for nitrate of each of the genomes in the brown module in relation to the membership of those genomes within the WGCNA network modules. Below, bar charts show the VIP score (≥1) of the different organisms in the brown module. **(B)** A Virus bacteria ratio (VBR) plot of a viral genome within the brown module that was predicted to infect a Proteobacteria genome. Below it, bar plots show the total coverage across all samples for both the vMAG and the MAG, and a line graph shows the measured nitrate concentrations that these genomes predict.

downstream from the wastewater outlet (Mueller et al., 2021), we note that we also clustered a similar proportion of viruses to other viral genomes from river systems. Notably, this similar clustering proportion for River Erpe viruses was not observed with the TARA ocean viruses (Figure 2C, Supplemental Figure 3). These results hint at possible ecosystem filtering that may affect the biogeographical patterns of freshwater viruses. Our results also underscore the importance of customized, ecosystem relevant databases in environmental viromics for extending the ecological relevance of these ecosystem modulators, and further understanding the major drivers for river microbiomes.

Temporally and spatially resolved metagenomics coupled to metabolites and geochemistry enhances our understanding of river microbiome structure

Sampling with a Eulerian method allowed us to detect microbiomes passing through the same space over time in the SW and PW samples. Due to the flow rate of SW, and the potential that PW communities may be more biofilm impacted, we might have expected to see greater microbial and viral changes in the surface compartment than the sediments over the sampled time period. On the contrary, both vMAGs and MAGs were more persistent and had more stable abundance patterns over time in the SW of the River Erpe (Figures 4A, B). A possible explanation is that the strong influence of the wastewater treatment plant, where inputs were relatively uniform and continuous over time (Mueller et al., 2021), could contribute to the increased temporal stability we observed. It is also possible that the mixing in the PW hyporheic zone was more frequent than the flow rate within this channel. In support of the former, we did observe strong clustering between our viral genomes and wastewater treatment viral genomes throughout the timeseries (Figure 2). Our study is consistent with previous research showing surface water microbiomes are not unstable, or intractable (Graham et al., 2017), and could thus be important for the poorly resolved indices of river health and biogeochemistry that currently exist.

Previous reports using non genome-resolved strategies highlight that richness in river PW and sediments are generally higher than those in the SW for bacterial communities (Abia et al., 2018). Contrary to this, our data shows the opposite trends in the Erpe river for both viral and microbial communities (Supplementary Figure 1). One possible explanation could be methodological due to the PW being sampled or assembled less completely as a result of genomic extraction bias caused by fine grain sediments, less sampling volume, or strain level complexity. However, our species area curves did not signify an obvious difference in sampling exhaustion between these compartments (Supplemental Figure 1), leaving open the possibility that this finding may be biological.

A possible biological explanation could be that the effluent of the Münchehofe WWTP is altering viral and microbial community diversity. Our geochemical data showed elevated total nitrogen and soluble reactive phosphorous concentrations, which are commonly reported for WWTP impacted systems (Fox et al., 1989; Effler et al., 2010). However, the role of these WWTP influences on river microbiome diversity are variable, with some studies reporting that WWTPs reduce bacterial diversity and overall nutrient concentrations (Atashgahi et al., 2015; Carles et al., 2022; Xie et al., 2022), and other studies showing increases in diversity resulting from eutrophication (Garnier et al., 1992; Marti and Balcázar, 2014).

Our presented and previously published geochemical data inferred more anoxic conditions in the porewater compared to the surface water (Mueller et al., 2021), while the FTICR-MS data indicated a higher concentration of non-labile, microbially inaccessible carbon in the pore water. Additionally, sediment profiles for our samples ranged from 85.8%-96.6% clay content (Mueller et al., 2021), which may impact groundwater and surface water exchanges, resulting in altered nutrient fluxes to inhibit microbial growth (Newcomer et al., 2016; Huettel et al., 1998). Taken together, it is possible that limited nutrient and carbon accessibility contributes to the decreased microbial and viral diversity observed in the PW compared to the SW.

Viruses have the potential to regulate river biogeochemical cycles by predation and metabolic reprograming of microbial hosts

Although river viral ecology is only recently becoming appreciated, early works suggested that viruses likely play key roles in the structuring of river microbial communities (Peduzzi and Luef, 2009; Peduzzi, 2016). By using a combination of computational methods, we show viruses infect microorganisms that encode a wide array of metabolic functionality critical to river biogeochemistry (e.g., methanogens, denitrifiers, oxygen respirers). In addition to predation, viral auxiliary metabolic genes are recognized across aquatic systems to play key roles in host metabolic reprogramming and can encompass a wide range of processes from photosynthesis to the oxidation of sulfur (Sullivan et al., 2006; Anantharaman et al., 2014). We add to the existing literature and show AMGs in urban river systems may also impact reactions involving nitrogen, carbon, and sulfur cycling. Additionally, one of the vMAGs that was predicted to infect a *Patescibacteria* genome encoded a ribosomal protein, a finding that has been previously reported in other systems for different bacteria (Mizuno et al., 2019). Candidate phyla radiation (CPR) organisms like *Patescibacteria* are present in wastewater treatment plants (Wang et al., 2023) and contain non-redundant, small genomes (Tian et al., 2020; Wang et al., 2023). As such, our results of ribosomal AMGs in *Patescibacteria*-infecting viruses hint at the possibility that viruses may help maintain those small genome sizes by encoding necessary host genes, a concept previously demonstrated for the virus-host dependency of cyanobacterial photosynthesis in oceans (Sullivan et al., 2006).

Other works looking at vMAGs from freshwater lakes and estuaries have shown that some viruses exhibit endemism for

certain environments, meaning their distribution is limited to a small geographic area (Ruiz-Perez et al., 2019). This points to an interesting idea that perhaps AMGs may also be tuned to the specific ecological functions of the sampled habitat, and as such that we could expect some degree of endemism in the AMGs. A recent study from an estuary identified significant partitioning of AMG functions between habitat types (water particle and sediment) (Luo et al., 2022). In support of these findings, we identified a subset of unique AMGs within the SW (e.g., flagellar assembly proteins, sugar metabolism) that could potentially be more associated with a lifestyle supported by favorable carbon, and aquatic environments that favor mobility. On the other hand, in the PW we detected AMGs that encoded for plant hemicellulose degradation (Supplementary Table S2), an adaptation that could sustain metabolism in a litter impacted, sediment habitat. However, most AMG functional categories from our dataset were highly similar across compartments suggesting some conservation within River Erpe compartments. As such, it is possible that due to the constant mixing of surface and HZ water in this river, and possibly others, stratification at the genomic potential may be less notable, and expression information may be necessary to capture habitat specific differences. Ultimately, this study highlights how moving forward annotation resolution and expanding reference database(s) are important factors to consider when extrapolating AMG inferences across datasets (Hurwitz and U'Ren, 2016; Shaffer et al., 2020).

In conclusion, our results highlight the power of temporally resolved metagenomics in understanding river microbiome dynamics. Leveraging the community-sequenced dataset of the River Erpe, we provide insights regarding compartment-level microbiome stability and show surface water microbiomes may not be as "untraceable" or "unstable" as previously thought. This stability at a genome-resolved view, suggests microbial content could add to the growing body of indicators for river wellness. Ultimately, this research provides a strong scaffolding foundation for future temporally resolved river studies that couple microbial omics measurements to biogeochemical rates to bridge the gap in understand overall ecosystem functionality.

Data availability statement

The datasets supporting the conclusions of this article are publicly available and collected as part of the Worldwide Hydrobiogeochemistry Observation Network for Dynamic River Systems (WHONDRS) collective sequencing project and are all publicly available on ESS-Dive (Wells et al., 2019). The 125 MAGs are deposited in NCBI BioProject, accession number PRJNA946291, and the 1230 vMAGs have been deposited in NCBI BioProject, under ID SAMN34000891. The 125 MAGs, the 1230 vMAGs, the raw annotations for each genome, and the dataset of freshwater and wastewater viruses we used to cluster to the HUM-V viruses are hosted on Zenodo (https://doi.org/10.5281/zenodo.7709817). All scripts, commands, and input data used for this manuscript are available at https://github.com/jrr-microbio/erpe_river.

Author contributions

JR-R: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing - Original Draft, Writing - Review and Editing, Visualization. AO: Conceptualization, Methodology, Writing - Original Draft, Formal analysis, Writing - Review and Editing, MB: Methodology, Software, Writing - Review and Editing. RD: Methodology, Investigation, Writing - Review and Editing. BM: Investigation, Writing - Review and Editing. HS: Investigation, Writing - Review and Editing. JE: Formal analysis, Writing - Review and Editing. RF: Software, Writing - Review and Editing. RAD: Formal analysis, Writing - Review and Editing. LS: Writing - Review and Editing, MS: Software. AG: Conceptualization, Methodology, Supervision, Writing - Review and Editing. JL: Conceptualization, Methodology, Supervision, Writing - Review and Editing. JS: Conceptualization, Methodology, Supervision, Funding acquisition, Writing - Review and Editing. KW: Conceptualization, Methodology, Resources, Writing - Original Draft, Writing - Review and Editing, Supervision, Project administration, Funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

JR-R, AO, MB, RF, RAD, JE, LS, and MS were fully or partially supported by awards to KW, including those from DOE Office of Science, Office of Biological and Environmental Research (BER), grant nos. DE-SC0021350 and DE-SC0023084, as well as the National Science Foundation, grant no. 2149506. A portion of this work was performed by MB, RD, AG, and JS at Pacific Northwest National Laboratory (PNNL) and funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, and Environmental System Science (ESS) Program. This contribution originates from the River Corridor Scientific Focus Area (SFA) project at Pacific Northwest National Laboratory (PNNL). PNNL is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract No. DE-AC05-76RL01830. A subcontract to KW from the River Corridor SFA also supported a portion of this work. Metagenomic sequencing was performed by the University of Colorado Anschutz's Genomics Shared Resource supported by the Cancer Center Support Grant (P30CA046934).

Acknowledgments

Data used in this manuscript were collected as a part of the WHONDRS 2018 sampling campaign and we thank those that participated in the design and implementation of that effort. Samples were sequenced and processed as a part of the Genome Resolved Open Watersheds (GROW) effort to sequence world rivers. We also thank Tyson Claffey and Richard Wolfe for Colorado State University server management.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1199766/full#supplementary-material

References

Abia, A. L. K., Alisoltani, A., Keshri, J., and Ubomba-Jaswa, E. (2018). Metagenomic analysis of the bacterial communities and their functional profiles in water and sediments of the Apies River, South Africa, as a function of land use. *Sci. Total Environ.* 616–617, 326–334. doi: 10.1016/j.scitotenv.2017.10.322

Ahlgren, N. A., Ren, J., Lu, Y. Y., Fuhrman, J. A., and Sun, F. (2017). Alignment-free d_2^* oligonucleotide frequency dissimilarity measure improves prediction of hosts from metagenomically-derived viral sequences. *Nucleic Acids Res.* 45, 39–53. doi: 10.1093/nar/gkw1002

Allen, G. H., and Pavelsky, T. M. (2018). Global extent of rivers and streams. *Science* 361, 585–588. doi: 10.1126/science.aat0636

Alonso, C., Román, A., Bejarano, M. D., Garcia de Jalon, D., and Carolli, M. (2017). A graphical approach to characterize sub-daily flow regimes and evaluate its alterations due to hydropeaking. *Sci. Total Environ.* 574, 532–543. doi: 10.1016/j.scitotenv.2016.09.087

Anantharaman, K., Duhaime, M. B., Breier, J. A., Wendt, K. A., Toner, B. M., and Dick, G. J. (2014). Sulfur oxidation genes in diverse deep-sea viruses. *Science* 344, 757–760. doi: 10.1126/science.1252229

Andrews, S. *FastQC: A quality control analysis tool for high throughput sequencing data*. Available at: https://github.com/s-andrews/FastQC (Accessed October 19, 2022).

Atashgahi, S., Aydin, R., Dimitrov, M. R., Sipkema, D., Hamonts, K., Lahti, L., et al. (2015). Impact of a wastewater treatment plant on microbial community composition and function in a hyporheic zone of a Eutrophic river. *Sci. Rep.* 5, 17284. doi: 10.1038/srep17284

Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on Earth. *Proc. Natl. Acad. Sci. U. S. A.* 115, 6506–6511. doi: 10.1073/pnas.1711842115

Berg, M., Goudeau, D., Olmsted, C., McMahon, K. D., Yitbarek, S., Thweatt, J. L., et al. (2021). Host population diversity as a driver of viral infection cycle in wild populations of green sulfur bacteria with long standing virus-host interactions. *ISME J.* 15, 1569–1584. doi: 10.1038/s41396-020-00870-1

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J. H., Roux, S., Adriaenssens, E. M., et al. (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.* 37, 632–639. doi: 10.1038/s41587-019-0100-8

Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35, 725–731. doi: 10.1038/nbt.3893

Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doulcier, G., Acinas, S. G., Alberti, A., et al. (2015). Ocean plankton. Patterns and ecological drivers of ocean viral communities. *Science* 348, 1261498. doi: 10.1126/science.1261498

Bushnell, B. (2014) *BBMap: A Fast, Accurate, Splice-Aware Aligner* (Berkeley, CA (United States: Lawrence Berkeley National Lab. (LBNL). Available at: https://www.osti.gov/servlets/purl/1241166 (Accessed March 23, 2022).

Carles, L., Wullschleger, S., Joss, A., Eggen, R. I. L., Schirmer, K., Schuwirth, N., et al. (2022). Wastewater microorganisms impact microbial diversity and important ecological functions of stream periphyton. *Water Res.* 225, 119119. doi: 10.1016/j.watres.2022.119119

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. doi: 10.1093/bioinformatics/btz848

Chow, C.-E. T., and Fuhrman, J. A. (2012). Seasonality and monthly dynamics of marine myovirus communities. *Environ. Microbiol.* 14, 2171–2183. doi: 10.1111/j.1462-2920.2012.02744.x

Chow, C.-E. T., and Suttle, C. A. (2015). Biogeography of viruses in the sea. *Annu. Rev. Virol.* 2, 41–66. doi: 10.1146/annurev-virology-031413-085540

Chu, H., Gao, G.-F., Ma, Y., Fan, K., and Delgado-Baquerizo, M. (2020). Soil microbial biogeography in a changing world: recent advances and future perspectives. *mSystems* 5. doi: 10.1128/mSystems.00803-19

Chung, D., Chun, H., and Keles, S. (2012). Spls: sparse partial least squares (SPLS) regression and classification. *R package version* 2, 1–1. doi: 10.1111/j.1467-9868.2009.00723.x

Cissoko, M., Desnues, A., Bouvy, M., Sime-Ngando, T., Verling, E., and Bettarel, Y. (2008). Effects of freshwater and seawater mixing on virio- and bacterioplankton in a tropical estuary. *Freshw. Biol.* 53, 1154–1162. doi: 10.1111/j.1365-2427.2007.01930.x

Coenen, A. R., Hu, S. K., Luo, E., Muratore, D., and Weitz, J. S. (2020). A primer for microbiome time-series analysis. *Front. Genet.* 11, 310. doi: 10.3389/fgene.2020.00310

Corinaldesi, C., Dell'Anno, A., Magagnini, M., and Danovaro, R. (2010). Viral decay and viral production rates in continental-shelf and deep-sea sediments of the Mediterranean Sea. *FEMS Microbiol. Ecol.* 72, 208–218. doi: 10.1111/j.1574-6941.2010.00840.x

Effler, S. W., O'Donnel, S. M., Prestigiacomo, A. R., O'Donnell, D. M., Gelda, R. K., and Matthews, D. A. (2010). The effect of municipal wastewater effluent on nitrogen levels in Onondaga Lake, a 36-year record. *Water Environ. Res.* 82, 3–19. doi: 10.2175/106143009X407384

Elbehery, A. H. A., and Deng, L. (2022). Insights into the global freshwater virome. *Front. Microbiol.* 13, 953500. doi: 10.3389/fmicb.2022.953500

Erbilgin, O., Bowen, B. P., Kosina, S. M., Jenkins, S., Lau, R. K., and Northen, T. R. (2017). Dynamic substrate preferences predict metabolic properties of a simple microbial consortium. *BMC Bioinf.* 18, 57. doi: 10.1186/s12859-017-1478-2

Fox, I., Malati, M. A., and Perry, R. (1989). The adsorption and release of phosphate from sediments of a river receiving sewage effluent. *Water Res.* 23, 725–732. doi: 10.1016/0043-1354(89)90206-6

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., et al. (2022). Global carbon budget 2021. *Earth Syst. Sci. Data* 14, 1917–2005. doi: 10.5194/essd-14-1917-2022

Garnier, J., Servais, P., and Billen, G. (1992). Bacterioplankton in the Seine River (France): impact of the Parisian urban effluent. *Can. J. Microbiol.* 38, 56–64. doi: 10.1139/m92-009

Gibson, B., Wilson, D. J., Feil, E., and Eyre-Walker, A. (2018). The distribution of bacterial doubling times in the wild. *Proc. Biol. Sci.* 285. doi: 10.1098/rspb.2018.0789

Graham, E. B., Crump, A. R., Resch, C. T., Fansler, S., Arntzen, E., Kennedy, D. W., et al. (2017). Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. *Environ. Microbiol.* 19, 1552–1567. doi: 10.1111/1462-2920.13720

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et al. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature* 532, 465–470. doi: 10.1038/nature16942

Guo, J., Bolduc, B., Zayed, A. A., Varsani, A., Dominguez-Huerta, G., Delmont, T. O., et al. (2021a). VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* 9, 37. doi: 10.1186/s40168-020-00990-y

Guo, J., Vik, D., Pratama, A. A., Roux, S., and Sullivan, M. (2021b) *Viral sequence identification SOP with VirSorter2. protocols.io*. Available at: https://www.protocols.io/view/viral-sequence-identification-sop-with-virsorter2-5qpvoyqebg4o/v3 (Accessed July 5, 2022).

He, A., and Rosazza, J. P. N. (2003). GTP cyclohydrolase I: purification, characterization, and effects of inhibition on nitric oxide synthase in nocardia species. *Appl. Environ. Microbiol.* 69, 7507–7513. doi: 10.1128/AEM.69.12.7507-7513.2003

Hendrix, R. W., Smith, M. C. M., Neil Burns, R., Ford, M. E., and Hatfull, G. F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2192–2197. doi: 10.1073/pnas.96.5.2192

Hewson, I., O'Neil, J. M., Fuhrman, J. A., and Dennison, W. C. (2001). Virus-like particle distribution and abundance in sediments and overlying waters along eutrophication gradients in two subtropical estuaries. *Limnol. Oceanogr.* 46, 1734–1746. doi: 10.4319/lo.2001.46.7.1734

Holmfeldt, K., Nilsson, E., Simone, D., Lopez-Fernandez, M., Wu, X., de Bruijn, I., et al. (2021). The Fennoscandian Shield deep terrestrial virosphere suggests slow motion "boom and burst". *cycles. Commun. Biol.* 4, 307. doi: 10.1038/s42003-021-01810-1

Hou, Z., Nelson, W. C., Stegen, J. C., Murray, C. J., Arntzen, E., Crump, A. R., et al. (2017). Geochemical and microbial community attributes in relation to hyporheic zone geological facies. *Sci. Rep.* 7, 12006. doi: 10.1038/s41598-017-12275-w

Huettel, M., and Ziebis Forster Luther., W. S.G.W. (1998). Advective transport affecting metal and nutrient distributions and interfacial fluxes in permeable sediments. *Geochimica et Cosmochimica Acta* 62 (4), 613–31. doi: 10.1038/s41598-017-12275-w

Hurwitz, B. L., and U'Ren, J. M. (2016). Viral metabolic reprogramming in marine ecosystems. *Curr. Opin. Microbiol.* 31, 161–168. doi: 10.1016/j.mib.2016.04.002

Joshi NA, F. J. N. (2011) *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files*. Available at: https://github.com/najoshi/sickle (Accessed October 29, 2020).

Kaevska, M., Videnska, P., Sedlar, K., and Slana, I. (2016). Seasonal changes in microbial community composition in river water studied using 454-pyrosequencing. *Springerplus* 5, 409. doi: 10.1186/s40064-016-2043-6

Kang, D., Li, F., Kirton, E. S., Thomas, A., Egan, R. S., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7, e7359. doi: 10.7717/peerj.7359

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858. doi: 10.1038/nprot.2015.053

Kim, S., Kramer, R. W., and Hatcher, P. G. (2003). Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Anal. Chem.* 75, 5336–5344. doi: 10.1021/ac034415p

Lammers, W. T. (1992). "Stimulation of Bacterial Cytokinesis by Bacteriophage Predation," *Hydrobiologia* 235(1), 261–65.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. doi: 10.1186/1471-2105-9-559

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lewandowski, J., Putschew, A., Schwesig, D., Neumann, C., and Radke, M. (2011). Fate of organic micropollutants in the hyporheic zone of a eutrophic lowland stream: results of a preliminary field study. *Sci. Total Environ.* 409, 1824–1835. doi: 10.1016/j.scitotenv.2011.01.028

Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033

Liu, S., Kuhn, C., Amatulli, G., Aho, K., Butman, D. E., Allen, G. H., et al. (2022). The importance of hydrology in routing terrestrial carbon to the atmosphere via global streams and rivers. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2106322119. doi: 10.1073/pnas.2106322119

Lu, Q., Mao, J., Xia, H., Song, S., Chen, W., and Zhao, D. (2022). Effect of wastewater treatment plant discharge on the bacterial community in a receiving river. *Ecotoxicol. Environ. Saf.* 239, 113641. doi: 10.1016/j.ecoenv.2022.113641

Lundquist, J. D., and Cayan, D. R. (2002). Seasonal and spatial patterns in diurnal cycles in streamflow in the western United States. *J. Hydrometeorol.* 3, 591–603. doi: 10.1175/1525-7541(2002)003<0591:SASPID>2.0.CO;2

Luo, X.-Q., Wang, P., Li, J.-L., Ahmad, M., Duan, L., Yin, L.-Z., et al. (2022). Viral community-wide auxiliary metabolic genes differ by lifestyles, habitats, and hosts. *Microbiome* 10, 190. doi: 10.1186/s40168-022-01384-y

Luo, X., Xiang, X., Huang, G., Song, X., Wang, P., Yang, Y., et al. (2020). Bacterial community structure upstream and downstream of cascade dams along the Lancang River in southwestern China. *Environ. Sci. pollut. Res. Int.* 27, 42933–42947. doi: 10.1007/s11356-020-10159-7

Malki, K., Sawaya, N. A., Tisza, M. J., Coutinho, F. H., Rosario, K., Székely, A. J., et al. (2021). Spatial and temporal dynamics of prokaryotic and viral community assemblages in a lotic system (Manatee springs, Florida). *Appl. Environ. Microbiol.* 87, e0064621. doi: 10.1128/AEM.00646-21

Marti, E., and Balcázar, J. L. (2014). Use of pyrosequencing to explore the benthic bacterial community structure in a river impacted by wastewater treatment plant discharges. *Res. Microbiol.* 165, 468–471. doi: 10.1016/j.resmic.2014.04.002

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., and Azzi, M. (2017). "RAWGraphs: A visualisation platform to create open outputs," in *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter CHItaly '17* (New York, NY, USA: Association for Computing Machinery), 1–5.

Metagenome Assembly Workflow (v1.0.1) and NMDC Workflows 0.2a documentation. Available at: https://nmdc-workflow-documentation.readthedocs.io/en/latest/chapters/3_MetaGAssemly_index.html (Accessed October 19, 2022).

Mizuno, C. M., Guyomar, C., Roux, S., Lavigne, R., Rodriguez-Valera, F., Sullivan, M. B., et al. (2019). Numerous cultivated and uncultivated viruses encode ribosomal proteins. *Nat. Commun.* 10, 752. doi: 10.1038/s41467-019-08672-6

Mueller, B. M., Schulz, H., Danczak, R. E., Putschew, A., and Lewandowski, J. (2021). Simultaneous attenuation of trace organics and change in organic matter composition in the hyporheic zone of urban streams. *Sci. Rep.* 11, 4179. doi: 10.1038/s41598-021-83750-8

Munn, C. B. (2006). Viruses as pathogens of marine organisms—from bacteria to whales. *J. Mar. Biol. Assoc. U. K.* 86, 453–467. doi: 10.1017/S002531540601335X

Mushegian, A. R. (2020). Are there 1031 virus particles on earth, or more, or fewer? *J. Bacteriol.* 202. doi: 10.1128/JB.00052-20

Naegeli, M. W., and Uehlinger, U. (1997). Contribution of the hyporheic zone to ecosystem metabolism in a prealpine gravel-bed-river. *J. North Am. Benthol. Soc* 16, 794–804. doi: 10.2307/1468172

Nelson, A. R., Sawyer, A. H., Gabor, R. S., Saup, C. M., Bryant, S. R., Harris, K. D., et al. (2019). Heterogeneity in hyporheic flow, pore water chemistry, and microbial community composition in an alpine streambed. *J. Geophys. Res. Biogeosci.* 124, 3465–3478. doi: 10.1029/2019JG005226

Newcomer, M. E., Hubbard, S. S., Fleckenstein, J. H., Maier, U., Schmidt, C., Thullner, M., et al. (2016). Simulating bioclogging effects on dynamic riverbed permeability and infiltration. *Water Resour. Res.* 52, 2883–2900. doi: 10.1002/2015WR018351

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2016). *vegan: Community Ecology Package. R package version 2.4-3* (Vienna: R Foundation for Statistical Computing).

Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 11, 2864–2868. doi: 10.1038/ismej.2017.126

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

Peduzzi, P. (2016). Virus ecology of fluvial systems: a blank spot on the map? *Biol. Rev. Camb. Philos. Soc* 91, 937–949. doi: 10.1111/brv.12202

Peduzzi, P., and Luef, B. (2008). Viruses, bacteria and suspended particles in a backwater and main channel site of the Danube (Austria). *Aquat. Sci.* 70, 186–194. doi: 10.1007/s00027-008-8068-3

Peduzzi, P., and Luef, B. (2009). "Viruses," in *Encyclopedia of Inland Waters*. Ed. G. E. Likens (Oxford: Academic Press), 279–294.

Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi: 10.1093/bioinformatics/bts174

Pusch, M. T., and Schwoerbel, J. (1994) *Community respiration in hyporheic sediments of a mountain stream (Steina, Black Forest)*. Available at: https://www.semanticscholar.org/paper/42cf63746d7d54d34807e2e7eaa9c6a5c19ce17e (Accessed January 13, 2020).

Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., and Finn, R. D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* 46, D624–D632. doi: 10.1093/nar/gkx1134

R Core Team (2018). *A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing). Available at: https://www.R-project.org.

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11, R25. doi: 10.1186/gb-2010-11-3-r25

Rodríguez-Ramos, J. A., Borton, M. A., McGivern, B. B., Smith, G. J., Solden, L. M., Shaffer, M., et al. (2022). Genome-resolved metaproteomics decodes the microbial and viral contributions to coupled carbon and nitrogen cycling in river sediments. *mSystems* 7, e0051622. doi: 10.1128/msystems.00516-22

Rosentreter, J. A., Borges, A. V., Deemer, B. R., Holgerson, M. A., Liu, S., Song, C., et al. (2021). Half of global methane emissions come from highly variable aquatic ecosystem sources. *Nat. Geosci.* 14, 225–230. doi: 10.1038/s41561-021-00715-2

Roux, S., Adriaenssens, E. M., Dutilh, B. E., Koonin, E. V., Kropinski, A. M., Krupovic, M., et al. (2018). Minimum information about an uncultivated virus genome (MIUViG). *Nat. Biotechnol.* doi: 10.1038/nbt.4306

Roux, S., Chan, L.-K., Egan, R., Malmstrom, R. R., McMahon, K. D., and Sullivan, M. B. (2017). Ecogenomics of virophages and their giant virus hosts assessed through time series metagenomics. *Nat. Commun.* 8, 858. doi: 10.1038/s41467-017-01086-2

Rowe, J. M., DeBruyn, J. M., Poorvin, L., LeCleir, G. R., Johnson, Z. I., Zinser, E. R., et al. (2012). Viral and bacterial abundance and production in the Western Pacific Ocean and the relation to other oceanic realms. *FEMS Microbiol. Ecol.* 79, 359–370. doi: 10.1111/j.1574-6941.2011.01223.x

Ruiz-Perez, C. A., Tsementzi, D., Hatt, J. K., Sullivan, M. B., and Konstantinidis, K. T. (2019). Prevalence of viral photosynthesis genes along a freshwater to saltwater transect in Southeast USA. *Environ. Microbiol. Rep.* 11, 672–689. doi: 10.1111/1758-2229.12780

Shaffer, M., Borton, M. A., McGivern, B. B., Zayed, A. A., La Rosa, S. L., Solden, L. M., et al. (2020). DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* 48, 8883–8900. doi: 10.1093/nar/gkaa621

Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PloS One* 11, e0163962. doi: 10.1371/journal.pone.0163962

Shi, L.-D., Dong, X., Liu, Z., Yang, Y., Lin, J.-G., Li, M., et al. (2022). A mixed blessing of viruses in wastewater treatment plants. *Water Res.* 215, 118237. doi: 10.1016/j.watres.2022.118237

Skennerton, C. T., Imelfort, M., and Tyson, G. W. (2013). Crass: identification and reconstruction of CRISPR from unassembled metagenomic data. *Nucleic Acids Res.* 41, e105. doi: 10.1093/nar/gkt183

Stegen, J. C., and Goldman, A. E. (2018). WHONDRS: a community resource for studying dynamic river corridors. *mSystems* 3. doi: 10.1128/mSystems.00151-18

Sullivan, M. B., Lindell, D., Lee, J. A., Thompson, L. R., Bielawski, J. P., and Chisholm, S. W. (2006). Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PloS Biol.* 4, e234. doi: 10.1371/journal.pbio.0040234

Suttle, C. A. (2007). Marine viruses–major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750

Tian, R., Ning, D., He, Z., Zhang, P., Spencer, S. J., Gao, S., et al. (2020). Small and mighty: adaptation of superphylum Patescibacteria to groundwater environment drives their genome simplicity. *Microbiome* 8, 51. doi: 10.1186/s40168-020-00825-w

Tolić, N., Liu, Y., Liyu, A., Shen, Y., Tfaily, M. M., Kujawinski, E. B., et al. (2017). Formularity: software for automated formula assignment of natural and other organic matter from ultrahigh-resolution mass spectra. *Anal. Chem.* 89, 12659–12665. doi: 10.1021/acs.analchem.7b03318

Tomalski, P., Tomaszewski, E., Wrzesiński, D., and Sobkowiak, L. (2021). Relationships of hydrological seasons in rivers and groundwaters in selected catchments in Poland. *Water* 13, 250. doi: 10.3390/w13030250

Trubl, G., Kimbrel, J. A., Liquet-Gonzalez, J., Nuccio, E. E., Weber, P. K., Pett-Ridge, J., et al. (2021). Active virus-host interactions at sub-freezing temperatures in Arctic peat soil. *Microbiome* 9, 208. doi: 10.1186/s40168-021-01154-2

Villa, J. A., Smith, G. J., Ju, Y., Renteria, L., Angle, J. C., Arntzen, E., et al. (2020). Methane and nitrous oxide porewater concentrations and surface fluxes of a regulated river. *Sci. Total Environ.* 715, 136920. doi: 10.1016/j.scitotenv.2020.136920

Vincent, F., and Vardi, A. (2023). Viral infection in the ocean-A journey across scales. *PloS Biol.* 21, e3001966. doi: 10.1371/journal.pbio.3001966

Wang, J., Atolia, E., Hua, B., Savir, Y., Escalante-Chong, R., and Springer, M. (2015). Natural variation in preparation for nutrient depletion reveals a cost-benefit tradeoff. *PloS Biol.* 13, e1002041. doi: 10.1371/journal.pbio.1002041

Wang, J., Fan, H., He, X., Zhang, F., Xiao, J., Yan, Z., et al. (2021). Response of bacterial communities to variation in water quality and physicochemical conditions in a river-reservoir system. *Global Ecol. Conserv.* 27, e01541. doi: 10.1016/j.gecco.2021.e01541

Wang, Y., Zhang, Y., Hu, Y., Liu, L., Liu, S.-J., and Zhang, T. (2023). Genome-centric metagenomics reveals the host-driven dynamics and ecological role of CPR bacteria in an activated sludge system. *Microbiome* 11, 56. doi: 10.1186/s40168-023-01494-1

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28, 127–181. doi: 10.1016/j.femsre.2003.08.001

Weinbauer, M. G., and Rassoulzadegan, F. (2004). Are viruses driving microbial diversification and diversity? *Environ. Microbiol.* 6, 1–11. doi: 10.1046/j.1462-2920.2003.00539.x

Wells, J. R., Goldman, A. E., Chu, R. K., Danczak, R. E., Garayburu-Caruso, V. A., Graham, E. B., et al. (2019). WHONDRS 48 hour diel cycling study at the Erpe river, Germany. doi: 10.15485/1577260

Xie, Y., Liu, X., Wei, H., Chen, X., Gong, N., Ahmad, S., et al. (2022). Insight into impact of sewage discharge on microbial dynamics and pathogenicity in river ecosystem. *Sci. Rep.* 12, 6894. doi: 10.1038/s41598-022-09579-x

Zielezinski, A., Deorowicz, S., and Gudyś, A. (2021). PHIST: fast and accurate prediction of prokaryotic hosts from metagenomic viral sequences. *Bioinformatics*. doi: 10.1093/bioinformatics/btab837

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 22 September 2023
DOI 10.3389/frmbi.2023.1242895

OPEN ACCESS

EDITED BY
Graciela Gonzalez-Gil,
King Abdullah University of Science and Technology, Saudi Arabia

REVIEWED BY
Emilie Muller,
Génomique et Microbiologie (GMGM), France
Duntao Shu,
Northwest A&F University, China

*CORRESPONDENCE
Kylie B. Bodle
kyliebodle@montana.edu

RECEIVED 19 June 2023
ACCEPTED 08 September 2023
PUBLISHED 22 September 2023

CITATION
Bodle KB, Mueller RC, Pernat MR and Kirkland CM (2023) Treatment performance and microbial community structure in an aerobic granular sludge sequencing batch reactor amended with diclofenac, erythromycin, and gemfibrozil. *Front. Microbiomes* 2:1242895. doi: 10.3389/frmbi.2023.1242895

Treatment performance and microbial community structure in an aerobic granular sludge sequencing batch reactor amended with diclofenac, erythromycin, and gemfibrozil

Kylie B. Bodle[1,2]*, Rebecca C. Mueller[2,3], Madeline R. Pernat[1,2] and Catherine M. Kirkland[1,2]

[1]Department of Civil Engineering, Montana State University, Bozeman, MT, United States, [2]Center for Biofilm Engineering, Montana State University, Bozeman, MT, United States, [3]United States Department of Agriculture (USDA) Agricultural Research Service, Western Regional Research Center, Albany, CA, United States

This study characterizes the effects of three commonly detected pharmaceuticals—diclofenac, erythromycin, and gemfibrozil—on aerobic granular sludge. Approximately 150 µg/L of each pharmaceutical was fed in the influent to a sequencing batch reactor for 80 days, and the performance of the test reactor was compared with that of a control reactor. Wastewater treatment efficacy in the test reactor dropped by approximately 30-40%, and ammonia oxidation was particularly inhibited. The relative abundance of active *Rhodocyclaceae, Nitrosomonadaceae*, and *Nitrospiraceae* families declined throughout exposure, likely explaining reductions in wastewater treatment performance. Pharmaceuticals were temporarily removed in the first 12 days of the test via both sorption and degradation; both removal processes declined sharply thereafter. This study demonstrates that aerobic granular sludge may successfully remove pharmaceuticals in the short term, but long-term tests are necessary to confirm if pharmaceutical removal is sustainable.

KEYWORDS

aerobic granular sludge, wastewater treatment, pharmaceuticals and personal care products, emerging contaminants, biodegradation, bioremediation, microbial activity, microbiome

1 Introduction

Pharmaceutical consumption has increased concomitantly with population growth (Bernot et al., 2016). A natural consequence of this is increasing pharmaceutical contamination in the environment, due in part to incomplete metabolism by humans, followed by poor removal by conventional wastewater treatment systems (Xia et al., 2005;

Bernot et al., 2016). As such, improved wastewater treatment methods are worthy of investigation, as exposure to pharmaceuticals can cause antibiotic resistance gene proliferation as well as numerous other adverse effects on plants, animals, and microbiota (Kim and Aga, 2007; Bexfield et al., 2019; Świacka et al., 2021).

Aerobic granular sludge (AGS) is an emerging wastewater treatment biotechnology that may be capable of enhancing pharmaceutical removal from wastewater (Zhao et al., 2015; Wang et al., 2018). AGS is highly diverse with populations of nitrifying, denitrifying, and phosphate-accumulating organisms that self-aggregate into spherical biofilms. The gel-like extracellular polymeric substances (EPS) secreted by bacteria in AGS confer protection from toxins and enhance AGS density, resulting in short settling times and high biomass retention (Adav et al., 2008). Furthermore, the EPS in AGS may provide a sorptive medium for organic compound removal (Kong et al., 2015; Kent and Tay, 2019). However, the body of literature on AGS-driven pharmaceutical treatment is limited, and therefore more information is needed on how granules respond to a wide range of pharmaceuticals.

The pharmaceuticals diclofenac (DCF), erythromycin (ERY), and gemfibrozil (GEM) were selected for use in this study because, though each is frequently detected in the environment (Fang et al., 2012; Schafhauser et al., 2018; Sathishkumar et al., 2020), few studies exist on each compound's interaction with AGS. Diclofenac is a commonly used non-steroidal anti-inflammatory drug (NSAID) that is poorly removed (under 40%) in conventional wastewater treatment systems (Paxéus, 2004). It has also been shown to act synergistically with antibiotics to prevent biofilm formation (Pawar et al., 2019). Erythromycin is a macrolide antibiotic commonly used in both human and veterinary medicine, and has been shown to bioaccumulate in multiple aquatic species (Schafhauser et al., 2018). Lastly, gemfibrozil is a fibrate, or lipid regulator, that has been shown to inhibit growth and cause endocrine disruption in various aquatic organisms (Zurita et al., 2007). All three compounds have been detected in wastewater treatment plant influents at concentrations as high as 64 µg/L (Fang et al., 2012; Schafhauser et al., 2018; Sathishkumar et al., 2020).

There were three objectives of this study: (1) Identify how the three common, but relatively unstudied, pharmaceuticals listed above impact conventional wastewater treatment by lab-grown AGS; (2) investigate the fate of each pharmaceutical by tracking aqueous and solid phase parent compounds and degradation products; and (3) track microbial community and activity changes throughout exposure. To our knowledge, no studies have confirmed pharmaceutical biodegradation by AGS using detection of degradation products. Notably, the byproducts formed during degradation of pharmaceuticals may be more toxic than the parent pharmaceuticals, and therefore it is vital to improve understanding of the intermediates and products formed. Abiotic removal of the dosed pharmaceuticals was monitored in this study but is not discussed in order to allow sufficient discussion of biotic processes here. In particular, this study sought to link shifts in the active microbial community in pharmaceutical-exposed granules with changes in wastewater treatment efficacy and pharmaceutical fate.

2 Methods

2.1 AGS reactor operation

AGS sequencing batch reactor (SBR) operation is detailed in Bodle et al., 2022. In brief, AGS was grown in two identical glass SBRs with a working volume of 3.4 L. Both SBRs were operated in repeating three-hour cycles: 72 minutes anaerobic feed, 100 minutes aeration at a gas flow rate of 5 L/min, three minutes settling, and five minutes effluent discharge. The hydraulic residence time in both reactors was approximately 6.4 hours, and the solids residence time was controlled at approximately 25 ± 5 days. SBR operating parameters are consistent with those in other lab-scale studies (e.g., Kong et al., 2015; Margot et al., 2015; Rodriguez-Sanchez et al., 2017). During aeration, pH and dissolved oxygen were controlled with LabVIEW software (National Instruments) at 7.0 ± 0.3 and 1.75 ± 0.25 mg/L, respectively. Influent media were identical to those described by de Kreuk and van Loosdrecht, 2004, except that influent sodium acetate was increased to 10.3 mM, resulting in an organic loading rate of 2.5 g C/L*d. Both SBRs were initially seeded with AGS from an AquaNereda® treatment plant in Utrecht, The Netherlands and operated at steady state (i.e., complete nitrogen and phosphate removal) for over 300 days prior to beginning experimentation. Immediately before starting experimentation, both SBRs were emptied and AGS was combined, mixed, and redistributed so that granule qualities would be as similar as possible in both reactors.

For 80 days, the test reactor received 46 mL of pharmaceutical media with the influent medium, resulting in an influent concentration of approximately 150 µg/L of each pharmaceutical. Individual stock solutions of each pharmaceutical were prepared first in methanol at 10 g/L, then diluted into nanopure water to obtain two pharmaceutical media: one solution ("DG") containing 17.86 mg/L each of diclofenac sodium (Acros Organics) and gemfibrozil (Acros Organics), and a second solution ("ERY") containing 35.8 mg/L erythromycin (TCI Chemicals). Twenty-three mL of each solution (46 mL total) were delivered with the influent medium throughout the anaerobic feed period. Approximately 0.13 mg/L methanol was also present in the combined influent stream. The pharmaceutical media were protected from light to prevent photolytic degradation and prepared fresh every 8-10 days.

Results discussed in Bodle et al., 2022 showed that the pharmaceuticals tested herein sorbed to different lab materials with different affinities; therefore, to ensure the accuracy of pharmaceutical dosing, ERY solution was pumped into the reactor using silicone tubing (Masterflex). DG solution was pumped into the reactor using PharmaPure tubing (Masterflex). Despite different pharmaceutical stock solution concentrations, sorption to tubing resulted in fairly stable influent concentrations of approximately 150 µg/L of each pharmaceutical. Although each pharmaceutical is typically measured in wastewater treatment plant

influents at approximately 1-10 µg/L (Fang et al., 2012; Schafhauser et al., 2018; Sathishkumar et al., 2020), pharmaceutical sorption to tubing drove usage of this elevated influent concentration: 150 µg/L was the minimum concentration of pharmaceuticals that could be consistently dosed to the test reactor without significant losses from sorption. Influent samples were taken from a sampling port in the tubing located at the base of the reactor (reactor schematic available in Supplementary Information Figure S1) and were extracted and quantified per methods detailed in Section 2.3.

2.2 Analytical methods – conventional wastewater analytes

Influent and effluent samples from both reactors were regularly taken and filtered through 0.45 µm regenerated cellulose filters prior to analyses for ammonia, nitrite, nitrate, phosphate, and dissolved organic carbon. Ammonia was quantified with Hach kit TNT 832 and a Hach DR 3900 spectrophotometer, equivalent to US EPA Method 350.1. Other anions listed were quantified with a Dionex ICS-1100 anion chromatography system equipped with an IonPac AS22 RFIC column (4 x 250 mm) and IonPac guard column (4 x 50 mm). Dissolved organic carbon (DOC), defined as that which could pass through a 0.45 µm filter, was measured with a SKALAR Formacs^{HT} Total Organic Carbon analyzer system.

2.3 Analytical methods – pharmaceutical analyses

Aqueous influent and effluent samples were prepared for mass spectrometry (MS) analyses by solid phase extraction (SPE) as detailed in Bodle et al., 2022. It is important to note that the periodic flow conditions that are characteristic of SBR systems allow constant sorption and desorption of pharmaceuticals within influent tubing. For this reason, the most accurate method of quantifying influent concentrations would have required extraction of the entire influent volume, which would been too destructive to reactor operation. For that reason, 100 mL influent or effluent sample were filtered with a 1.5 µm glass fiber filter (Hach) to remove solids and loaded on preconditioned Waters Oasis HLB cartridges (30 mg, 20 mL) at 10 mL/min using a vacuum manifold system. Loaded cartridges were washed, dried, and frozen at -18°C until elution (no longer than 14 days). Pharmaceutical losses to the lab materials used during extraction (glassware, syringes, pipette tips, and glass fiber filters) were tested for and were found to be minimal (Bodle et al., 2022).

Influent and effluent pharmaceutical samples were periodically taken in duplicate to account for possible pharmaceutical losses during the extraction process, as detailed in Bodle et al., 2022. In brief, one sample from each duplicate set was pre-spiked with pharmaceutical stock solution to a final nominal concentration of 100 µg/L prior to extraction. After extraction, the unspiked sample was split in half, and one half was post-spiked to a final nominal concentration equal to 100 µg/L multiplied by each samples' concentration factor. Recovery was then determined as follows:

$$Recovery = \frac{Prespike\ \ concentration - unspiked\ \ concentration}{Postspike\ \ concentration - unspiked\ \ concentration}$$

Measured pharmaceutical concentrations in unspiked samples were thereby corrected for recovery of each analyte. Recovery was 97 ± 9% for DCF, 117 ± 17% for ERY, and 98 ± 11% for GEM (Table 1). Triplicate influent and effluent samples were taken once per month to assess accuracy. Pharmaceutical quantification and detection were performed with an Acquity I Class Plus ultra-performance liquid chromatograph (UPLC) coupled to a Waters Synapt XS quadrupole time-of-flight mass spectrometer (QToF-MS) in positive ion mode. Chromatographic analysis methods were adapted from Sodré and Sampaio, 2020. In brief, pharmaceuticals were separated with an Agilent Eclipse Plus C18 column (2.1 x 100 mm) at 30°C under gradient elution. Formic acid-enriched methanol and ultrapure water (0.1% v/v) were used as mobile phases at 0.7 mL/min, and the concentration of organic solvent was increased from 1% initially to 95% at 10.1 minutes, then returned to initial conditions and re-equilibrated for 3 minutes (total run length of 13 minutes). Sample injection volume was 8 µL. Retention times are summarized in Table 1.

A calibration curve relating each pharmaceuticals' peak area with its nominal concentration in analytical standards was used to determine parent compounds' concentrations. Standards were prepared in both methanol and water. The instrument limits of quantification were under 10 µg/L for all compounds. The open-source software MZmine and R were used to analyze and compile mass spectra data (RStudioTeam, 2009; Pluskal et al., 2010).

All samples were also screened for DCF, ERY, and GEM biodegradation products using a personal compound database

TABLE 1 Physical and chemical properties of tested pharmaceuticals, as well as average extraction recoveries and retention times.

	Diclofenac	Erythromycin	Gemfibrozil
Chemical formula	$C_{14}H_{11}Cl_2NO_3$	$C_{37}H_{67}NO_{13}$	$C_{15}H_{22}O_3$
Molecular weight (g/mole)	296.1	733.9	250.3
Octanol-water partition coefficient (Log K_{ow})	4.51 (Avdeef et al., 1998)	3.06 (Schafhauser et al., 2018)	4.77 (Fang et al., 2012)
Aqueous phase extraction recovery	97 ± 9%	117 ± 17%	98 ± 11%
Solid phase extraction recovery	75 ± 8%	56 ± 22%	68 ± 8%
UPLC retention time (minutes)	8.8	6.8	9.5

and library (PCDL) developed from the literature (SI Table S1). Degradation products are reported when mass errors were less than 5 ppm and the signal-to-noise ratio of peaks was greater than or equal to 10. The goal of this approach is not to quantify degradation products' concentrations but simply to use the presence of degradation products as an indicator of biodegradation. The relative concentrations of aqueous degradation products were tracked over time by calculating corrected peak areas as follows:

$$Aqueous\ \ corrected\ \ peak\ \ area = \frac{Degradation\ \ product\ \ peak\ \ area \times Concentration\ \ factor}{Peak\ \ area\ \ of\ \ relevant\ \ 100\frac{\mu g}{L}\ standard}$$

where the "relevant standard" term refers to the peak area of the related parent compound at 100 µg/L. For example, the peak area of an aqueous ERY degradation product would be divided by the peak area of the 100 µg/L ERY standard measured during the same mass spectrometry run and multiplied by the sample's concentration factor (e.g., a 100 mL sample extracted and concentrated to 1 mL had a concentration factor of 100).

Pharmaceuticals were also extracted from granules to quantify solid phase concentrations per methods adapted from Martin et al., 2010. Extraction methods are summarized in the Supplementary Information. All solid phase samples were screened for degradation products as described above. Solid phase samples were extracted in triplicate once per month and periodically evaluated for recovery using spike-recovery testing, as described above. Average recoveries are summarized in Table 1. Peak areas of degradation products in the solid phase were corrected as follows:

$$Solid\ \ corrected\ \ peak\ \ area = \frac{Degradation\ \ product\ \ peak\ \ area}{Peak\ \ area\ \ of\ \ relevant\ \ 100\frac{\mu g}{L}\ standard\ \times Sample\ \ dry\ \ weight\ \ (g)}$$

Aqueous and solid samples were periodically analyzed from the control reactor for pharmaceutical parents and degradation products. Except when noted otherwise, pharmaceuticals were not detected in any phase in the control reactor. Likewise, degradation products were generally not found in the influent to the test reactor.

Lastly, the toxicities of detected degradation products were estimated using the US Environmental Protection Agency's Toxicity Estimation Software Tool (TEST). TEST estimates chemical toxicity using quantitative structure activity relationships (Martin et al., 2020).

2.4 Bacterial community composition

2.4.1 DNA/RNA extraction

Granules from both reactors were periodically sampled for molecular characterization of the prokaryotic microbial community using high throughput sequencing of 16S rRNA genes and transcripts. Granules were collected during the aeration phase to ensure samples were fully mixed and representative of communities in the entire reactor. Samples were stored at -80°C prior to extraction. Nucleic acids were extracted from approximately 10 granules at each time point. Extraction and analyses of replicate nucleic acid samples were beyond the scope of this study. All samples were extracted at once by bead beating in DNA/RNA Shield (Zymo Research), then DNA was purified from the lysate with the DNA Clean and Concentrator Kit (Zymo Research). RNA samples were extracted from the same lysate with the Direct-zol RNA Miniprep Kit (Zymo Research), digested with the TURBO DNA-free Kit (Invitrogen), and purified with the RNA Clean and Concentrator Kit (Zymo Research). RNA was then reverse transcribed to cDNA using the ProtoScript II First Strand cDNA Synthesis Kit (New England Biolabs). DNA and RNA concentrations were quantified with Qubit dsDNA HS and RNA HS kits (Invitrogen), respectively. For all kits listed, the manufacturer-provided protocols were followed.

2.4.2 Metagenome characterization

Metagenomic sequencing of granules was conducted on granular inoculum approximately 300 days before beginning the experimentation described herein; therefore, metagenome data provides an approximation of microbe functionality within granules (given the time between metagenome sampling and the onset of the experiment). DNA was extracted using the Zymobiomics DNA Miniprep kit and quantitated with the Qubit dsDNA kit (Invitrogen). Libraries were prepared for shotgun metagenomic sequencing on the Illumina NovaSeq platform using paired end 150bp sequencing at Novogene with a target of 15Gb of data. To quality filter fastq files, prinseq was used to remove sequences with more than 10 Ns, mean quality scores below 20, sequences shorter than 50 nt, and to trim ambiguous bases from the ends of reads. Paired end sequences were then assembled using metaSPades (v 3.15.5, Nurk et al., 2017) with kmer sizes from 21 to 99 and all reads were mapped against the resulting scaffolds using Bowtie2 (Langmead and Salzberg, 2012). The sam files were converted to bam format using samtools (Li et al., 2009). We binned scaffolds into putative metagenome assembled genomes (MAGs) using MetaBat2 (v 2.2.15, Kang et al., 2019). CheckM (v 1.2.2, Parks et al., 2015) was used to calculate MAG quality, the relative abundance of each bin (based on the number of reads mapped to each MAG), and to construct a multigene phylogeny. Any MAG with completeness greater than 50% and with contamination above 10% was refined using Anvio (v 7.1, Eren et al., 2015). To classify the taxonomy of putative bins, we used GTDB-tk (Chaumeil et al., 2019) against the v207 database release.

To gain insights into the potential functional pathways contained within the abundant MAGs, putative genes were identified using prodigal and compared to the Kyoto Encyclopedia of Genes and Genomes (KEGG) with the program KOFAM scan (Aramaki et al., 2019). Pathway completeness measures were generated using KEGG Decoder (Graham et al., 2018). This analysis was focused on MAGs with completeness over 90%, combined with MAGs that had lower completeness but made up over 1% of the community. Our analysis focused primarily on nitrogen cycling, based on previous findings showing shifts in these pathways in reactors with pharmaceutical-exposed wastewater (Jiang et al., 2021).

2.4.3 16S library preparation and sequencing

To characterize the prokaryotic community, Phusion Hot Start II DNA polymerase (Thermo Scientific) was used to target the V4 region of the 16S rRNA gene with the primers 515F-A (GTGYCA GCMGCCGCGGTAA) and 806R-b (GGACTACVSGGGTATC TAAT). Reactions were 20 µL each with final concentrations of 0.4 mM dNTPs, 0.2 mM primers, and 1 U polymerase. Thermocycling conditions consisted of an initial denaturation at 98°C for 10 seconds followed by 30 cycles of denaturing at 98°C for 20 seconds, annealing at 60°C for 10 seconds, and elongating at 72° C for 30 seconds. A final extension was then performed at 72°C for 5 minutes. PCR products were purified with Mag-Bind TotalPure NGS beads (Omega), and samples were then barcoded with the Nextera XT Index Kit v2 Set D (Illumina). Barcoded samples were again purified with Mag-Bind beads, quantified, and pooled at equimolar concentration into a sample library containing approximately 30 ng DNA from each sample. Sequencing was performed onsite at Montana State University with the Illumina MiSeq platform using the v3 600 cycle kit. Raw sequence files were deposited to GenBank (BioProject ID PRJNA985155).

2.4.4 Statistics and data analysis

USEARCH software was used to merge forward and reverse reads, quality filter with a max score of 1, trim primer sequences, and dereplicate sequences. The UNOISE3 algorithm was used to identify zero-radius OTUs (ZOTUs) and construct an OTU table. To classify ZOTUs, we used SINTAX against a reference database of 16S sequences from the Genome Taxonomy Database (GTDB, release 202, Parks et al., 2021) with outgroup sequences for chloroplast and mitochondria. Sequences identified as Eukarya or with a bootstrap value less than 70% at the phylum level were removed from downstream analyses. A phylogenetic tree was constructed using a reference maximum likelihood tree generated from full length and near-full length 16S sequences downloaded from GenBank and GTDB using RAxML (Stamatakis, 2014). MAFFT (Katoh and Standley, 2013) was used to align the OTUs to the reference sequences; OTUs were then inserted into the reference tree with pplacer (Matsen et al., 2010).

The goal of targeting both rRNA genes and transcripts (cDNA) was to examine both the total microbial community, defined as the community recovered in DNA reads, and the active microbial community, defined as ZOTUs with a ratio of 16S rRNA transcripts to rRNA genes greater than or equal to 1 (Bowsher et al., 2019). Rarefied ZOTU matrices of rRNA transcripts (cDNA) and rRNA genes paired by sample were used to calculate the transcript to gene ratios and identify active OTUs across 100 rarefaction trials. To account for so-called phantom taxa, or taxa detected in rRNA but not rRNA gene sequences, values were set to 100 based on methods described in Bowsher et al., 2019 prior to calculating mean values across trials. Due to the potential for bias in sequence numbers arising during cDNA transcription, DNA read numbers were used to calculate diversity indices. DNA read numbers for phantom taxa were also included in the active community. To examine differences among the community pools, non-metric multidimensional scaling ordination was used to compare the total (DNA-based), RNA-based, and active communities. Taxa were grouped at the family level to calculate changes in relative abundance over time.

To link shifts in the active microbiome with changes in nitrogen and phosphorus levels, as well as pharmaceutical degradation, we used vector fitting of effluent concentrations of DCF, ERY, and GEM using the function envfit in vegan (Oksanen, 2010). More specifically, to examine if specific microbial groups could be linked to the strong decline in pharmaceutical removal observed between days 5 and 17, we calculated response ratios of the active ZOTUs with greater than 0.1% relative abundance between these time points. These values ranged from -1 to 1, with 1 being a strong increase in relative abundance at day 17, and -1 being a strong decrease. Values between -0.5 and 0.5 were considered neutral. The calculated values were added to the phylogeny as a dataset within the ITOL annotation platform to examine the potential for cohesive negative or positive responses within specific clades. All statistical analyses were conducted using R software (version 4.2.2, RStudioTeam, 2009) with packages vegan (Oksanen, 2010), picante (Kembel et al., 2010), and phyloseq (McMurdie and Holmes, 2013).

3 Results

3.1 Pharmaceutical impacts on granular wastewater treatment performance

Immediately after pharmaceutical dosing to the test reactor began, nitrogen removal dropped to approximately 70% (Figure 1). There was a brief recovery in nitrogen removal from day 6 to 20, during which removal peaked at 93%, though removal declined thereafter and stabilized at approximately 70%. Poor nitrogen removal was due mainly to incomplete ammonia oxidation (SI Figure S2). Nitrite and nitrate concentrations in both control and test reactors were approximately equal for the entire experimental duration (Figure S3), indicating that nitrite oxidation and denitrification were minimally impacted by pharmaceutical exposure. Nitrogen removal data for the control reactor are not plotted past day 40 because the software controlling both reactors experienced an error on day 40, stopping dissolved oxygen control in the test reactor and resulting in an acid overdose to the control reactor. DO control to the test reactor was interrupted for just 24 hours; however, the acid overdose to the control reactor caused shutdown of that reactor for 10 days. It is reasonable to expect that complete nitrogen removal would have continued in the control reactor if not for this disruption.

Phosphate removal also declined sharply in the test reactor and remained noisy for the entire experimental duration (Figure 1). Despite this, DOC consumption was over 90% in both reactors (Figure S4), and most carbon continued to be consumed anaerobically in the test reactor (Figure S5). Near-complete anaerobic carbon consumption suggests the potential activity of glycogen accumulating organisms, discussed more in section 4.1.

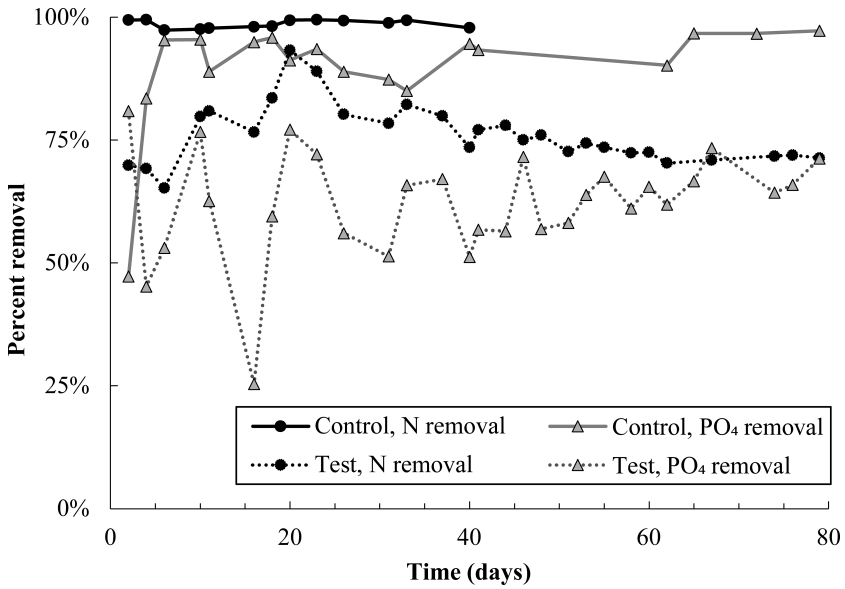


FIGURE 1
Total nitrogen and phosphate removal in the control and test SBR during the pharmaceutical dosing period. Note that N removal data in the control reactor is not plotted after day 40 because an acid overdose in the control reactor severely inhibited nitrifying populations. Given trends prior to the acid overdose, it is likely that nitrogen removal in the control would have proceeded at 100% if not for this issue. Nitrogen and phosphate removal in the test reactor averaged out at 73 ± 2% and 63 ± 6%, respectively, over the last 40 days of the experiment.

3.2 Pharmaceutical removal

All pharmaceuticals were partially removed in the first 10-12 days of dosing, evidenced by lower effluent than influent concentrations (Figure 2, Figure S6). Removal was calculated by performing a mass balance on influent and effluent concentrations at each time point (SI Equations S1 and S2). Degradation products were also detected during the first 12 days (Figure 3). However, from day 12-23, effluent concentrations of all pharmaceuticals spiked to approximately twice influent concentrations, indicating a release of retained pharmaceuticals from the reactor. Solid phase data collected over this time frame also appear to indicate desorption of pharmaceuticals from AGS (Figure S6). Pharmaceutical degradation is discussed in more detail in the following sections.

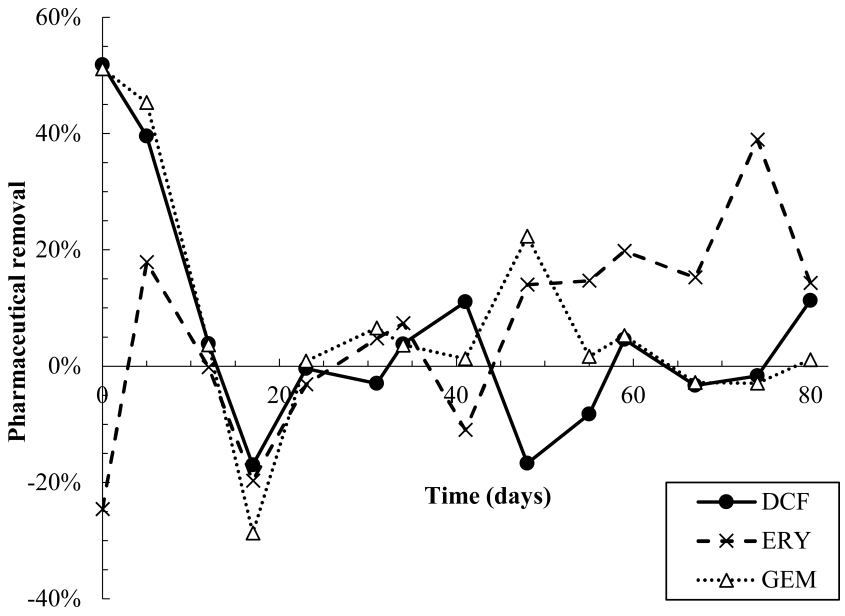


FIGURE 2
Pharmaceutical removal versus time. Removal was calculated based on a mass balance using measured influent and effluent concentrations (SI equations S1 and S2). Negative removal percentages indicate that effluent concentrations were higher than predicted by the mass balance.

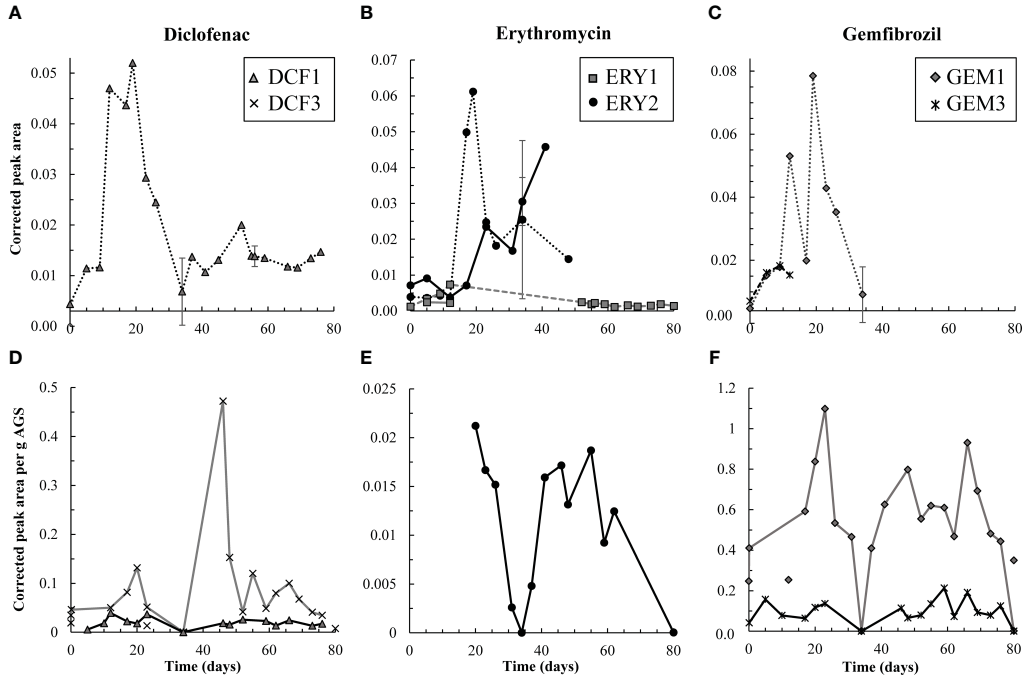


FIGURE 3
Top row, **(A-C)** aqueous degradation products detected over time in the effluent (dashed and dotted lines) from the test reactor for DCF, ERY, and GEM, respectively. Y-axes all reflect corrected peak area but differ in scale. ERY-associated degradation products were temporarily detected in the influent (solid lines) to the test reactor. However, corrected peak areas of both ERY-associated products were generally higher in the effluent than influent, and therefore AGS-driven biodegradation of these compounds was likely occurring. Bottom row, **(D-F)** solid phase degradation products for DCF, ERY, and GEM, respectively. Note that y-axes differ in scale though each reflects corrected peak area per g AGS. Points not connected by a line indicate detections in control granules. DCF3 was detected three times in control granules, likely due to cross-contamination during mass spectrometry analyses. Likewise, GEM1 was detected twice in control granules. Error bars, representing standard deviation of triplicate samples, are present for both aqueous and solid data on days 34, 56, and 80, and at times are smaller than sample points. Points on these days are averages.

3.2.1 Diclofenac

Pharmaceutical fate was most clearly interpreted for DCF: over the first 12 days, solid phase DCF concentrations increased, while effluent concentrations remained lower than influent ones, and removal peaked at approximately 50% (Figure 2). The DCF degradation products hydroxy-DCF ($C_{14}H_{11}Cl_2NO_3$, "DCF1") and 1-(4-hydroxyphenyl)-2-,3-dihydro-1H-indol-2-one ($C_{14}H_{11}NO_2$, "DCF3") were also present in the aqueous and solid phases in the first 12 days (Figure 3). Taken together, these data indicate removal via sorption and biodegradation.

From days 17-23, desorption and a decline in biodegradation capacity likely caused DCF to spike in the effluent: on days 17 and 19, effluent concentrations peaked while solid phase DCF concentrations sharply declined (Figure S6), suggesting desorbing DCF contributed to higher concentrations in the effluent. Likewise, aqueous phase DCF1 peaked from days 12-19 (Figure 3), suggesting that bacteria in AGS were less capable of converting this product to further intermediates. Solid phase peak areas of DCF1 also increased slightly over the same timeframe, likely because aqueous phase concentrations were higher and therefore increased sorption was possible.

Influent and effluent DCF concentrations were near equal from day 23 on, indicating negligible DCF removal. Interestingly, aqueous and solid phase DCF degradation products were detected for the remainder of the experiment. The presence of these products may indicate partial biodegradation, albeit not at rates significant enough to result in measurable removal.

The degradation products detected represent initial and tertiary products, according to a pathway proposed in Jewell et al., 2016: DCF1 is formed first via mono-oxygenation. DCF1 may be present as two isomers, 4'-hydroxy-DCF and/or 5-hydroxy-DCF, but the UPLC-QToF-MS method used herein was unable to differentiate between the two. It is most likely that DCF1 was present as the 4' isomer, as DCF3 is formed from degradation of this isomer. 4'-hydroxy-DCF is then converted to 1-(2-chloro-4-hydroxyphenyl)-3H-indol-2-one ($C_{14}H_{10}NO_2Cl$, "DCF2") via reductive dechlorination and amidation, and DCF3 is then formed via further reductive dechlorination (Jewell et al., 2016). Further degradation of DCF3 is hypothesized in Jewell et al., 2016, but intermediate structures are unknown.

Pharmaceutical degradation products may have had an inhibitory effect on bacteria in AGS; however, the impacts of many degradation products on wastewater bacteria or other environmental receptors are not well understood. Regarding the herein-detected DCF degradation products, a study by Syed et al., 2016 showed that 4'-hydroxy-DCF inhibits ATP synthesis in rat liver mitochondria, though concentrations tested were higher than would be expected in humans and significantly higher than those in the environment. Conversely, 4'-hydroxy-DCF did not have any inhibitory impact on *Vibrio fischeri* bacteria at up to 20 mg/L

(Grandclément et al., 2020). The toxicity of DCF3 has not been established in the literature, therefore TEST was used. Toxicity estimates for all detected degradation products are summarized in SI Table S2. In general, the predicted toxicities for detected degradation products are similar to or slightly less toxic than those for DCF, with the exception of DCF1, which is predicted to be more bioaccumulative and toxic to fathead minnows.

3.2.2 Erythromycin

Similar to DCF, ERY was removed in the first 12 days of dosing via both sorption and biodegradation. Effluent concentrations were lower than influent ones for the first five days, and aqueous phase degradation products were also measured at low levels over the initial 12-day period (Figure 2, Figure 3, Figure S6). Notably, ERY removal was negative on day zero, indicating that measured effluent concentrations were higher than those predicted by a mass balance (Figure 2). It is probable that negative removal on day zero is an artifact of noisy influent pharmaceutical concentrations, as discussed in Section 2.3 and shown in Figure S6. Influent ERY concentrations on day zero were likely higher than measured, given that influent concentrations then stabilize at approximately 207 ± 22 µg/L for the next 20 days. Between days 12-23, aqueous ERY degradation products then spiked, as did effluent ERY concentrations, evidenced by negative removal (Figure 3).

Low levels of the primary and secondary products 3-depyranosyloxy ERY ($C_{29}H_{53}NO_{10}$, "ERY1") and 7,12-dyhydroxy-6-deoxyerythronolide B ($C_{21}H_{38}O_8$, "ERY2"), respectively (Llorca et al., 2015; Ren et al., 2021), were present up to day 12, indicative of biodegradation. On days 17 and 19, ERY2 concentrations spiked in the effluent, likely indicating that conversion of this secondary product to further intermediates was no longer occurring at the same capacity. Interestingly, both products were also present in the influent (solid lines in Figure 3B). Influent ERY1 concentrations were lower than effluent ones, indicating that AGS contributed to formation of ERY1. ERY2 concentrations were higher in the influent than effluent and continued to increase until day 40. The presence of both ERY1 and ERY2 in the influent likely indicates photodegradation or ERY biodegradation by contaminating biomass in the influent tubing. ERY2 was also present in the solid phase throughout the test.

After day 20, effluent ERY concentrations declined to near influent ones. Approximately 19 ± 9.7% removal was sustained after day 48, perhaps due to slightly elevated solid phase concentrations over the same time (Figure 2, Figure S6). The presence of aqueous and solid phase ERY degradation products for the remaining test duration may also indicate partial removal via biodegradation.

ERY degradation pathways potentially used by wastewater bacteria and proposed in Ren et al., 2021 and Ren et al., 2022 hypothesize that ERY1 is formed first via cleavage of the cladinose sugar from ERY. Cladinose was not detected in this study, likely because it is readily metabolizable. ERY1 then undergoes further degradation to ERY2 via cleavage of the desosamine sugar from ERY1, and the final product of ERY degradation is 2,4,6,8,10,12-hexamethyl-3,5,6,11,12,13-hexahydroxy-9-ketopentadecanoic acid ($C_{21}H_{40}O_9$, "ERY3"). Desosamine and ERY3 were also not detected, again likely because both are readily metabolizable, or because the mass spectrometry method used was not suitable for these compounds. Although ERY removal peaked at 18% in the first 12 days of dosing, the low levels of both ERY1 and ERY2 detected in that period suggest that intermediates completed the degradation pathway and end products were fully metabolized.

The TEST-predicted toxicities of ERY1 and ERY2 are summarized in SI Table S2. ERY1 and ERY2 appear to be less toxic to aquatic organisms than ERY, though both are predicted to be more toxic to rats.

3.2.3 Gemfibrozil

Much like DCF and ERY, GEM removal occurred due to a combination of sorption and biodegradation in the first 12 days of dosing, followed by a spike in effluent concentrations from days 12-23 (Figure 2). Removal was then near zero for the remainder of the test. Aqueous GEM degradation products were not detected in the effluent after day 34 (Figure 3). The primary and tertiary products 5-[2-(hydroxymethyl)phenoxy]-2,2-dimethylpentanoic acid ($C_{15}H_{22}O_4$, "GEM1") and 2-[(4-carboxy-4-methylpentyl)oxy] benzoic acid ($C_{15}H_{20}O_5$, "GEM3"), respectively (Kjeldal et al., 2016), were both measured up to day 12. GEM1 then spiked in the effluent. This pattern also mimics that seen for DCF and ERY degradation products: Primary products were converted to downstream intermediates until day 12, after which conversion stopped occurring to the same extent, causing primary products to wash out in the effluent.

The same primary and tertiary products were present in the solid phase for the entire experimental duration (Figure 3), which may indicate a preference of these compounds for the solid phase and/or continuous low levels of biodegradation. Regardless, biodegradation was not significant enough to impact GEM removal. GEM1 was also detected twice in control granules, likely due to carry over during mass spectrometry analyses: GEM1 was not detected in control influent or effluent, and relative solid phase concentrations were lower in control samples than test granules.

Kjeldal et al., 2016 isolated a GEM-degrading *Bacillus* species from activated sludge and proposed a degradation pathway in which GEM is first converted to GEM1 via hydroxylation. GEM1 is next oxidized to 5-(2-formylphenoxy)-2,2-dimethylpentanoic acid (GEM2), which is then further oxidized to GEM3. According to the degradation pathway proposed in Kjeldal et al., 2016, GEM3 degradation may undergo two to three further reactions before reaching final products. Based on this, the GEM degradation observed in the first 12 days may have proceeded to approximately 50% completion, though degradation rates were insufficient for complete removal. Further GEM degradation may have also occurred, but UPLC-MS methods may not have been suited to detect other products. TEST-predicted toxicities of GEM2 and GEM3 are summarized in SI Table S2. Predicted toxicities and bioconcentration factors are generally lower for degradation products than for GEM.

3.3 Functional potential of granules from shotgun metagenomics

Metagenomic sequencing generated approximately 25 million high quality reads, resulting in a total of 19 bins with greater than 90% completeness and less than 5% contamination, and an additional 38 with completeness greater than 50% with less than 10% contamination. All MAGs were classified as members of the bacteria, including known members of glycogen accumulating organisms (GAOs) such as species in the *Competibacter* and *Contendibacter*. Phosphorus accumulating organisms (PAOs) were also present, such as species within the *Accumulibacter*. A MAG identified as a species in the *Accumulibacter* was the most abundant organism detected, at 19% of the population based on the number of raw reads that mapped to the MAG.

Despite the presence of multiple MAGs identified as known ammonia oxidizing bacteria (AOB), including *Nitrosospira* and *Nitrosomonas*, this pathway was not uncovered in the KEGG analysis of constructed MAGs, though the complete ammonia oxidation pathway was found in the total shotgun community (Figure 4). The scaffolds with identified ammonia monooxygenase (*amo*) genes (*amoA, amoB*, and *amoC*) were relatively short (< 6000 bp). One scaffold contained all three *amo* genes, and another scaffold contained just *amoA* and *amoC*. Based on BLAST (Altschul et al., 1990), both were highly similar to sequences identified from *Nitrosomonas* species; one scaffold was similar to a MAG identified as *Nitrosomonas* from a metagenomic analysis of biofilms in wastewater treatment plants (Suarez et al., 2022). The absence of this pathway in putative AOB in the assembled MAGs could be due to the relatively low genome completeness calculated for these taxa, as the *Nitrosomonas* MAG (Granule57, Figure 4) was 87% complete; however, the relative abundance of this organism was 0.66%, and was not included in the overall KEGG analysis. In addition, this MAG had the pathway for nitrite reduction and nitric oxide reduction associated with NOB. We also identified the complete pathway for hydroxylamine oxidation, the second step in nitrification, in the total community, but not in the assembled MAGs.

Nitrate reductase genes were widespread in the granule community, as seen in previous metagenomic studies of wastewater (Singleton et al., 2021). Twelve MAGs contained genes for the complete dissimilatory nitrate reduction pathway and of those, ten also contained the full nitrite oxidation pathway (Figure 4) including members of the *Rhodocyclaceae*, a

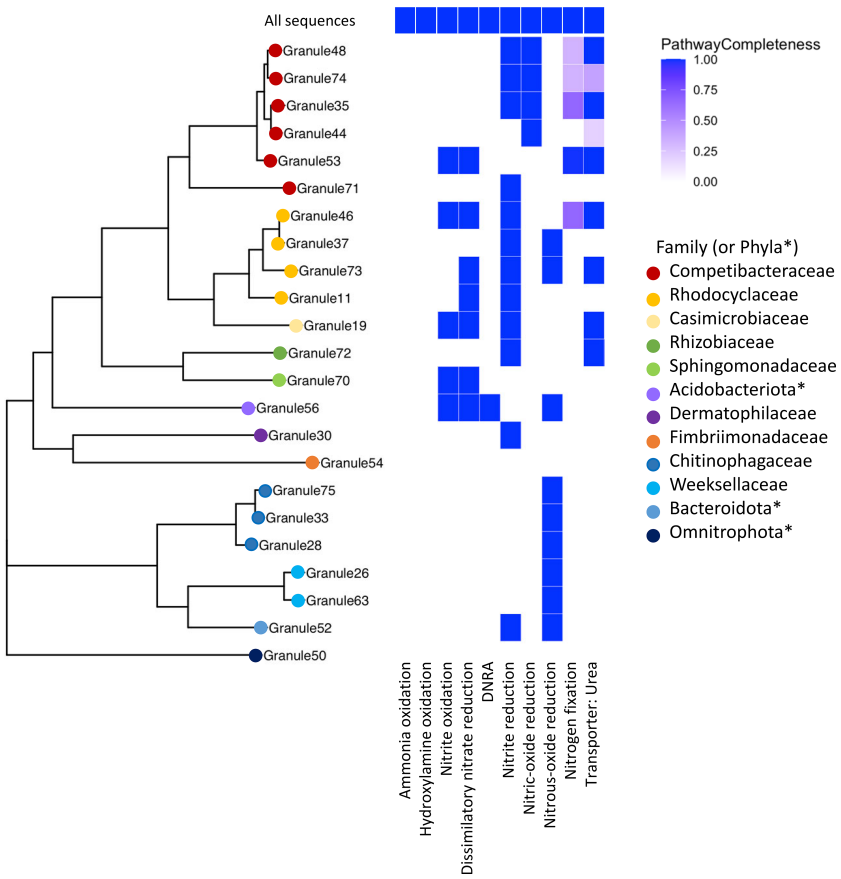


FIGURE 4
A multigene phylogeny of assembled MAGs and the level of completeness for nitrogen cycling pathways. Organisms identified from the assembled MAGs were from a range of families, with a large number in the Pseudomonadota (e.g., Competibacteraceae, Rhodocyclaceae) and the Bacteroidota. *Phyla, rather than family, are listed when the GTDB-provided family name corresponds to an uncultivated organism. None of the assembled MAGs contained the ammonia oxidation pathway, but multiple MAGs had complete denitrification pathways, primarily for dissimilatory nitrate reduction. Note that granule numbering is arbitrary and was simply used to differentiate different samples.

Competibacter, and a MAG classified as a member of the Thermoanaerobaculia. Only a single MAG (Granule56) had the pathway for dissimilatory nitrate reduction to ammonia (DNRA), identified as a species in the Thermoanaerobaculia in the phylum Acidobacteriota, which also contained the dissimilatory nitrate reduction pathway. An organism related to the Thermoanaerobaculia was also detected in the active community, but at very low levels (< 0.1% relative abundance).

3.4 Active microbial community response to pharmaceuticals

Shifts in the active microbial community were evaluated using 16S sequencing of rRNA genes and transcripts throughout pharmaceutical exposure. As described in Section 2.4.4, the active microbial community was defined as ZOTUs with a ratio of 16S rRNA transcripts (i.e. cDNA) to rRNA genes greater than or equal to one. The 16S rRNA gene read numbers from ZOTUs identified as "active" were then used to calculate relative abundances within the total active community. Multidimensional ordination (SI Figure S7) showed separation of the three pools (DNA, RNA, and active), but also showed separation of the treatment and control SBRs.

At day zero, the distributions of active families in both the control and test reactors were similar: *Competibacteraceae*, *Rhodocyclaceae*, and *Chitinophagaceae* were most abundant (Figure 5). Notably, the *Rhodocyclaceae* family includes PAOs (*Candidatus Accumulibacter)* and a potential denitrifying genus (*Zoogloea*). The role of bacteria in the *Chitinophagaceae* family is unclear—multiple OTUs matching *Niabella*, *Terrimonas*, and *Flavipsychrobacter* genera were detected, but to date, the function of these genera in AGS is unknown.

After five days, active families in the test and control SBRs diverged greatly. In the test reactor, active *Rhodocyclaceae* spiked in abundance at day five, from 14 to 46% (Figure 5), which may be linked with the brief recovery in phosphate removal between days four and 10 (Figure 1). Active *Rhodocyclaceae* then declined sharply in abundance in the test reactor for the remainder of the dosing period, stabilizing at approximately 1.5 ± 0.7% abundance for the second half of the test. In contrast, active *Rhodocyclaceae* in the control reactor were present at 10.2 ± 5% for the entire test duration.

Despite large shifts in the test reactor community, significant correlations between microbial community composition and effluent pharmaceutical concentrations were generally not found. A significant, though small, correlation between effluent GEM concentrations and the microbial community was observed (R^2 = 0.28, p = 0.048).

Despite the lack of overall correlations, when times with the largest differences in pharmaceutical removal (specifically days 5 through 17, Figure 2) were focused on, the strongest declines in relative abundance were concentrated within the Gammaproteobacteria and primarily in *Rhodocyclaceae* (Figure 5, Figure S8). The most striking shift for any single ZOTU was observed for a sequence classified as a member of *Azonexus (*family *Rhodocyclaceae*), which made up 40% of the active community at day 5 but was absent by day 17. Large changes in relative abundance were also observed for two members of *Accumulibacter*, which were present at over 1% of the active community on day 5 but were not detected by day 17.

In contrast, positive responses were distributed across the phylogeny, with the largest shift occurring in a ZOTU identified as a *Methylotenera* species: relative abundance increased from under 0.1% at day 5 to 22% at day 17. Likewise, a ZOTU identified as a Bacteroidia species (UBA2475_sp013816615) increased from under 0.1% to 6% over the same time period. Positive shifts were also observed for members of the Chloroflexota, Gammaproteobacteria, and Acidobacteriota (Figure S8).

We had limited success linking the ZOTUs to the organisms identified in the metagenome assemblies. Of the top ten most abundant ZOTUs in the active community, three had no close relatives in the metagenome. The most abundant active ZOTU in

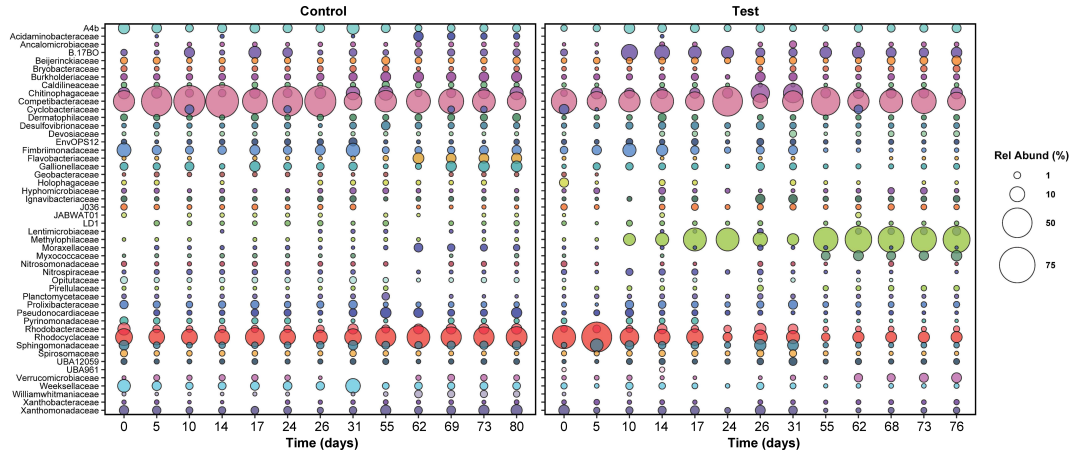


FIGURE 5
The relative abundance of active families in both reactors over time. Only families with relative abundance greater than or equal to 1% at at least one time point are plotted. Sampling times were matched across both reactors; for this reason, relative abundance data between days 31 and 55 are not plotted, since the control reactor was shut down during this time.

the test reactor (genus *Methylotenera*, order Burkholderiales), was not detected in the assembled MAGs; however, this is likely due to the length of time elapsed between metagenome sampling and the onset of experimentation as well as the presence of pharmaceuticals. Four ZOTUs were classified as *Competibacter A denitrificans* with varying levels of confidence. The remaining ZOTUs with a close relative in the metagenome included a *Xanthomonadaceae* (genus SCMT01), a species of *Accumulibacter*, and an *Azonexus* species.

4 Discussion

4.1 Links between active bacterial community, wastewater treatment, and pharmaceutical degradation

Methylophilaceae became the dominant active family in the test reactor after day 17 and included *Methylotenera* and *Methylophilus* genera. Members of this family may have proliferated due to the presence of trace methanol in the influent from pharmaceutical dosing (0.13 mg/L influent concentration); however, some *Methylophilaceae* species can also consume acetate (Jenkins et al., 1987; Dedysh et al., 2005), and therefore it is more likely that *Methylophilaceae* proliferated due to increased acetate availability in the aerobic phase of SBR operation (Figure S5). Some *Methylotenera* species are also capable of aerobic denitrification (Mustakhimov et al., 2013), which may explain why nitrate did not accumulate in the test reactor. It is worth noting that this family did not appear to be affected by pharmaceuticals' presence, but also did not proliferate until after pharmaceutical removal stopped—*Methylophilaceae* were therefore likely not responsible for the pharmaceutical degradation observed in the first 12 days of the test.

Competibacteraceae was the second most abundant active family in the test reactor, at 21.8 ± 10.5% abundance over the last 40 days of exposure. Bacteria in the *Competibacteraceae* family are glycogen accumulating organisms, or GAOs, that compete with PAOs for anaerobic carbon consumption but do not aerobically consume phosphate. For this reason, their proliferation is typically linked with reduced phosphate removal (McIlroy et al., 2014). It is likely that *Competibacteraceae* activity sustained anaerobic carbon consumption in the test reactor; however, it is surprising that the relative abundance of this family was lower in the test reactor than in the control reactor: active *Competibacteraceae* were present at 30.8 ± 18.6% abundance in the control reactor, despite complete phosphate removal (Figure 1). It is possible that *Competibacteraceae* activity levels were similar in both reactors; the increased abundance of active *Methylophilaceae* in the test reactor may make the relative abundance of active *Competibacteraceae* appear smaller.

The relative abundance of active nitrifying families *Nitrospiraceae* and *Nitrosomonadacea* was higher in the test reactor than the control for the first 38 days of pharmaceutical exposure (Figure 6), which may explain why nitrogen removal briefly recovered between days zero and 20 (Figure 1). Nitrifier activity may also be responsible for pharmaceutical degradation in the first 12 days of the test: the ammonia monooxygenase enzyme used by both ammonia oxidizing bacteria and archaea (AOA) is known to react non-specifically with aromatic compounds, typically through an oxidation reaction (Keener and Arp, 1994; Schatteman et al., 2022). The primary degradation product of DCF, and all detected GEM degradation products, are formed through oxidation reactions. Aqueous GEM degradation products were not detected after day 34, approximately the same time that active nitrifier abundance dropped. Likewise, aqueous DCF1 levels fell and remained low after day 34. It is possible that decreased levels of DCF and GEM degradation products after day 34 are linked with decreased abundance of active nitrifiers over the same time period. It should be noted, however, that active *Nitrosomonadaceae* (the only AOB family detected) were not detected for most timepoints in both the control and test SBR. The lack of active AOB data,

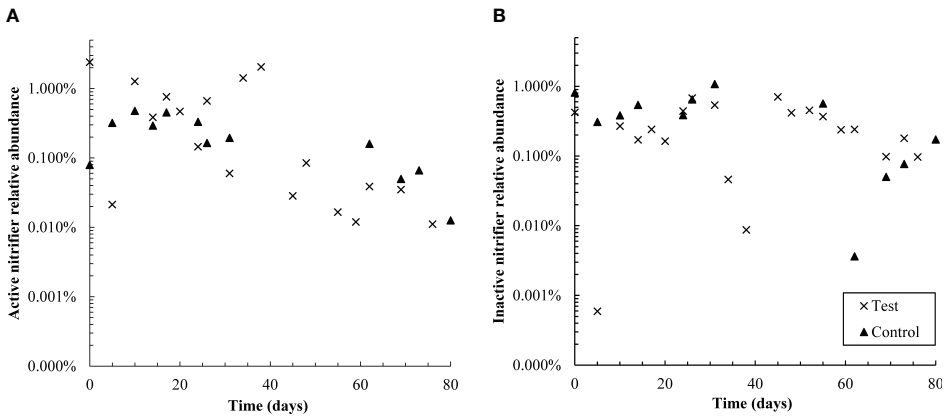


FIGURE 6
Semi-log plots of the relative abundance of active **(A)** and non-active **(B)** nitrifiers (summed Nitrosomonadaceae and Nitrospiraceae families) over time in control and test reactors. Over the first 38 days, active nitrifiers were more abundant in the test reactor than the control; however, from days 55-80, active nitrifiers were more abundant in the control SBR. Active nitrifiers also declined in relative abundance in the test reactor throughout the dosing period. The control SBR was shut down from days 40-50 and therefore data is not available in that timeframe; this is also likely why active nitrifiers in the control reactor are present at slightly lower relative abundances in the second half of the experiment than the first. Inactive nitrifiers were present at similar levels in both reactors throughout the test.

however, is contradicted by complete ammonia oxidation in the control reactor and partial ammonia oxidation in the test reactor, as well as active *Nitrospiraceae* abundance patterns over the dosing period. Several active AOA genera were also detected, which may further explain why ammonia oxidation continued throughout dosing. AOA abundance is discussed in more detail in section 4.2.

4.2 Patterns in taxonomic shifts in microbial communities

The addition of pharmaceutical compounds led to strong shifts in the microbial community and concomitant reductions in nitrogen and phosphorus removal. Based on the assessment of potential functional pathways, no high-quality MAGs were found with the ammonia oxidation pathway. Although AOB, such as *Nitrosomonas*, were detected in both the shotgun metagenomic and metabarcoding datasets, we were unable to construct a MAG with over 90% completeness, which limited our ability to determine the role of these organisms. However, despite their limited abundance overall, it is likely that *Nitrosomonas* is the main driver of ammonia oxidation in the reactors, given their presence in seed granules and the homology of the identified functional genes.

Several archaeal taxa were also found in the metabarcoding dataset though archaea were not detected in the granule metagenome. Archaea detected were primarily species within the *Thermoproteota*, including members of the order *Nitrosophaerales*. *Nitrosophaerales* contains numerous AOA, such as *Nitrosotenuis* and *Nitrosocaldus* genera. AOA have been posited to be important for the oxidation of ammonia to nitrite in commercial wastewater treatment plants (Limpiyakorn et al., 2013), though a study by Bai et al., 2012 found that AOA dominated in smaller scale treatment plants while AOB dominated in commercial ones. The authors hypothesized that AOA may be more sensitive to the toxic compounds more frequently measured in larger scale treatment plants.

Although archaea made up less than 1% of the sequences identified, 47 ZOTUs were identified as active taxa. We focused primarily on abundant taxa, but the ability of rare organisms to drive ecosystem function is well documented (Jousset et al., 2017); for example, rare taxa may be particularly important in pollutant degradation. The large shifts in relative abundance of wastewater-treating taxa, such as *Rhodocyclaceae* and *Methylophilaceae* (Figure 5), suggest that additional studies in which microbe functionality is explored could provide important insights into community interactions that might otherwise be overlooked.

5 Conclusions

Lab-grown AGS was exposed to three commonly found, but relatively unstudied pharmaceuticals at approximately 150 μg/L each. The fate of each pharmaceutical and its degradation products in the aqueous and solid phases were monitored, and pharmaceutical impacts on wastewater treatment performance and microbial communities were evaluated.

All pharmaceuticals were partially removed via both biodegradation and sorption in the first 12 days of the study. Biodegradation capacity then declined irreversibly, indicated by washout of degradation products, declining production, and negligible pharmaceutical removal. Exposure to the pharmaceutical mixture negatively impacted wastewater treatment efficacy and the relative abundance of active wastewater treating families. Nitrogen and phosphate removal declined to approximately 73% and 63%, respectively, though carbon removal was not impacted. Declining nitrogen removal was due mainly to inhibited ammonia oxidation and likely also related to the declining abundance of active nitrifiers. Similarly, active *Rhodocyclaceae* declined in abundance in the test reactor, which likely contributed to poor phosphate removal.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, PRJNA985155.

Author contributions

All authors contributed to design, conception, and investigation of this study. KB, RM, and MP collected data and performed analyses. CK acquired funding for the study and both CK and RM provided supervision. KB wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by Montana INBRE, which is funded by the National Institute of General Medical Sciences division of the National Institutes of Health under Award Number P20GM103474. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgments

The authors would like to thank staff at the Mass Spectrometry Facility at Montana State University (MSU) for training and assistance. Funding for the Proteomics, Metabolomics and Mass Spectrometry Facility used in this publication was made possible in part by the MJ Murdock Charitable Trust, the National Institute of General Medical Sciences of the National Institutes of Health under Award Numbers P20GM103474 and S10OD28650, and the MSU Office of Research, Economic Development and Graduate Education.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1242895/full#supplementary-material

References

Świacka, K., Michnowska, A., Maculewicz, J., Caban, M., and Smolarz, K. (2021). Toxic effects of NSAIDs in non-target species: A review from the perspective of the aquatic environment. *Environ. pollut.* 273, 115891. doi: 10.1016/j.envpol.2020.115891

Adav, S. S., Lee, D. J., Show, K. Y., and Tay, J. H. (2008). Aerobic granular sludge: Recent advances. *Biotechnol. Adv.* 26 (5), 411–423. doi: 10.1016/j.bioteChadv.2008.05.002

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215 (3), 403–410. doi: 10.1016/s0022-2836(05)80360-2

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2019). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36 (7), 2251–2252. doi: 10.1093/bioinformatics/btz859

Avdeef, A., Box, K. J., Comer, J. E. A., Hibbert, C., and Tam, K. Y. (1998). Determination of liposomal membrane-water partition coefficients of ionizable drugs. *Pharm. Res.* 15 (2), 209–215. doi: 10.1023/A:1011954332221

Bai, Y., Sun, Q., Wen, D., and Tang, X. (2012). Abundance of ammonia-oxidizing bacteria and archaea in industrial and domestic wastewater treatment systems. *FEMS Microbiol. Ecol.* 80 (2), 323–330. doi: 10.1111/j.1574-6941.2012.01296.x

Bernot, M. J., Becker, J. C., Doll, J., and Lauer, T. E. (2016). A national reconnaissance of trace organic compounds (TOCs) in United States lotic ecosystems. *Sci. Total Environ.* 572, 422–433. doi: 10.1016/j.scitotenv.2016.08.060

Bexfield, L. M., Toccalino, P. L., Belitz, K., Foreman, W. T., and Furlong, E. T. (2019). Hormones and pharmaceuticals in groundwater used as a source of drinking water across the United States. *Environ. Sci. Technol.* 53 (6), 2950–2960. doi: 10.1021/acs.est.8b05592

Bodle, K. B., Pernat, M. R., and Kirkland, C. M. (2022). Pharmaceutical sorption to lab materials may overestimate rates of removal in lab-scale bioreactors. *Water Air Soil pollut.* 233 (12), 505. doi: 10.1007/s11270-022-05974-2

Bowsher, A. W., Kearns, P. J., and Shade, A. (2019). 16S rRNA/rRNA gene ratios and cell activity staining reveal consistent patterns of microbial activity in plant-associated soil. *mSystems* 4 (2), e00003–e00019. doi: 10.1128/mSystems.00003-19

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36 (6), 1925–1927. doi: 10.1093/bioinformatics/btz848

Dedysh, S. N., Knief, C., and Dunfield, P. F. (2005). Methylocella species are facultatively methanotrophic. *J. Bacteriol* 187 (13), 4665–4670. doi: 10.1128/jb.187.13.4665-4670.2005

de Kreuk, M. K., and van Loosdrecht, M. C. M. (2004). Selection of slow growing organisms as a means for improving aerobic granular sludge stability. *Water Sci. Technol.* 49 (11-12), 9–17. doi: 10.2166/wst.2004.0792

Eren, A. M., Esen Ö, C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3, e1319. doi: 10.7717/peerj.1319

Fang, Y., Karnjanapiboonwong, A., Chase, D. A., Wang, J., Morse, A. N., and Anderson, T. A. (2012). Occurrence, fate, and persistence of gemfibrozil in water and soil. *Environ. Toxicol. Chem.* 31 (3), 550–555. doi: 10.1002/etc.1725

Graham, E. D., Heidelberg, J. F., and Tully, B. J. (2018). Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J.* 12 (7), 1861–1866. doi: 10.1038/s41396-018-0091-3

Grandclément, C., Piram, A., Petit, M.-E., Seyssiecq, I., Laffont-Schwob, I., Vanot, G., et al. (2020). Biological Removal and Fate Assessment of Diclofenac Using Bacillus subtilis and Brevibacillus laterosporus Strains and Ecotoxicological Effects of Diclofenac and 4′-Hydroxy-diclofenac. *J. Chem.* 2020, 9789420. doi: 10.1155/2020/9789420

Jenkins, O., Byrom, D., and Jones, D. (1987). Methylophilus: a new genus of methanol-utilizing bacteria. *Int. J. Systematic Evolutionary Microbiol.* 37 (4), 446–448. doi: 10.1099/00207713-37-4-446

Jewell, K. S., Falås, P., Wick, A., Joss, A., and Ternes, T. A. (2016). Transformation of diclofenac in hybrid biofilm–activated sludge processes. *Water Res.* 105, 559–567. doi: 10.1016/j.watres.2016.08.002

Jiang, Y., Shi, X., and Ng, H. Y. (2021). Aerobic granular sludge systems for treating hypersaline pharmaceutical wastewater: Start-up, long-term performances and metabolic function. *J. Hazardous Materials* 412, 125229. doi: 10.1016/j.jhazmat.2021.125229

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., et al. (2017). Where less may be more: how the rare biosphere pulls ecosystems strings. *ISME J.* 11 (4), 853–862. doi: 10.1038/ismej.2016.174

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7, e7359. doi: 10.7717/peerj.7359

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30 (4), 772–780. doi: 10.1093/molbev/mst010

Keener, W. K., and Arp, D. J. (1994). Transformations of aromatic compounds by nitrosomonas europaea. *Appl. Environ. Microbiol.* 60 (6), 1914–1920. doi: 10.1128/aem.60.6.1914-1920.1994

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26 (11), 1463–1464. doi: 10.1093/bioinformatics/btq166

Kent, J., and Tay, J. H. (2019). Treatment of 17α–ethinylestradiol, 4–nonylphenol, and carbamazepine in wastewater using an aerobic granular sludge sequencing batch reactor. *Sci. Total Environ.* 652, 1270–1278. doi: 10.1016/j.scitotenv.2018.10.301

Kim, S., and Aga, D. S. (2007). Potential ecological and human health impacts of antibiotics and antibiotic-resistant bacteria from wastewater treatment plants. *J. Toxicol. Environ. Health Part B* 10 (8), 559–573. doi: 10.1080/15287390600975137

Kjeldal, H., Zhou, N. A., Wissenbach, D. K., von Bergen, M., Gough, H. L., and Nielsen, J. L. (2016). Genomic, proteomic, and metabolite characterization of gemfibrozil-degrading organism bacillus sp. GeD10. *Environ. Sci. Technol.* 50 (2), 744–755. doi: 10.1021/acs.est.5b05003

Kong, Q., Wang, Z.-b., Shu, L., and Miao, M.-s. (2015). Characterization of the extracellular polymeric substances and microbial community of aerobic granulation sludge exposed to cefalexin. *Int. Biodeterioration Biodegradation* 102, 375–382. doi: 10.1016/j.ibiod.2015.04.020

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9 (4), 357–359. doi: 10.1038/nmeth.1923

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi: 10.1093/bioinformatics/btp352

Limpiyakorn, T., Fürhacker, M., Haberl, R., Chodanon, T., Srithep, P., and Sonthiphand, P. (2013). amoA-encoding archaea in wastewater treatment plants: a review. *Appl. Microbiol. Biotechnol.* 97 (4), 1425–1439. doi: 10.1007/s00253-012-4650-7

Llorca, M., Rodríguez-Mozaz, S., Couillerot, O., Panigoni, K., de Gunzburg, J., Bayer, S., et al. (2015). Identification of new transformation products during enzymatic treatment of tetracycline and erythromycin antibiotics at laboratory scale by an on-line turbulent flow liquid-chromatography coupled to a high resolution mass spectrometer LTQ-Orbitrap. *Chemosphere* 119, 90–98. doi: 10.1016/j.chemosphere.2014.05.072

Margot, J., Lochmatter, S., Barry, D. A., and Holliger, C. (2015). Role of ammonia-oxidizing bacteria in micropollutant removal from wastewater with aerobic granular sludge. *Water Sci. Technol.* 73 (3), 564–575. doi: 10.2166/wst.2015.514

Martin, T. M., Harten, P., Venkatapathy, R., and Young, D. (2020). *Toxicity estimation software tool. 5.1.2 ed* (Cincinnati: US EPA).

Martin, J., Santos, J. L., Aparicio, I., and Alonso, E. (2010). Multi-residue method for the analysis of pharmaceutical compounds in sewage sludge, compost and sediments by sonication-assisted extraction and LC determination. *J. Sep Sci.* 33 (12), 1760–1766. doi: 10.1002/jssc.200900873

Matsen, F. A., Kodner, R. B., and Armbrust, E. V. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinf.* 11 (1), 538. doi: 10.1186/1471-2105-11-538

McIlroy, S. J., Albertsen, M., Andresen, E. K., Saunders, A. M., Kristiansen, R., Stokholm-Bjerregaard, M., et al. (2014). 'Candidatus Competibacter'-lineage genomes retrieved from metagenomes reveal functional metabolic diversity. *Isme J.* 8 (3), 613–624. doi: 10.1038/ismej.2013.162

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8 (4), e61217. doi: 10.1371/journal.pone.0061217

Mustakhimov, I., Kalyuzhnaya, M. G., Lidstrom, M. E., and Chistoserdova, L. (2013). Insights into Denitrification in Methylotenera mobilis from Denitrification Pathway and Methanol Metabolism Mutants. *J. Bacteriology* 195 (10), 2207–2211. doi: 10.1128/JB.00069-13

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27 (5), 824–834. doi: 10.1101/gr.213959.116

Oksanen, J. (2010) *Vegan : community ecology package*. Available at: http://vegan.r-forge.r-project.org/.

Parks, D. H., ChuvoChina, M., Rinke, C., Mussig, A. J., Chaumeil, P.-A., and Hugenholtz, P. (2021). GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* 50 (D1), D785–D794. doi: 10.1093/nar/gkab776

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25 (7), 1043–1055. doi: 10.1101/gr.186072.114

Pawar, H. V., Tetteh, J., Debrah, P., and Boateng, J. S. (2019). Comparison of in *vitro* antibacterial activity of streptomycin-diclofenac loaded composite biomaterial dressings with commercial silver based antimicrobial wound dressings. *Int. J. Biol. Macromolecules* 121, 191–199. doi: 10.1016/j.ijbiomac.2018.10.023

Paxéus, N. (2004). Removal of selected non-steroidal anti-inflammatory drugs (NSAIDs), gemfibrozil, carbamazepine, b-blockers, trimethoprim and triclosan in conventional wastewater treatment plants in five EU countries and their discharge to the aquatic environment. *Water Sci. Technol.* 50 (5), 253–260. doi: 10.2166/wst.2004.0335

Pluskal, T., Castillo, S., Villar-Briones, A., and Oresic, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinf.* 11, 395. doi: 10.1186/1471-2105-11-395

Ren, J., Deng, L., Niu, D., Wang, Z., Fan, B., Taoli, H., et al. (2021). Isolation and identification of a novel erythromycin-degrading fungus, Curvularia sp. RJJ-5, and its degradation pathway. *FEMS Microbiol. Lett.* 368 (1). doi: 10.1093/femsle/fnaa215

Ren, J., Ni, S., Shen, Y., Niu, D., Sun, R., Wang, C., et al. (2022). Characterization of the erythromycin degradation pathway and related enzyme in Rhodococcus gordoniae rjjtx-2. *J. Cleaner Production* 379, 134758. doi: 10.1016/j.jclepro.2022.134758

Rodriguez-Sanchez, A., Margareto, A., Robledo-Mahon, T., Aranda, E., Diaz-Cruz, S., Gonzalez-Lopez, J., et al. (2017). Performance and bacterial community structure of a granular autotrophic nitrogen removal bioreactor amended with high antibiotic concentrations. *Chem. Eng. J.* 325, 257–269. doi: 10.1016/j.cej.2017.05.078

RStudioTeam (2009) *a language and environment for statistical computing*. Available at: http://www.R-project.org.

Sathishkumar, P., Meena, R. A. A., Palanisami, T., Ashokkumar, V., Palvannan, T., and Gu, F. L. (2020). Occurrence, interactive effects and ecological risk of diclofenac in environmental compartments and biota - a review. *Sci. Total Environ.* 698, 134057. doi: 10.1016/j.scitotenv.2019.134057

Schafhauser, B. H., Kristofco, L. A., de Oliveira, C. M. R., and Brooks, B. W. (2018). Global review and analysis of erythromycin in the environment: Occurrence, bioaccumulation and antibiotic resistance hazards. *Environ. pollut.* 238, 440–451. doi: 10.1016/j.envpol.2018.03.052

Schatteman, A., Wright, C. L., Crombie, A. T., Murrell, J. C., and Lehtovirta-Morley, L. E. (2022). Hydrazines as substrates and inhibitors of the archaeal ammonia oxidation pathway. *Appl. Environ. Microbiol.* 88 (8), e0247021. doi: 10.1128/aem.02470-21

Singleton, C. M., Petriglieri, F., Kristensen, J. M., Kirkegaard, R. H., Michaelsen, T. Y., Andersen, M. H., et al. (2021). Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing. *Nat. Commun.* 12 (1) 2009. doi: 10.1038/s41467-021-22203-2

Sodré, F. F., and Sampaio, T. R. (2020). Development and application of a SPE-LC-QTOF method for the quantification of micropollutants of emerging concern in drinking waters from the Brazilian capital. *Emerging Contaminants* 6, 72–81. doi: 10.1016/j.emcon.2020.01.001

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30 (9), 1312–1313. doi: 10.1093/bioinformatics/btu033

Suarez, C., Sedlacek, C. J., Gustavsson, D. J. I., Eiler, A., Modin, O., Hermansson, M., et al. (2022). Disturbance-based management of ecosystem services and disservices in partial nitritation-anammox biofilms. *NPJ Biofilms Microbiomes* 8 (1), 47. doi: 10.1038/s41522-022-00308-w

Syed, M., Skonberg, C., and Hansen, S. H. (2016). Mitochondrial toxicity of diclofenac and its metabolites via inhibition of oxidative phosphorylation (ATP synthesis) in rat liver mitochondria: Possible role in drug induced liver injury (DILI). *Toxicol. Vitro* 31, 93–102. doi: 10.1016/j.tiv.2015.11.020

Wang, X. C., Chen, Z. L., Kang, J., Zhao, X., and Shen, J. M. (2018). Removal of tetracycline by aerobic granular sludge and its bacterial community dynamics in SBR. *RSC Adv.* 8 (33), 18284–18293. doi: 10.1039/c8ra01357h

Xia, K., Bhandari, A., Das, K., and Pillar, G. (2005). Occurrence and fate of pharmaceuticals and personal care products (PPCPs) in biosolids. *J. Environ. Qual.* 34 (1), 91–104. doi: 10.2134/jeq2005.0091

Zhao, X., Chen, Z., Wang, X., Li, J., Shen, J., and Xu, H. (2015). Remediation of pharmaceuticals and personal care products using an aerobic granular sludge sequencing bioreactor and microbial community profiling using Solexa sequencing technology analysis. *Bioresource Technol.* 179, 104–112. doi: 10.1016/j.biortech.2014.12.002

Zurita, J. L., Repetto, G., Jos, A., Salguero, M., López-Artíguez, M., and Cameán, A. M. (2007). Toxicological effects of the lipid regulator gemfibrozil in four aquatic systems. *Aquat Toxicol.* 81 (1), 106–115. doi: 10.1016/j.aquatox.2006.11.007

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 27 September 2023
DOI 10.3389/frmbi.2023.1094865

OPEN ACCESS

EDITED BY
Kian Mau Goh,
University of Technology Malaysia, Malaysia

REVIEWED BY
Paltu Kumar Dhal,
Jadavpur University, India
Tianjie Yang,
Nanjing Agricultural University, China
Amira Suriaty Yaakop,
University of Science Malaysia (USM), Malaysia

*CORRESPONDENCE
Emma R. Master
emma.master@utoronto.ca

RECEIVED 10 November 2022
ACCEPTED 01 September 2023
PUBLISHED 27 September 2023

CITATION
Wong MT, Nesbø CL, Wang W, Couturier M, Lombard V, Lapebie P, Terrapon N, Henrissat B, Edwards EA and Master ER (2023) Taxonomic composition and carbohydrate-active enzyme content in microbial enrichments from pulp mill anaerobic granules after cultivation on lignocellulosic substrates.
Front. Microbiomes 2:1094865.
doi: 10.3389/frmbi.2023.1094865

Taxonomic composition and carbohydrate-active enzyme content in microbial enrichments from pulp mill anaerobic granules after cultivation on lignocellulosic substrates

Mabel T. Wong[1], Camilla L. Nesbø[1], Weijun Wang[1], Marie Couturier[2], Vincent Lombard[3,4], Pascal Lapebie[3], Nicolas Terrapon[3,4], Bernard Henrissat[3,4,5,6], Elizabeth A. Edwards[1] and Emma R. Master[1,7]*

[1]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada, [2]Université Grenoble Alpes, Centre National de la Recherche Scientifique (CNRS), Centre de Recherches sur les Macromolécules Végétales (CERMAV), Grenoble, France, [3]Aix-Marseille Université, Centre National de la Recherche Scientifique (CNRS), UMR7257, Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB), Marseille, France, [4]L'Institut National de Recherche our l'Agriculture, l'Alimentation et l'Environnement (INRAE), USC1408 Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB), Marseille, France, [5]Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, [6]Technical University of Denmark, Danmarks Tekniske Universitet (DTU) Bioengineering, Lyngby, Denmark, [7]Department of Bioproducts and Biosystems, Aalto University, Espoo, Finland

Metagenomes of lignocellulose-degrading microbial communities are reservoirs of carbohydrate-active enzymes relevant to biomass processing. Whereas several metagenomes of natural digestive systems have been sequenced, the current study analyses metagenomes originating from an industrial anaerobic digester that processes effluent from a cellulose pulp mill. Both 16S ribosomal DNA and metagenome sequences were obtained following anaerobic cultivation of the digester inoculum on cellulose and pretreated (steam exploded) poplar wood chips. The community composition and profile of predicted carbohydrate-active enzymes were then analyzed in detail. Recognized lignocellulose degraders were abundant in the resulting cultures, including populations belonging to Clostridiales and Bacteroidales orders. Poorly defined taxonomic lineages previously identified in other lignocellulose-degrading communities were also detected, including the uncultivated Firmicutes lineage *OPB54* which represented nearly 10% of the cellulose-fed enrichment even though it was not detected in the bioreactor inoculum. In total, 3580 genes encoding carbohydrate-active enzymes were identified through metagenome sequencing. Similar to earlier enrichments of animal digestive systems, the profile encoded by the bioreactor inoculum following enrichment on pretreated wood was distinguished from the cellulose counterpart by a higher

occurrence of enzymes predicted to act on pectin. The majority (> 93%) of carbohydrate-active enzymes predicted to act on plant polysaccharides were identified in the metagenome assembled genomes, permitting taxonomic assignment. The taxonomic assignment revealed that only a small selection of organisms directly participates in plant polysaccharide deconstruction and supports the rest of the community.

KEYWORDS

carbohydrate active enzymes, metagenomics, lignocellulose, anaerobe, bioconversion

Introduction

Wood and agricultural fibre comprise cellulose, hemicelluloses and lignins (i.e., lignocellulose) and represent major resources for the production of renewable fuels, chemicals and materials. Biological deconstruction of lignocellulose is catalyzed by the concerted action of carbohydrate-active enzymes (CAZymes) and metagenomes of lignocellulose-degrading communities are an especially rich source of CAZymes that could be used for lignocellulose processing (Terrapon et al., 2018; Garron and Henrissat, 2019; Drula et al., 2022).

Metagenomic studies aimed at identifying new CAZymes have sequenced grass-feeding gut microbiota (Al-Masaudi et al., 2017; Deusch et al., 2017; Wang et al., 2019) and wood degrading gut microbiomes of termites (Warnecke et al., 2007; Liu et al., 2019; Romero Victorica et al., 2020), beetle (Scully et al., 2013), wood wasp (Adams et al., 2011) moose (Svartström et al., 2017; Wong et al., 2017), and beaver (Wong et al., 2017; Armstrong et al., 2018). Most pertinent CAZymes for lignocellulose processing can be uncovered through sequencing metagenomes of enrichment cultures (Jimenéz et al., 2016; Wang et al., 2016; Schultz-Johansen et al., 2018; Tomazetto et al., 2020; Borjigin et al., 2022). For instance, in our previous study of digestive microbiomes from Canadian beaver (*Castor canadensis*) and North American moose (*Alces americanus*), we report substrate-induced convergence of taxonomic profiles and CAZyme compositions following anaerobic cultivation of corresponding inocula on cellulose or pretreated wood fibre.

Large-scale bioreactor systems that process industrial and municipal lignocellulosic materials represent additional, compelling sources of CAZymes (Wilkens et al., 2017). Besides municipal bioreactors, anaerobic digesters that transform pulp mill effluent to biogas are an untapped source of enzymes for lignocellulose processing. In most cases, pulp mill effluent is mechanically processed to remove suspended solids and then biologically treated through an aerobic activated sludge process (Thompson et al., 2001). The generated biosludge is subsequently dewatered and incinerated for power generation or else landfilled. Increasingly, anaerobic bioconversion of pulp mill effluent prior to aerobic treatment is employed to reduce the accumulation of secondary biosludge while also generating biogas (Hagelqvist, 2013; Meyer and Edwards, 2014). About 10% of pulp mills worldwide have installed anaerobic treatment technologies, and in particular, internal circulation reactors (Meyer and Edwards, 2014). These high-rate reactors enable effective bioconversion of organic compounds into biogas via the sequential activity of hydrolytic bacteria, acetogens, and methanogenic archaea (Tauseef et al., 2013; Kamali et al., 2016).

Herein, we investigated whether microbial communities enriched from pulp mill anaerobic bioreactors encode an assemblage of CAZymes distinct from those identified through metagenomic analysis of natural digestive systems. To address this question, microbial granules were collected from an anaerobic internal circulation bioreactor located at a pulp mill and enriched for three years on multiple lignocellulosic carbon sources. Subsequent metagenome sequencing and metagenome assembled genomes (MAGs) permitted CAZyme assignment to specific microbial species originating from the pulp mill bioreactor. The current study underpins the significance of both environmental inoculum and enrichment condition on the taxonomic distribution and functional potential of resulting microbial communities.

Materials and methods

Collection of pulp mill anaerobic granules

Anaerobic granules were collected in 2009 from an internal circulation bioreactor located at a pulp mill in Québec, Canada. The bioreactor typically received 15,000 m^3/day of mixed wastewater including acid condensate from the evaporator system and bleached chemi-thermomechanical pulp effluent.

Set up and maintenance of lignocellulose-degrading enrichment cultures from pulp mill anaerobic granules

The anaerobic granules were used as inoculum for enrichment cultures grown at 36 °C under anaerobic conditions as previously described (Wong et al., 2016). Briefly, sulphide-reduced mineral medium (pH 7.0) was prepared and purged with 80% N_2, 20% CO_2 gas mixture (Wong et al., 2017). Approximately 15 mL of anaerobic

granules were transferred to 160 mL Wheaton glass serum bottles amended with a lignocellulosic substrate (average 36.1 mg chemical oxygen demand (COD) equivalent per bottle) and 45 mL of mineral medium. The lignocellulosic amendments were i) microcrystalline cellulose (Avicel PH101, Sigma-Aldrich, MO, USA), ii) cellulose + lignosulphonate, iii) cellulose + tannic acid (Sigma-Aldrich, MO, USA), and iv) steam-exploded poplar (SunOpta Inc., Canada). Biogas production was regularly monitored using a pressure transducer (Omega PX725 Industrial Pressure Transmitter, Omega DP24-E Process Meter). When biogas production ceased, the microbial community was transferred to a new bottle with fresh anaerobic medium and lignocellulose carbon source (Wong et al., 2016). This process was repeated eight times over a period of 3 years prior to DNA extraction and sequencing.

DNA extraction and sequencing

DNA was extracted from each culture after three years of enrichment on respective substates. Samples (10 mL) were taken just as biogas production began to slow down and centrifuged at 15,000 x g for 15 min at 4 °C. Total community DNA was then extracted from the resulting pellets using the QIAamp DNA Stool Mini Kit. The concentration and quality of the extracted DNA was assessed by measuring the 260/280 absorbance ratio using a Nanodrop 2000 spectrophotometer (Thermo Scientific, MA, United States) before storing the DNA samples at -80 °C. The DNA samples were used for 16S rRNA gene amplicon sequencing and for metagenomic shotgun sequencing (Wong et al., 2016). The V6-8 hypervariable region of 16S rRNA genes was amplified with 926 Forward (5'- AAACTYAAAKGAATTGACGG) and 1392 Reverse (5'- ACGGGCGGTGTGTRC) primers and multiplexed with 10-nt Roche barcodes by polymerase chain reaction (Table S1; DeAngelis et al., 2012). Multiplex-pyrosequencing was performed in 2013 on a 454 GS FLX platform (454 Life Sciences-a Roche Company, Branford, CT, USA) at the Génome Québec Innovation Centre. Illumina paired-end sequencing with TruSeq library was performed in 2015 using a Illumina HiSeq 2000 (Illumina Inc., San Diego, CA, USA) at the Génome Québec Innovation Centre.

Analyses of 16S rRNA gene pyrosequences

Pyrosequencing output was converted to sequence reads and quality scores using Roche 454 Life Science propriety software (http://www.454.com) and then analyzed by QIIME 1.8.0 (Caporaso et al., 2010) as previously described (Wong et al., 2016). Combining the taxonomic profiles from our previous analysis on lignocellulose-degrading microcosms from beaver dropping and moose rumen (Wong et al., 2017), relative abundances of genera representing ≥ 1% for at least one of the samples were extracted for hierarchical clustering (correlation clustering and average linkage) and Principal Component Analysis using R statistics in ClustVis (Metsalu and Vilo, 2015). Non-parametric Kruskal-Wallis test was conducted using R script (R Development Core Team, 2010).

Metagenome assembly, annotation of CAZyme families, multi-modular sequences, and polysaccharide utilization loci

Quality trimming and assembly of metagenomics shotgun sequences were first done using Abyss v 1.3 (Simpson et al., 2009) as described in Wong et al., 2017. In addition, the raw reads (70,817,105 pairs for cellulose-fed enrichments, and 78,824,461 pairs for pretreated poplar-fed enrichments) were quality trimmed and assembled using the Anvi'o v6.2 snakemake metagenomic workflow (Eren et al., 2021). This pipeline uses the read quality control methods developed in Minoche et al., 2011. Three assemblers were used; megahit v.1.1.2 (Li et al., 2015), spades v. 3.12.0 (metaspades mode, Nurk et al., 2017), idba v. 1.1.3 (Peng et al., 2012), and the assemblies were performed both on the samples separately and combined in co-assemblies. The assemblies were binned into metagenome assembled genomes (MAGs) using metabat2 (Kang et al., 2019) and maxbin2 (Wu et al., 2016), and the resulting MAGs were compared and dereplicated using dRep (Olm et al., 2017) with a pairwise ANI cut-off at 99%, % completion > 75% and contamination < 25%. The resulting collection of MAGs were combined into one dataset; reads were mapped back to the MAGs using bowtie2 v 2.4.2 to obtain coverage information and visualized in Anvi'o. Taxonomic classification was assigned to each MAG using the GTDB-TK tool kit which compares the genomes to the genome taxonomy database (Chaumeil et al., 2020). Taxonomic composition of metagenomic reads based on rRNA reads extracted from the libraries and classified to order level was performed using the PhyloFlash v. 3.4 (Gruber-Vodicka et al., 2020) software and the SILVA v.138 database.

Prediction of open reading frames, CAZyme families, PULs as well as taxonomic assignment of the predicted CAZymes were performed on the Abyss-assembly and visualized in accordance to the methodology described in Wong et al., 2017. Count of identified sequences in a given CAZyme family was normalized by the number of open reading frames. Normalized count of predicted plant polysaccharide-active CAZyme families from the earlier beaver dropping and moose rumen metagenomes (Wong et al., 2017) were combined with the current dataset for hierarchical clustering (correlation clustering and average linkage) and PCA using R statistics in ClustVis (Metsalu and Vilo, 2015). The identified genes encoding CAZymes were mapped to the MAGs using BLASTN.

Results and discussion

Establishment of biogas-producing microbial enrichments

Anaerobic enrichment cultures were established and biogas production was sustained over three years on four lignocellulosic carbon sources. As observed previously (Wong et al., 2016), biogas yield per COD added dropped during early stages of enrichment, likely due to depletion of readily digestible COD present in the inoculum (Figure 1). The subsequent increase in biogas yield was consistent with microbial acclimatization to the amended carbon sources. Following the three year acclimatization period (i.e., by the

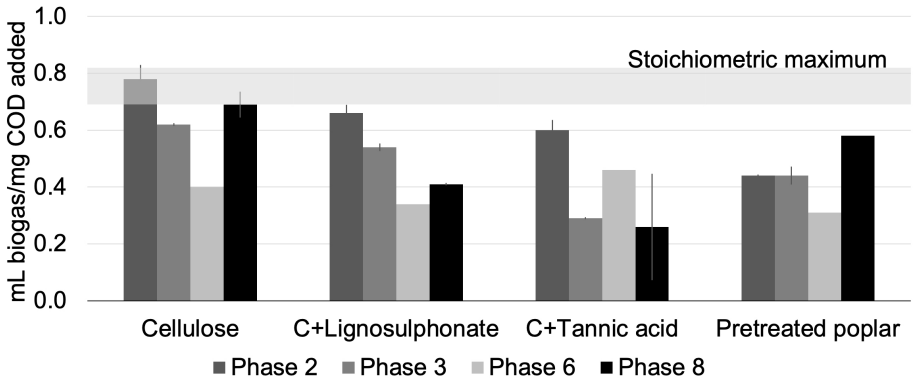


FIGURE 1
Biogas production by microcosms fed with lignocellulosic carbon sources over 3-years. DNA was extracted from Phase 8 for this study. The estimated stoichiometric maximum biogas yield is shown by the horizontal grey band to provide a reference for the extent of conversion of each substrate (Symons and Buswell, 1933; Wong et al., 2016); error bars indicate standard deviation; n = 3.

eighth transfer), biogas yields were highest for cellulose-fed cultures (0.69 mL biogas/mg COD added), followed by pretreated poplar (0.58 mL biogas/mg COD added), cellulose + lignosulphonate (0.41 mL biogas/mg COD added), and cellulose + tannic acid (0.15 mL biogas/mg COD added). The apparent inhibition of biogas production in cultures amended with tannic acid was also observed in enrichment cultures from beaver droppings and moose rumen (Wong et al., 2016).

Microbial diversity in lignocellulose-degrading microcosms enriched from pulp mill anaerobic granules

A total of 99,218 high-quality 16S rRNA gene pyrosequences from the established anaerobic enrichments were assigned to 5,359 OTUs at 97% similarity threshold (Table S2A). Beta-diversity based UPGMA (unweighted pair group method with arithmetic mean) clustering of the microbial communities revealed two main clades: i) microcosms fed with cellulose or cellulose + tannic acid, and ii) inoculum and microcosms fed with cellulose + lignosulphonate and pretreated poplar (Figure 2). The clustering of biological replicates for a given enrichment further reflected the statistical differences between the communities following the amendments, as well as the reproducibility of enrichments and 16S rRNA gene analysis. The following paragraph highlights the relative abundances of main microorganisms in each culture to uncover shifts in populations that correlate with the amended substrate.

Prior to enrichment, pulp mill anaerobic granules were dominated by *Clostridiales* (~ 40%), *Bacteroidales* (10%), *Synergistales* (6%), *Anaerolineales* (6%), *Syntrophobacterales* (5%) bacterial orders, as well as methanogens belonging to *Methanosarcinales* order (5%) (Figure 2; Table S2B). The species richness of the microbial communities decreased upon anaerobic enrichment on cellulose, and to a lesser extent, pretreated poplar (Figure S1). Nevertheless, lineages recognized as lignocellulose degraders were consistently abundant in all enrichment cultures. For example, populations belonging to *Clostridiales* and

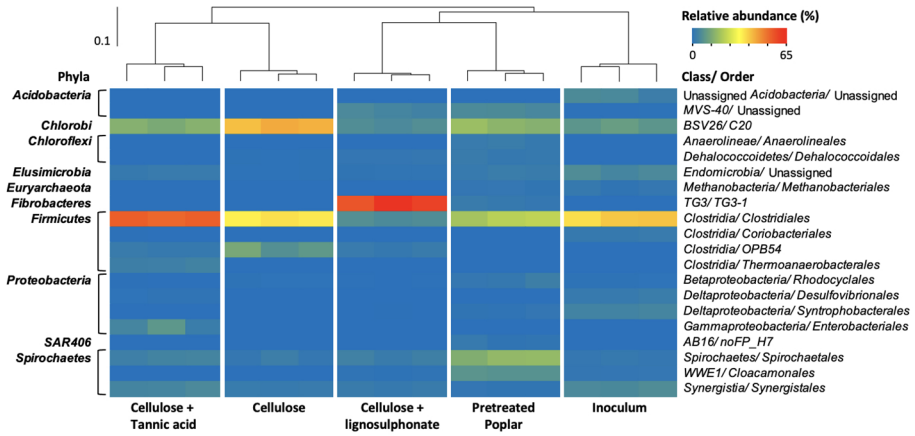


FIGURE 2
Heatmap of the microbial orders (≥ 1% in at least one sample) identified by 16S rRNA analysis of pulp mill anaerobic granules and corresponding enrichment cultures, with the sample similarities calculated by weighted UniFrac method and summarized by UPGMA clustering.

Bacteroidales orders represented more than 75% of the microbial community in the cellulose-fed enrichments and remained ~ 40% in those fed with pretreated poplar (Table S2B). Moreover, the uncultivated Firmicutes lineage *OPB54* represented up to 10% of the cellulose-fed enrichment and ~ 3% in all other enrichments even though it was not detected in the anaerobic granule inoculum (Table S2B). Members of *OPB54* are reported in diverse ecologies including the gut microbiomes of beaver and moose (Wong et al., 2017) and can ferment a wide variety of carbohydrates (Liu et al., 2014). Similarly, the 'termite group 3' (*TG3*) class of the *Fibrobacterota* phylum was abundant (60%) in cultures enriched on cellulose + lignosulphonate even though it was not detected in the inoculum. As its name implies, the *TG3* class was initially detected in termite guts (Hongoh et al., 2006) and later identified in diverse habitats, including anaerobic digesters (Rahman et al., 2016).

Metagenome sequencing and assembly

The cellulose-fed and pretreated poplar-fed enrichments exhibited highest biogas yields compared to the other enrichments generated herein and so corresponding metagenomic DNA was collected for sequencing and assembly. The metagenomic DNA from cellulose-fed and pretreated poplar-fed enrichments yielded 64 and 68 million high quality reads, respectively. The taxonomic composition of the metagenomes was assessed by extracting rRNA genes and comparing them to the SILVA database. The taxonomic composition at order level was similar to the community structure obtained from the amplicon data (Figure S2), with *Clostridiales* and *Bacteroidales* representing > 75% of the microbial community in the cellulose-fed enrichments and ~30% in those fed with pretreated poplar.

The metagenome reads were assembled using four assemblers; assembly statistics are provided in Table S3. The assembled contigs were binned into 2,417 bins, which were dereplicated using a 99% identity cut-off to 139 unique metagenome assembled genomes (MAGs) with completeness > 75% and contamination < 25% as determined by CheckM (Parks et al., 2015; Table S4). Of these, 127 MAGs had at least medium quality with less than 10% contamination (Bowers et al., 2017). Moreover, 83% of reads from cellulose-fed enrichments and 75% of reads from pretreated poplar-fed enrichments were mapped to the MAGs.

Taxonomy and abundances of each MAG are shown in Figure 3 and Table S4. Most abundant MAGs obtained from the cellulose-fed enrichment were classified as belonging to *Parabacteroides* (TG_idbaud_maxbin.144), *Herbinix* (TGCmsp_mb_bin.66) and *Acetivibrionaceae* (TGPidbaud_mb_bin_32). By comparison, the community structure was more even in the pretreated poplar-fed enrichment, and most abundant MAGs were classified to *Treponemataceae* Spiro-10 (TGP_msp_maxbin.003), *Bacteroidota* UBA10030 (TGP_msp_maxbin.004) and *Cloacimonas* sp. 002432865 (TGP_mgh_maxbin.005). One abundant MAG shared by the two metagenomes at the 99% identity was the *Acetivibrionaceae* MAG TGPidbaud_mb_bin_32. However, inspection of reads mapped to this MAG revealed that each enrichment contains a distinct, but highly similar strain of this taxon.

CAZyme family profiles of the cellulose- and poplar-fed microcosms

In total, 3,580 genes encoding 95 families of glycoside hydrolases (GHs), 12 families of carbohydrate esterases (CEs), 13 families of polysaccharide lyases (PLs), 32 families of glycosyltransferases (GTs) and 33 families of carbohydrate binding modules (CBMs) were predicted from the two metagenomes (Table S5). This corresponded to approximately 3% of open reading frames predicted in the cellulose-fed and pretreated poplar-fed enrichments. No auxiliary redox enzymes were identified as expected for anaerobic microbial communities.

Among the predicted CAZymes, 245 genes from the cellulose-fed enrichment and 555 from the pretreated poplar-fed enrichment were predicted to be involved in plant polysaccharide deconstruction (Figure 4). Similar to cellulose-fed and pretreated poplar-fed enrichments of moose rumen samples and beaver droppings (Wong et al, 2017), nearly 30% of these CAZymes belonged to families GH2, GH3, GH5, GH9, GH43, CE1, and CE4 (Figure 4A). Thus, a similar CAZyme profile emerged from industrial and natural digestive systems following their enrichment on the same cellulosic and pretreated-poplar substrates. Closer inspection of the CAZyme profiles reported herein, however, distinguished the metagenome of the pretreated poplar-fed enrichment from the cellulose counterpart by its higher occurrence of CAZyme families predicted to act on pectin, including more than nine times the number of sequences belonging to family PL1 (pectate lyases), and more than twice the number of sequences belonging to families CE8 (pectin methylesterase), GH28 (polygalacturonases) and GH105 (unsaturated glucuronyl/galacturonyl hydrolases) (Figure 4B).

The majority (> 93%) of the predicted CAZymes acting on plant polysaccharides were included in the MAGs constructed from the metagenomes (Figure 3, Table S6). MAGs from cellulose-fed enrichments that carry the highest number of predicted CAZymes targeting plant polysaccharides were also the most abundant in the metagenome; however, this was not the case for pretreated poplar-fed enrichments (Tables S6, S7). For example, the most abundant MAG in the cellulose-fed enrichment is classified as *Acetivibrionaceae* (TGPidbaud_mb_bin_32) and carries 79 of the 245 genes from the corresponding metagenome that are predicted to encode CAZymes acting on plant polysaccharides. The MAG belonging to *Acetivibrionaceae* from the pretreated poplar-fed enrichment also encoded the highest number of CAZymes predicted to deconstruct plant polysaccharides; however, this MAG was low in abundance (Tables S6, S8). Instead, the most abundant MAGs in the pretreated poplar-fed enrichment were the *Bacteroidota* UBA10030 MAG (TGP_msp_maxbin.004) that carried only 61 CAZymes predicted to act on plant polysaccharides, *Treponemataceae* Spiro-10 MAG (TGP_msp_maxbin.003) that carried only 17 and the *Cloacimonas* sp. 002432865 MAG (TGP_mgh_maxbin.005) that carried none.

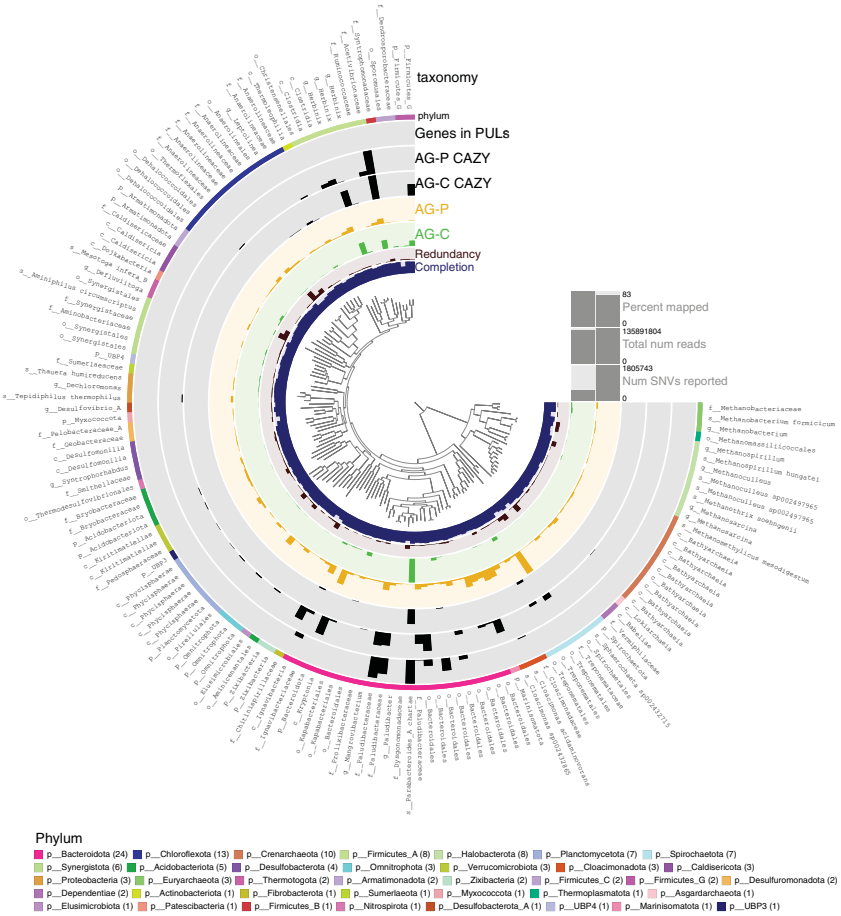


FIGURE 3
Metagenome Assembled Genomes (MAGs) from anaerobic granules (AG) following enrichment on cellulose (AG-C) or pretreated poplar (AG-P). The 139 dereplicated MAGs from the AG-C and AG-P metagenomes are visualized using the Anvi'o software (Eren et al., 2021). The phylogenetic tree at the centre of the figure is a maximum likelihood tree based on concatenated ribosomal proteins calculated using FastTree (Price et al., 2010). From inner to outer ring: The first ring shows the % completion of each MAG calculated based on single-copy core genes (SCGs) with maximum set to 100%. The second ring indicates the level of redundancy (or possible contamination) also based on SCGs with maximum set to 30%. The green and yellow bar for each MAG shows the mean coverage of each MAG in each sample, with maximum at 1284 and 587 for AG-C and AG-P, respectively. The AG-C CAZY and AG-P CAZY bar shows number of blast matches to carbohydrate-active enzyme sequences predicted to act on plant-polysaccharides, where the maximum number in each case is 79 and 80, respectively (pairwise sequence identity was 99%-100%). Genes in PULSs corresponds to the number of CAZymes encoded in MAGs that are predicted to act on plant polysaccharides and be localized in polysaccharide utilization loci (PULs). The outer circle shows the taxonomic domain by color, predicted by Anvio using the genome taxonomy database (gtdb, https://gtdb.ecogenomic.org/). Finally, the best taxonomic classification of each MAG is provided. A 'p' in front of the taxon name indicates this is a phylum classification, a 'c' indicates classification to class, 'f' indicates family level classification, 'o' indicates order level classification, 'g' indicates classification to genus and 's' indicates species level classification. The bars on the right-side shows percent and total number of reads mapped to the MAGs from each metagenome. SNV denotes single nucleotide variants and shows the level of diversity in all the MAGs when reads from the two metagenomes are mapped back to the best MAG.

Notably, the *Bacteroidota* MAG is distinguished by having a comparatively high content of predicted polysaccharide lyases and pectinases belonging to family GH28.

As mentioned in the above section, both *OPB54* and *TG3* species were identified in the enrichment cultures but not the anaerobic granule inoculum, and corresponding OTUs were previously reported in other lignocellulose-degrading communities. Herein, a MAG corresponding to the *OPB54* species was identified in the cellulose-fed enrichment (TGC_mgh_maxbin.005), which is assigned to the family UBA8346 that forms a novel Firmicutes lineage. The *OPB54* MAG encodes nearly 40 CAZyme sequences predicted to act on plant polysaccharides, and compared to most abundant MAGs in the cellulose-fed metagenome, it comprised a high number of predicted pectinolytic enzymes. Specifically, the *OPB54* MAG uniquely encodes sequences belonging to families PL22 (oligogalacturonan lyases) and PL26 (rhamnogalacturonan lyases), and 50% of sequences belonging to family GH28. *TG3* was identified in the pretreated poplar-fed metagenome, which assigned *TG3* to the *Fibrobacterota* phylum (TGPmsp_mb_bin.17). The *TG3* MAG encodes 44 CAZyme sequences predicted to act on plant polysaccharides, with a distribution close to that observed for the corresponding metagenome (Figure 4A, Table S8).

Members of the archaeal phylum *Bathyarchaeota* are found in many anoxic environments and reportedly transform both plant polysaccharides and lignin (Zhou et al., 2018). Several MAGs

10.3389/frmbi.2023.1094865

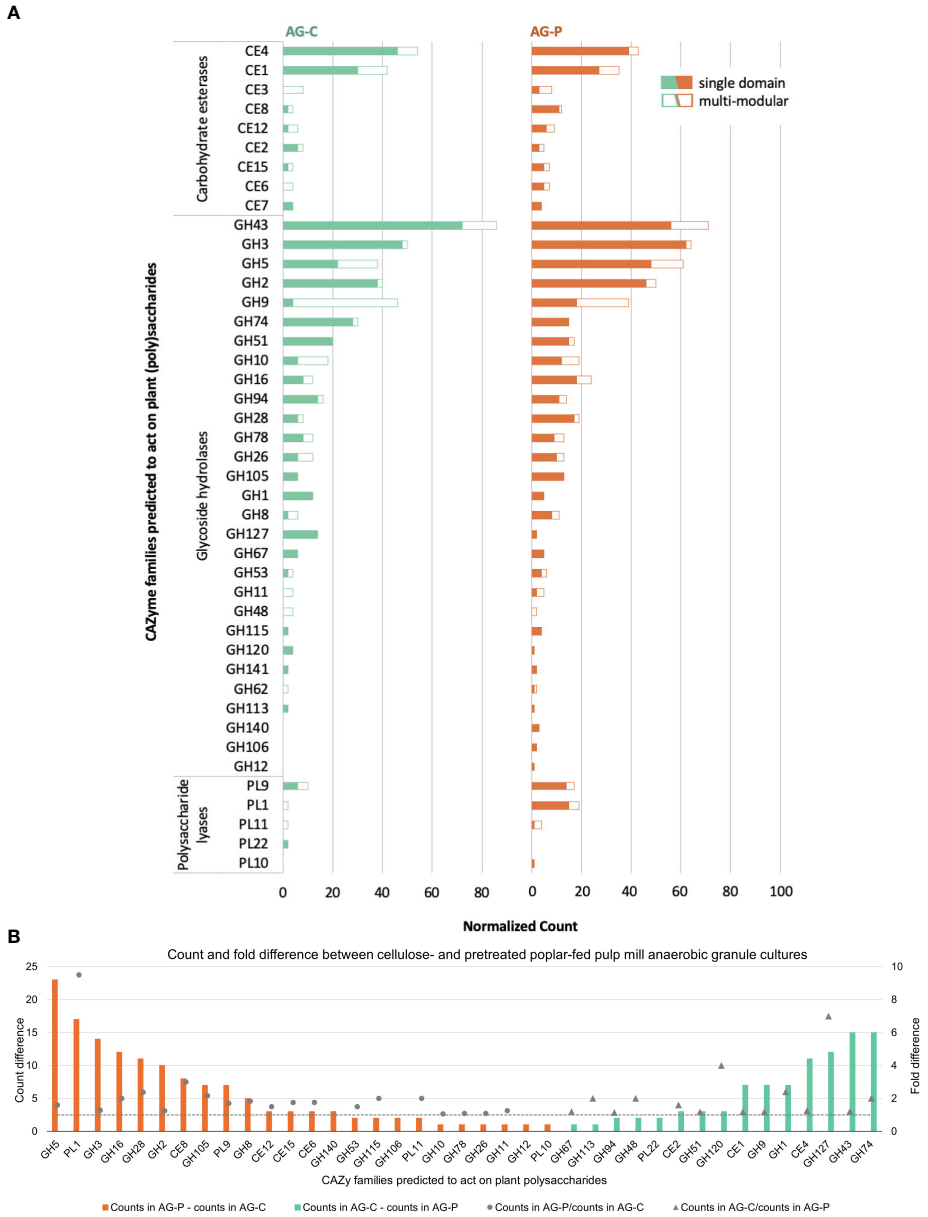


FIGURE 4
(A) Distribution of plant polysaccharide degrading CAZyme families as single and multi-modular domains. **(B)** Normalized count and fold difference of CAZyme families predicted to act on plant polysaccharides between pretreated poplar- and cellulose-fed pulp mill anaerobic granule cultures. Fold difference was only calculated for non-zero counts.

belonging to the *Bathyarchaeota* were identified in the metagenome sequences assembled herein; however, only one MAG (TGPidbaud_mb_bin.66) encoded a CAZyme predicted to act on plant-derived carbohydrates (belonging to family GH2; Tables S6, S8).

Prediction of polysaccharide utilization loci

As a specialized adaptation for polysaccharide degradation, genomes of many gram-negative bacteria within the *Bacteroidetes, Gemmatimonadetes, Ignavibacteriae, Gemmatimon, Balneolaeota* phyla feature PULs that comprise physically-linked genes that encode CAZymes for stepwise binding, hydrolysis and uptake of plant polysaccharides (Terrapon et al., 2018). Of the 800 sequences from cellulose- and pretreated poplar-fed enrichments that were predicted to encode CAZymes targeting plant polysaccharides, over 25% were present in identified PULs. Consistent with the overall distribution of predicted CAZyme sequences, those belonging to families GH3, GH2, and GH43 were most frequently identified in the predicted PULs (Figure 5). The PULs were identified in eight MAGs, where seven were classified as belonging to the phylum *Bacteroidota* and one containing a single two-gene PUL was classified as belonging to *Cloacimonas acidaminovorans* (TGP_mgh_maxbin.006) (Table S6). Whereas three of the PUL-containing MAGs were detected in the cellulose-fed enrichment,

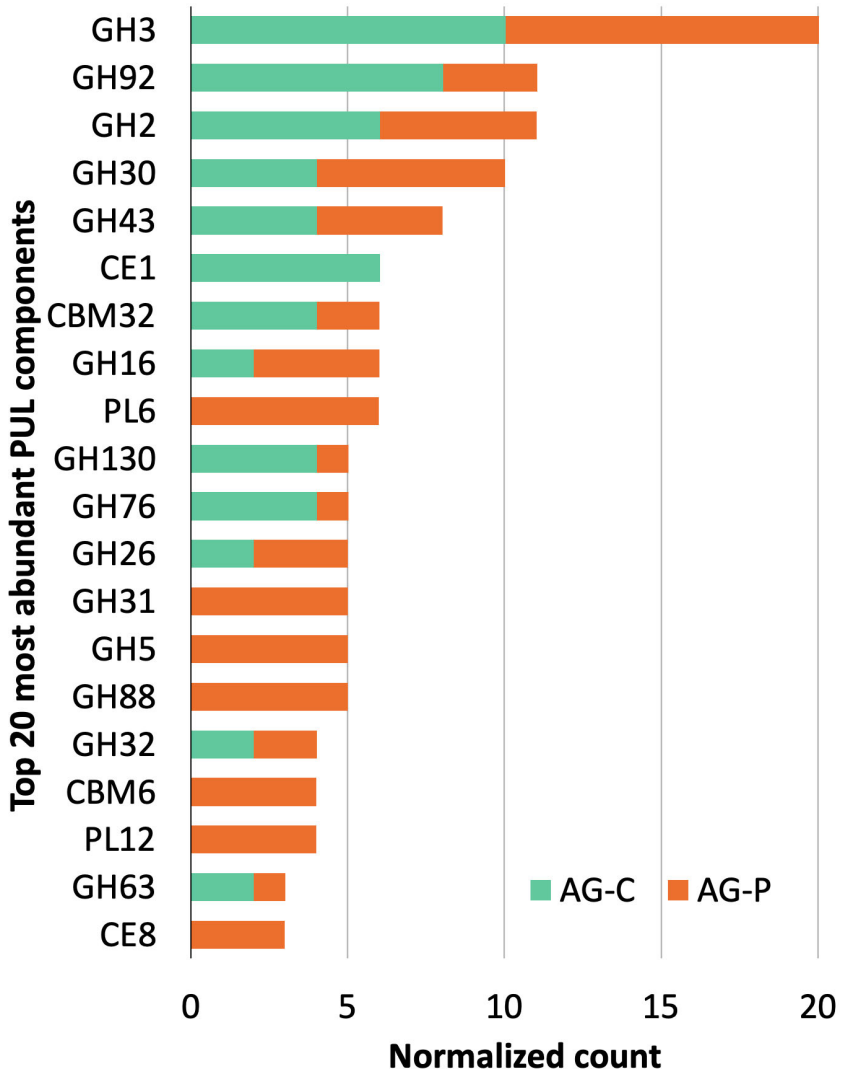


FIGURE 5
Top 20 most abundant CAZyme families identified in predicted PULs from cellulose- and pretreated poplar-fed pulp mill anaerobic granules microcosms.

including the highly abundant *Parabacteroides* MAG (TG_idbaud_maxbin.144), all eight were detected in the pretreated poplar-fed enrichment with the *Paludibacteraceae* MAG (TGP_msp_maxbin.009) and the *Ignavibacteria* MAG (TGP_idbaud_maxbin.007l) being the most enriched (Table S5).

Conclusion

Following cultivation on cellulose and pretreated poplar, microbial communities originating from anaerobic pulp mill digesters encoded a similar profile of predicted plant-polysaccharide active enzymes to communities originating from animal digestive systems (Wong et al., 2017). This observation underscores the considerable impact that substrate amendment has on the functional potential of microbial communities originating from disparate environmental sources. The metagenome assembled genomes of lignocellulose-degrading microbial communities completed herein highlight that OTU abundance does not always identify microorganisms that encode the highest complement of enzymes needed to access the amended lignocellulosic substrate. This was especially apparent for microbial communities enriched on pretreated poplar wood chips. Interestingly, most of the CAZymes predicted to act on plant polysaccharides were found in a small subset of the MAGs constructed. For example, 4 MAGs from the cellulose-fed enrichment encode 219 of the 245 CAZymes predicted to act on plant polysaccharides, while 8 MAGs from the pretreated poplar-fed enrichment encode 403 of the 555 predicted CAZymes. This suggests that a small selection of organisms directly participates in the plant polysaccharide deconstruction and supports the rest of the community. Lastly, the reported MAG analyses shed light on the functional contributions of OTUs that reoccur in lignocellulose-degrading communities (e.g., *TG3* and

OPB54), which can inform the selection of new CAZyme sequences for functional characterization.

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

MW performed the sequence analyses, data interpretation, and compiled the manuscript. CN generated the metagenome assembled genomes and MAG analyses. WW maintained the enrichment cultures, prepared DNA samples for sequencing. MC contributed to data interpretation. BH, NT, VL and PL contributed to the search and annotation of CAZyme modules and PULs, as well as data interpretation. EM and EE conceived and coordinated the study. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the Government of Ontario for the project "Forest FAB: Applied Genomics for Functionalized Fibre and Biochemicals" (ORF-RE-05-005), the Natural Sciences and Engineering Research Council of Canada for the Strategic Network Grant "Industrial Biocatalysis Network", a European Union's Seventh Framework Program (FP/2007/2013)/European Research Council (ERC) Grant Agreement 322820 to BH and an ERC Consolidator Grant to EM (BHIVE – 648925). MW was funded by a University of Toronto Fellowship and Professor William F. Graydon Memorial Graduate Fellowship. MC was funded by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Program.

Acknowledgments

We would like to thank C. Washer and O. Molenda for training on using the Trimmomatic and AbySS assembly programmes, as well as W. Gao and D. Robson for IT support.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1094865/full#supplementary-material

References

Adams, A. S., Jordan, M. S., Adams, S. M., Suen, G., Goodwin, L. A., Davenport, K. W., et al. (2011). Cellulose-degrading bacteria associated with the invasive woodwasp *Sirex noctilio*. *ISME J.* 5, 1323–1331. doi: 10.1038/ismej.2011.14

Al-Masaudi, S., El Kaoutari, A., Drula, E., Al-Mehdar, H., Redwan, E. M., Lombard, V., et al. (2017). A metagenomics investigation of carbohydrate-active enzymes along the gastrointestinal tract of Saudi sheep. *Front. Microbiol.* 8, 666. doi: 10.3389/fmicb.2017.00666

Armstrong, Z., Mewis, K., Liu, F., Morgan-Lang, C., Scofield, M., Durno, E., et al. (2018). Metagenomics reveals functional synergy and novel polysaccharide utilization loci in the Castor canadensis fecal microbiome. *ISME J.* 12, 2757–2769. doi: 10.1038/s41396-018-0215-9

Borjigin, Q., Zhang, B., Yu, X., Gao, J., Zhang, X., Qu, J., et al. (2022). Metagenomics study to compare the taxonomic composition and metabolism of a lignocellulolytic microbial consortium cultured in different carbon conditions. *World J. Microbiol. Biotechnol.* 38 (5), 78. doi: 10.1007/s11274-022-03260-1

Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35 (8), 725–731. doi: 10.1038/nbt.3893

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2020). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848

DeAngelis, K. M., Fortney, J. L., Borglin, S., Silver, W. L., Simmons, B. A., and Hazen, T. C. (2012). Anaerobic decomposition of switchgrass by tropical soil-derived feedstock-adapted consortia. *MBio* 3, e00249. doi: 10.1128/mBio.00249-11

Deusch, S., Camarinha-Silva, A., Conrad, J., Beifuss, U., Rodehutscord, M., and Seifert, J. (2017). A structural and functional elucidation of the rumen microbiome influenced by various diets and microenvironments. *Front. Microbiol.* 8,1605. doi: 10.3389/fmicb.2017.01605

Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 50, D571–D577. doi: 10.1093/nar/gkab1045

Eren, A. M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S. E., Schechter, M. S., et al. (2021). Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* 6 (1), 3–6. doi: 10.1038/s41564-020-00834-3

Garron, M. L., and Henrissat, B. (2019). The continuing expansion of CAZymes and their families. *Curr. Opin. Chem. Biol.* 53, 82–87. doi: 10.1016/j.cbpa.2019.08.004

Gruber-Vodicka, H. R., Seah, B. K. B., and Pruesse, E. (2020). phyloFlash: Rapid SSU rRNA profiling and targeted assembly from metagenomes. *mSystems* 5, e00920–e00920. doi: 10.1128/mSystems.00920-20

Hagelqvist, A. (2013). Batchwise mesophilic anaerobic co-digestion of secondary sludge from pulp and paper industry and municipal sewage sludge. *Waste Manage* 33, 820–824. doi: 10.1016/j.wasman.2012.11.002

Hongoh, Y., Deevong, P., Hattori, S., Inoue, T., Noda, S., Noparatnaraporn, N., et al. (2006). Phylogenetic diversity, localization, and cell morphologies of members of the candidate phylum *TG3* and a subphylum in the phylum *Fibrobacteres*, recently discovered bacterial groups dominant in termite guts. *Appl. Environ. Microbiol.* 72, 6780–6788. doi: 10.1128/AEM.00891-06

Jimenéz, D. J., Brossi, M. J. D., Schuckel, J., Kracun, S. K., Willats, W. G. T., and van Elsas, J. D. (2016). Characterization of three plant biomass-degrading microbial consortia by metagenomics- and metasecretomics-based approaches. *Appl. Microbiol. Biotechnol.* 100, 10463–10477. doi: 10.1007/s00253-016-7713-3

Kamali, M., Gameiro, T., Costa, M. E. V., and Capela, I. (2016). Anaerobic digestion of pulp and paper mill wastes - An overview of the developments and improvement opportunities. *Chem. Eng. J.* 298, 162–182. doi: 10.1016/j.cej.2016.03.119

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7, e7359. doi: 10.7717/peerj.7359

Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31 (10), 1674–1676. doi: 10.1093/bioinformatics/btv033

Liu, N., Li, H. J., Chevrette, M. G., Zhang, L., Cao, L., Zhou, H. K., et al. (2019). Functional metagenomics reveals abundant polysaccharide-degrading gene clusters and cellobiose utilization pathways within gut microbiota of a wood-feeding higher termite. *ISME J.* 13, 104–117. doi: 10.1038/s41396-018-0255-1

Liu, Y., Qiao, J. T., Yuan, X. Z., Guo, R. B., and Qiu, Y. L. (2014). *Hydrogenispora ethanolica* gen. nov., sp. nov., an anaerobic carbohydrate-fermenting bacterium from anaerobic sludge. *Int. J. Syst. Evol. Microbiol.* 64, 1756–1762. doi: 10.1099/ijs.0.060186-0

Metsalu, T., and Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res.* 43, W566–W570. doi: 10.1093/nar/gkv468

Meyer, T., and and Edwards, E. A. (2014). Anaerobic digestion of pulp and paper mill wastewater and sludge. *Water Res.* 65, 321–349. doi: 10.1016/j.watres.2014.07.022

Minoche, A. E., Dohm, J. C., and Himmelbauer, H. (2011). Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biol.* 12 (11), R112. doi: 10.1186/gb-2011-12-11-r112

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27 (5), 824–834. doi: 10.1101/gr.213959.116

Olm, M. R., Brown, C. T., Brooks, B., and and Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* 11, 2864–2868. doi: 10.1038/ismej.2017.126

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25 (7), 1043–1055. doi: 10.1101/gr.186072.114

Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi: 10.1093/bioinformatics/bts174

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS One* 5 (3), e9490. doi: 10.1371/journal.pone.0009490

Rahman, N. A., Parks, D. H., Vanwonterghem, I., Morrison, M., Tyson, G. W., and Hugenholtz, P. (2016). A phylogenomic analysis of the bacterial phylum *Fibrobacteres*. *Front. Microbiol.* 6.

R Development Core Team (2010). *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computer).

Romero Victorica, M., Soria, M. A., Batista-García, R. A., Ceja-Navarro, J. A., Vikram, S., Ortiz, M., et al. (2020). Neotropical termite microbiomes as sources of novel plant cell wall degrading enzymes. *Sci. Rep.* 10 (1), 3864. doi: 10.1038/s41598-020-60850-5

Schultz-Johansen, M., Cueff, M., Hardouin, K., Jam, M., Larocque, R., Glaring, M. A., et al. (2018). Discovery and screening of novel metagenome-derived GH107 enzymes targeting sulfated fucans from brown algae. *FEBS J.* 285, 4281–4295. doi: 10.1111/febs.14662

Scully, E. D., Geib, S. M., Hoover, K., Tien, M., Tringe, S. G., Barry, K. W., et al. (2013). Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. *PloS One* 8, e73827. doi: 10.1371/journal.pone.0073827

Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J. M., and Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome Res.* 19 (6), 1117–1123. doi: 10.1101/gr.089532.108

Svartström, O., Alneberg, J., Terrapon, N., Lombard, V., de Bruijn, I., Malmsten, J., et al. (2017). Ninety-nine *de novo* assembled genomes from the moose (*Alces alces*) rumen microbiome provide new insights into microbial plant biomass degradation. *ISME J.* 11, 2538–2551. doi: 10.1038/ismej.2017.108

Symons, G. E., and Buswell, A. M. (1933). The methane fermentation of carbohydrates. *J. Am. Chem. Soc.* 55, 2028–2036. doi: 10.1021/ja01332a039

Tauseef, S. M., Abbasi, T., and Abbasi, S. A. (2013). Energy recovery from wastewaters with high-rate anaerobic digesters. *Renew Sust Energ Rev.* 19, 704–741. doi: 10.1016/j.rser.2012.11.056

Terrapon, N., Lombard, V., Drula, E., Lapebie, P., Al-Masaudi, S., Gilbert, H. J., et al. (2018). PULDB: the expanded database of Polysaccharide Utilization Loci. *Nucleic Acids Res.* 46, D677–D683. doi: 10.1093/nar/gkx1022

Thompson, G., Swain, J., Kay, M., and Forster, C. F. (2001). The treatment of pulp and paper mill effluent: a review. *Bioresour Technol.* 77, 275–286. doi: 10.1016/S0960-8524(00)00060-2

Tomazetto, G., Pimentel, A. C., Wibberg, D., Dixon, N., and Squina, F. M. (2020). Multi-omic directed discovery of cellulosomes, polysaccharide utilization loci, and lignocellulases from an enriched rumen anaerobic consortium. *Appl. Environ. Microbiol.* 86 (18), e00199–e00120. doi: 10.1128/AEM.00199-20

Wang, C., Dong, D., Wang, H. S., Muller, K., Qin, Y., Wang, H. L., et al. (2016). Metagenomic analysis of microbial consortia enriched from compost: new insights into the role of Actinobacteria in lignocellulose decomposition. *Biotechnol. Biofuels* 9, 22. doi: 10.1186/s13068-016-0440-2

Wang, L. J., Zhang, G. N., Xu, H. J., Xin, H. S., and Zhang, Y. G. (2019). Metagenomic analyses of microbial and carbohydrate-active enzymes in the rumen of Holstein cows fed different forage-to-concentrate ratios. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.00649

Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M., Richardson, T. H., Stege, J. T., et al. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450, 560–U517. doi: 10.1038/nature06269

Wilkens, C., Busk, P. K., Pilgaard, B., Zhang, W. J., Nielsen, K. L., Nielsen, P. H., et al. (2017). Diversity of microbial carbohydrate-active enzymes in Danish anaerobic digesters fed with wastewater treatment sludge. *Biotechnol. Biofuels* 10, 158. doi: 10.1186/s13068-017-0840-y

Wong, M. T., Wang, W. J., Couturier, M., Razeq, F. M., Lombard, V., Lapebie, P., et al. (2017). Comparative metagenomics of cellulose- and poplar hydrolysate-degrading microcosms from gut microflora of the Canadian beaver (*Castor canadensis*) and North American moose (*Alces americanus*) after long-term enrichment. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.02504

Wong, M. T., Wang, W. J., Lacourt, M., Couturier, M., Edwards, E. A., and Master, E. R. (2016). Substrate-driven convergence of the microbial community in lignocellulose-amended enrichments of gut microflora from the Canadian beaver (*Castor canadensis*) and North American moose (*Alces americanus*). *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.00961

Wu, Y.-W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi: 10.1093/bioinformatics/btv638

Zhou, Z., Pan, J., Wang, F., Gu, J.-D., and Li, M. (2018). Bathyarchaeota: globally distributed metabolic generalists in anoxic environments. *FEMS Microbiol. Rev.* 42, 639–655. doi: 10.1093/femsre/fuy023

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 03 October 2023
DOI 10.3389/frmbi.2023.1248384

OPEN ACCESS

EDITED BY
Karin Jacobs,
Stellenbosch University, South Africa

REVIEWED BY
Vadim Lebedev,
Branch of Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russia
Itumeleng Moroenyane,
Stellenbosch University, South Africa

*CORRESPONDENCE
Daoyuan Zhang
zhangdy@ms.xjb.ac.cn
Honglan Yang
yanghonglan@ms.xjb.ac.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 27 June 2023
ACCEPTED 18 September 2023
PUBLISHED 03 October 2023

CITATION
Yang Q, Wang J, Zhang D, Feng H, Bozorov TA, Yang H and Zhang D (2023) Effects of multi-resistant *ScALDH21* transgenic cotton on soil microbial communities. *Front. Microbiomes* 2:1248384. doi: 10.3389/frmbi.2023.1248384

Effects of multi-resistant *ScALDH21* transgenic cotton on soil microbial communities

Qilin Yang[1,2,3†], Jiancheng Wang[1,2,4†], Dawei Zhang[5], Hui Feng[6,7], Tohir A. Bozorov[1,8], Honglan Yang[1,2,4*] and Daoyuan Zhang[1,2,4*]

[1]State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China, [2]Key Laboratory of Ecological Safety and Sustainable Development in Arid Lands, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China, [3]University of Chinese Academy of Sciences, Beijing, China, [4]Turpan Eremophytes Botanical Garden, Chinese Academy of Sciences, Turpan, China, [5]Research Institute of Economic Crops, Xinjiang Academy of Agricultural Sciences, Urumqi, China, [6]State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou, China, [7]Guangdong Provincial Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-sen University, Guangzhou, China, [8]Laboratory of Molecular and Biochemical Genetics, Institute of Genetics and Plants Experimental Biology, Uzbek Academy of Sciences, Tashkent Region, Uzbekistan

Transgenic crops are increasingly prevalent worldwide, and evaluating their impact on soil microbial communities is a critical aspect of upholding environmental safety. Our previous research demonstrated that overexpression of *ScALDH21* from desiccant-tolerant moss, *Syntrichia caninervis*, in cotton revealed multi-resistance to drought, salt, and biotic stresses. We conducted metabarcoding using high-throughput sequencing to evaluate the effect of *ScALDH21* transgenic cotton on soil microbial communities. We further conducted soil tests to analyze the chemical properties of transgenic and non-transgenic cotton, including the total content and availability of chemical elements (K, P, and N), organic matter, and pH value. Both transgenic and non-transgenic cotton fields exhibited soil pH values higher than 8. The presence of transgenic cotton significantly enhanced the availability of available K and the total content of total P in the soil. Alpha and beta diversity indices of soil microbiota showed no difference between two transgenic and non-transgenic cotton groups. Dominant clades of fungal and bacterial genera were equivalent at the phylum and genus levels in all three groups. The correlation analysis of microbial communities and soil environmental factors revealed the absence of significant differences between transgenic and non-transgenic cotton genotypes. Functional predictions of soil microbial communities indicated that microbial community function did not show significant differences between transgenic and non-transgenic cotton samples. These findings are essential for evaluating the environmental effects of transgenic crops and supporting the secure implementation of transgenic cotton.

KEYWORDS

environmental safety assessment, metagenome analysis, microbial community diversity, *ScALDH21*, soil microbial communities, transgenic cotton

1 Introduction

Cotton (*Gossypium hirsutum* L.) is an important fiber crop and oil source globally. The cotton yield accounts for about 30% of the world's total yield every year in China (Sultana et al., 2023). More than 87% of the total Chinese cotton yield was produced in the Xinjiang-Uyghur Autonomous Region of China in 2020. Therefore, cotton is a very important crop in this region. The genetic engineering approach is widely used for crop improvement, especially for yield and resistance (James, 2000; Bacha and Iqbal, 2023). At present, more and more genetically modified (GM) crops are produced and planted commercially around the world, and their potential risks to the environment have attracted much attention (Nap et al., 2003; Angelo et al., 2022; Bacha and Iqbal, 2023).

Environmental safety assessments of GM crops are the main content of commercialization (Walker et al., 2003; Tran et al., 2018; Mandal et al., 2020; Angelo et al., 2022). Soil is an important place for material circulation and energy transformation in the whole ecosystem. Soil ecosystem stability directly affects crop growth and development, and ultimately affects the stability of the whole agricultural system. Many safety assessments have focused on the impact of GM crops on soil ecosystems (Shen et al., 2006; Shen et al., 2015). Transgenic crops may directly affect soil microbial and fungal communities, which in turn may lead to changes in agricultural soil ecosystems. Therefore, soil microbial communities are an important indicator of the safety of transgenic materials (Velmourougane and Sahu, 2013; Zhang et al., 2015a).

Plants affect soil mainly through "rhizo-deposition" and plant residues (Shahmoradi et al., 2019). Plants' influence on the microbial population in the rhizosphere can be more than ten times greater than that in the bulk soil (Ali et al., 2013). This is because the root exudates and plant residues in the rhizosphere create a nutrient-rich and biologically active environment that promotes microbial diversity. Usually, the rich microbial diversity of the root is symbiotically associated with plants to promote better growth (Walker et al., 2003; Han et al., 2017). The soil microbial community of transgenic plants such as maize (*Zea mays*), soybean (*Glycine max*) (Sung et al., 2021),cotton (*Gossypium* sp.) (Shen et al., 2006), rice (*Oryza sativa* L.) (Wu et al., 2021) and tomato (*Solanum lycopersicum*) (Gao et al., 2015) were most studied. Most studies reported that transgenic crops had no or a minor effect on soil microorganisms and they had usually caused temporary changes in the root-rhizosphere microbiome, which disappear after a winter's recovery. The influence of field management, soil properties, and season was greater than that of a transgenic event on soil microorganisms (Shen et al., 2006; Balachandar et al., 2008; Kapur et al., 2010; Velmourougane and Sahu, 2013). Only a few studies have shown that GM crops have significant effects on soil microorganisms (Nap et al., 2003; Griffiths et al., 2007). Studies have found that transgenic *Bt* cotton (overexpressing endotoxin from *Bacillus thuringiensis*) can be cultured with higher bacterial and fungal diversity (Donegan et al., 1995; Velmourougane and Sahu, 2013; Liang et al., 2018).

With the development of high-throughput sequencing approaches, metagenomics has been applied to study soil microorganisms. It breaks through the limitations of unculturable microorganisms (Brooks et al., 2015). High-throughput sequencing of fungal 18S rRNA and internal transcribed spacer (ITS) is most suitable because of high copy number of ribosomal DNA and because it is easy to be amplify and separate. Ribosomal DNA contains conserved regions in 18S, 5.8S, and 28S but has high variability in the ITS region, where the evolution rate is 10 times higher than in 18S rDNA. This makes ITS the preferred method for studying fungal community diversity (Marti et al., 2013). Meanwhile, whole genome sequencing and 16S rDNA sequencing are two main strategies to study the bacterial macrogenomics. 16S rRNA is located on the small ribosomal subunits of prokaryotic cells, including 10 conserved regions and 9 hypervariable regions, which is considered to be the most suitable index for bacterial phylogeny and taxonomic identification (Youssef et al., 2009; Fang et al., 2015; Shen et al., 2015).

In our previous studies, the moss aldehyde dehydrogenase ALDH21 gene from the desiccation-tolerant moss *Syntrichia caninervis* was cloned and transformed into the upland cotton variety Xinnongmian 1 to obtain transgenic cotton (Yang et al., 2016; Yang et al., 2019; Yang et al., 2023). Transgenic cotton is drought-tolerant, salt-resistant and disease-resistant to *Verticillium dahliae*, that can decrease cotton yield in Xinjiang's arid land with limited water supply. However, environmental safety is an important index for transgenic crops, and the assessment of soil microorganisms is an important research direction for safety. In this research, we examined the chemical properties of the soil in cotton fields. We utilized high-throughput sequencing to evaluate the fungal and bacterial communities in soil after one growth period using ITS and 16S rRNA. Alpha and beta diversity were employed to describe the abundance and diversity of microbial communities in addition to exploring inter-sample correlations, species composition correlations with environmental factors, and microbial function predictions to assess the effects of transgenic material on soil microbial communities around the roots. Our objective was to evaluate the environmental safety of transgenic cotton and investigate potential changes in soil microbial populations due to transgenic cotton cultivation. Our findings confirmed that the cultivation of transgenic cotton did not cause significant ecological harm to these microbial communities.

2 Materials and methods

2.1 Experimental materials and design

Soil samples were collected from the artificial infection background of *ScALDH21* transgenic and non-transgenic cotton at the Manasi experiment station of the Xinjiang Academy of Agriculture Science. The Latin square design divided the cropping area into 9 plots where transgenic cotton lines 92, 96 (described as L1, L2, respectively) and non-transgenic cotton line 779 (described as NT) were planted, and each line had 9 repetitions. Since 2013, the same cotton lines have been continuously planted in the experimental plots used in this study. Bulk soil samples from the 5 centimeter periphery and 5-10 centimeter depth of cotton plants were collected after a growth season from the center of plots and stored at -80°C.

The soil chemical properties were detected in the Central Laboratory of Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, based on the methods of forest soil nitrogen (LY/T 1228-2015), phosphorus in forest soil (LY/T1232-2015), potassium in forest soil (LY/T1234-2015), soil organic matter (NY/T1121.6-2006 Soil Testing, part 6), and soil pH (NY/T1121.6-2006 Soil testing, part 2).

2.2 Genomic DNA extraction and PCR amplification

Total soil genomic DNA was extracted with the FastDNA SPIN Kit for Soil (MP Biomedicals, Inc., CA). DNA concentration and purity level were monitored on a 1% agarose gel and a Nanodrop 2000 spectrophotometer, respectively. Based on the concentration, DNA was diluted to 1 ng μL^{-1} using sterile water. Specific primer pairs (515F: 5'-GTGCCAGCMGCCGCGGTAA-3' and 806R: 5'-GGACTACHVGGGTWTCTAAT-3') for the V4 region of 16S rRNA were used to identify bacterial diversity, and ITS (ITS5-1737F: 5'-GGAAGTAAAAGTCGTAACAAGG-3'and ITS2-2043R: 5'-GCTGCGTTCTTCATCGATGC-3') was used to evaluate fungus diversity.

All PCR reactions were carried out in 30 μL reactions with 15 μL of Phusion® High-Fidelity PCR Master Mix (New England Biolabs). Thermal cycling consisted of initial denaturation at 98°C for 1 min, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 50°C for 30 s, and elongation at 72°C for 30 s. Finally, elongation was at 72°C for 7 min. The PCR products were purified with the GeneJETTM Gel Extraction Kit (Thermo Fisher Scientific). Sequencing libraries were generated using the Ion Plus Fragment Library Kit (Thermo Fisher Scientific) following the manufacturer's instructions. The library's quality was assessed on the Qubit@2.0 fluorometer (Thermo Fisher Scientific). Finally, the library was sequenced on an Ion S5TM XL platform, and 400/600 bp single-end reads were generated (Novogene company).

2.3 Sequencing, OTU cluster and species annotation

During analysis the barcodes and primers were removed from the raw sequencing data. After removing the barcode and primer sequences, the sample reads were merged using the FLASH software (V1.2.11, http://ccb.jhu.edu/software/FLASH/) (Magoc and Salzberg, 2011) to obtain Raw Tags. Then, quality control was performed on the Raw Tags using the fastp software to obtain high-quality Clean Tags. Finally, the Clean Tags with a database using the Usearch software to detect chimeric sequences and remove them (Haas et al., 2011), resulting in the final set of effective data known as Effective Tags.

The Effective Tags obtained above, was denoised with the DADA2 module using QIIME2 software (Li et al., 2020), and sequences with and abundance below 5 were filtered to obtain the final Amplicon Sequence Variants (ASVs) and feature table. Next, the classify-sklearn module in QIIME2 software was used align the obtained ASVs with a database in order to obtain species information for each ASV (Callahan et al., 2016). For the 16S regions, the Silva138.1 database (https://www.arb-silva.de/) was used and for the ITS region, used the UNITEv8.2 database (https://unite.ut.ee/).

Principal coordinate analysis (PCoA) and principal component analysis (PCA) analyses were performed using the vegan package in the R programming language (McMurdie and Holmes, 2013; Bro and Smilde, 2014). The vegan package provides specific functions, cmdscale and rda, that can be used to calculate PCoA and PCA. Non-metric multidimensional scaling (NMDS) analysis can be performed using the metaMDS function in the vegan package, which is capable of generating NMDS analysis based on the Bray-Curtis distance.

2.4 The analysis of microorganism diversity and complexity

Alpha diversity was assessed using five indices, including Chao1, Shannon, Simpson, ACE, and PD whole tree. QIIME2 was utilized for index calculation, while R software (version 2.15.3) was employed for dilution curve visualization and analysis of differences among groups based on the alpha diversity index. We utilized the Tukey HSD and Duncan methods of ANOVA to analyze the significance of differences between groups of the alpha diversity index.

Beta diversity was evaluated to determine differences in species complexity among samples. QIIME2 was employed to calculate the Unifrac distance and to construct a UPGMA sample clustering tree. Using R software (version 2.15.3). AMOVA, based on weighted Unifrac distance, was used to test for significant differences between the groups.

2.5 Correlation analysis of environmental factors

We conducted redundancy analysis to demonstrate the relationship between soil chemical properties and OTUs. The relative abundance data of fungal and bacterial communities at the OTU level were considered species data, while soil chemical data were viewed as soil environmental variables. The CANOCO5 software (Smilauer and Jan, 2014) was used to perform RDA analysis to evaluate the correlation between soil microbial communities and soil chemical factors. To assess differences in soil chemical data, ANOVA's Duncan method was employed for analysis of variance and multiple comparisons. The CANOCO5 software was utilized to display the impact of environmental factors on microbial community structure. Given the nonlinear relationship between variables, we computed correlation coefficients between the two variables using the Spearman correlation analysis method, employing the psych package in the R programming language.

2.6 Function prediction

The FunGuild software was utilized to predict fungal functional profiles based on ITS sequencing data using the default database and settings. FAPROTAX (Louca et al., 2016) was used to predict bacterial functional profiles based on 16S rRNA sequencing data using the default database and settings.

3 Results

3.1 Soil chemical properties

To determine chemical properties, soil samples were collected from 5 centimeters around the periphery and 5-10 centimeters deep. We measured eight soil chemical properties in plots planted with two transgenic lines (L1, L2) and non-transgenic cotton (NT). The analysis revealed that the levels of available nitrogen (N) (Figure 1A), available phosphorus (P) (Figure 1B), and available kalium (K) (Figure 1C) in the independent transformed line L2 were slightly higher compared to the transgenic line L1 and the non-transformed line NT (Figures 1A–C). The available N (Figure 1A) and available P (Figure 1B) content in each cotton line did not exhibit significant changes, while a significant difference was observed for available K (Figure 1C) between L2 and NT. The levels of soil organic matter in L1 (16.29 g/kg) and L2 (15.67 g/kg) were slightly higher than NT, although this was not significant (Figure 1D). Likewise, no significant difference was observed in total N and total K in the soil among the different cotton lines (Figures 1E, G). However, it should be noted that the total P content in L1 and L2 was significantly higher than that of NT (Figure 1F). The soil pH values of all cotton lines were found to be alkaline, exceeding 8, and there was no significant difference among the three groups (Figure 1H).

3.2 Sequencing results and quality control

After filtering out low-quality and short sequence reads, a total of 81073.85 and 83378.48 raw reads were obtained from fungal and bacterial sequencing, respectively. Next, after the initial quality control process, on average 80165.93 and 78516.26 clean reads were obtained, respectively. The average length of fungal ITS was 227.3 bp and that of bacterial 16S rRNA was 417 bp. The sequencing quality was high, with a sequencing error rate less than 1% in clean reads. The operational transcriptional unit (OTU) clustering was performed with 97% consistency, and fungal sequencing yielded 3566 OTUs (Table S1), and bacterial sequencing yielded 7481 OTUs (Table S2). The rationality of total OTUs of each soil sample was detected by rarefaction curves. With the increase in sequencing depth, OTU numbers of fungi (Figure 2A) slowly increased, and the curve tended to be smooth.

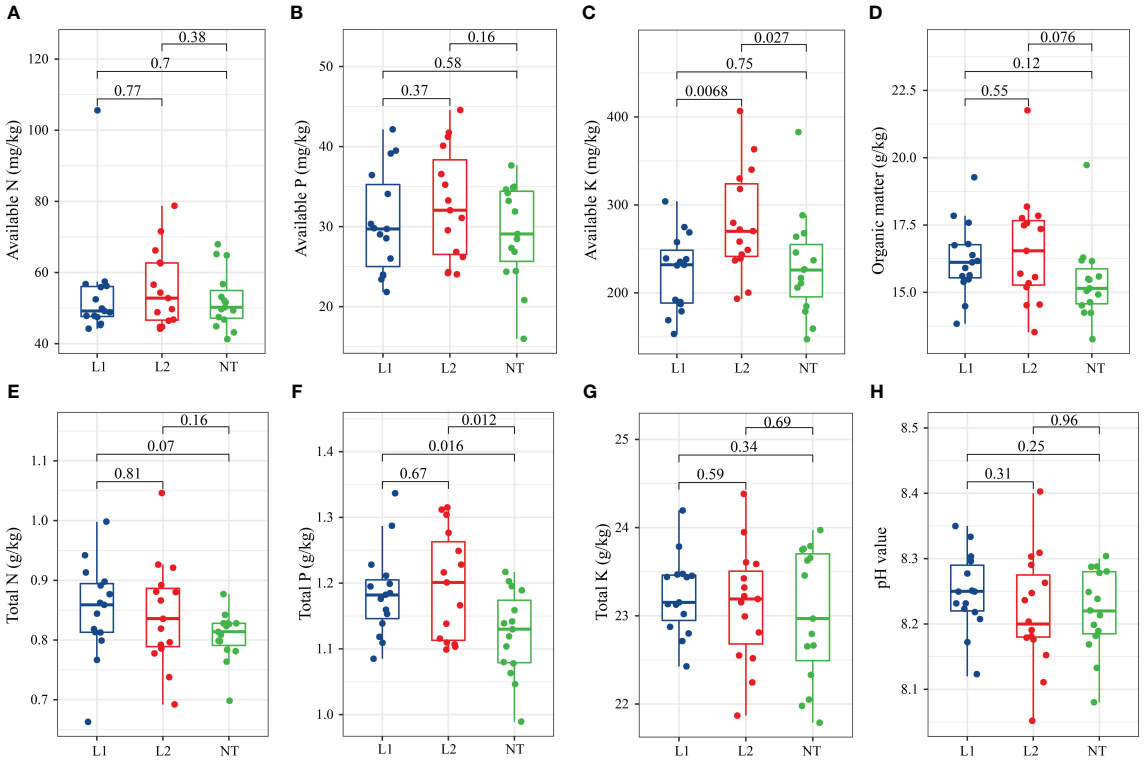


FIGURE 1
Soil chemical properties in cotton fields. **(A)** Available N (mg/kg) **(B)** Available P (mg/kg). **(C)** Available K (mg/kg). **(D)** Organic matter (g/kg). **(E)** Total N (g/kg). **(F)** Total P (g/kg). **(G)** Total K (g/kg). **(H)** pH value. The data were tested using the t-test, and a mean ± standard deviation was reported for each group with a sample size of n = 15 in each group. Statistical significance was expressed numerically, with a significance level of $P < 0.05$ indicating a significant difference between the two groups.

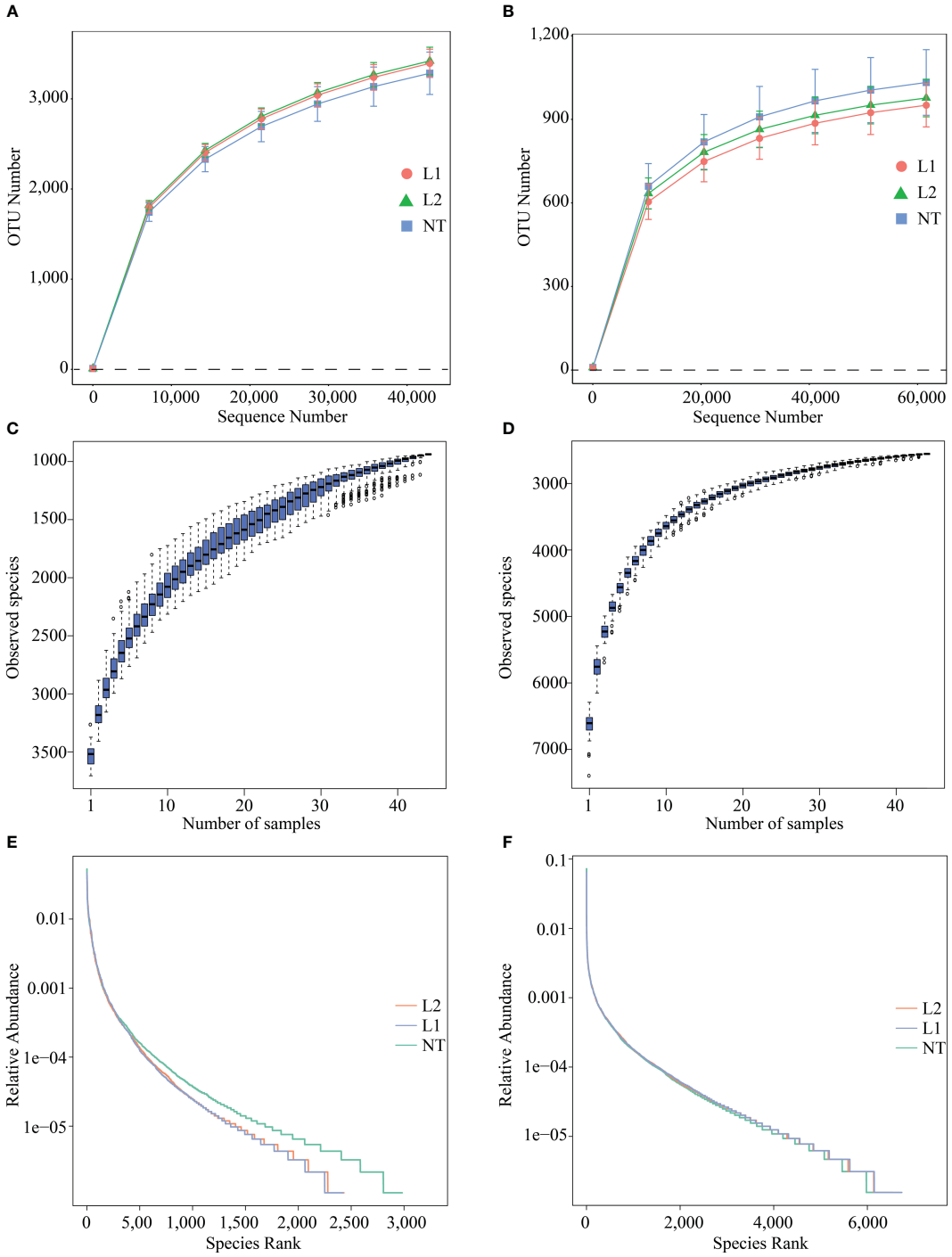


FIGURE 2
Species diversity analysis of sequenced samples. **(A)** Rarefaction Curve for fungal sequencing. **(B)** Rarefaction Curve for bacteria sequencing. **(C)** species accumulation boxplot of species sequenced by fungi. **(D)** species accumulation boxplot of species sequenced by bacteria. **(E)** Rank abundance curve for fungi sequencing. **(F)** Rank abundance curve for bacteria sequencing.

The rarefaction curves of L1, L2, and NT samples were consistent with this trend, and the sequencing depth of fungi was reasonable. The bacterial (Figure 2B) rarefaction curve was basically flat but still not saturated, and a small number of bacterial species may not be detected. The number of OTUs increases with the sample size, as depicted in the species accumulation boxplot (Figures 2C, D). The curve reaches stability as the number of OTUs saturates and fewer new OTUs are added to each sample, signifying the sequencing depth's accuracy in representing fungal and bacterial composition. We used rank abundance curves to compare the differences in species composition between transgenic and non-transgenic cotton microbiomes. The results showed that, in different groups, the curve shapes of fungal communities were roughly similar but deviated at the same abundance threshold (Figure 2E), while the curve shapes

of bacterial communities were also similar but did not deviate at the same abundance threshold (Figure 2F). In addition, the inflection points of the curves were analyzed, and they all appeared to have similar abundance values. The analysis of rank abundance curves indicated that the species composition and distribution patterns of root-associated microbiomes in transgenic and non-transgenic cotton were consistent.

3.3 Comparative analysis of multiple samples

The principal coordinate analysis (PCoA) was executed to explore the dissimilarity among microbial communities obtained from different soils. Variation in fungi was explained by 20.43% and 8.89% by PCoA1 and PCoA2, respectively (Figure 3A), while variation in bacteria was explained by 15.63% and 9.47% by PCoA1 and PCoA2, respectively (Figure 3B). PCoA analysis demonstrated the sample clustering of L1, L2, and NT, where soil samples from different groups clustered together, indicating a similarity in the microbial community structure between the different groups. We carried out principal component analysis (PCA) to evaluate the structural variances of fungal and bacterial communities in the different samples. PC1 and PC2 explained 9.74% and 5.6% of the total variation in fungi, respectively (Figure 3C), while PC1 and PC2 explained 5.6% and 4.65% of the total variation in bacteria, respectively (Figure 3D). PCA1 and PCA2 scatter plots showed no significant separation between the different groups, suggesting that no significant variation in fungal and bacterial community composition was observed between the samples. Non-metric multi-dimensional scaling (NMDS) was used to analyze the variability between fungal and bacterial communities at different ecological niches. MDS1 and MDS2 scatter plots indicated no significant variances in the fungal and bacterial communities of different groups and showed some similarity in their composition and relative abundance (Figures 3E, F).

3.4 Microbial diversity

In this study, we examined the alpha and beta diversity indices for fungi and bacteria in soil samples from transgenic (L1 and L2) and non-transgenic (NT) cotton. These indices reflect the abundance and diversity of fungi and bacteria in the soil. The coverage in all three samples exceeded 99.9%, indicating near-perfect coverage (Figure 4A). We used Shannon's diversity index (Shannon), the Chao1 index (Chao 1), and the abundance Coverage-Based Estimator (ACE) indices, which are widely used in ecology to assess fungal and bacterial abundance. Our findings show that the Shannon and Chao 1 indices did not exhibit significant variation between the transgenic and non-transgenic cotton soils (Figures 4B, C). The ACE fungal level was slightly higher than L1 in NT but not different from L2 (Figure 4D). Beta diversity analysis revealed variability between L2 and NT and L1 in the fungal community, resulting in reduced diversity, while L2 exhibited reduced bacterial diversity, but overall, the magnitude of change was not significant (Figure S1). In conclusion, our findings suggest that there were no substantial changes in the diversity of soil microorganisms between transgenic and non-transgenic cotton samples.

3.5 Composition of soil microbial community

The statistical analysis revealed that the L1, L2, and NT groups had 2422, 2431, and 2982 OTUs for fungal species, respectively, and 1777 OTUs were common to all groups (Figure 5A). In the case of bacterial OTUs, the statistical analysis revealed that L1, L2, and NT had 6726, 6682, and 6844 OTUs, respectively, and 5737 OTUs were common to all groups (Figure 5B). A statistical analysis was conducted at the phylum level for fungi and bacteria in the three groups. The results showed that the three groups had similar dominant species at phylum level in both fungi and bacteria. The fungal phyla *Ascomycota*, *Mortierellomycota*, *Basidiomycota*, *Kickxellomycota*, and *Mucoromycota* were widely distributed across all groups (Figure 5C). Similarly, *Proteobacteria*, *Actinobacteria*, *Acidobacteria*, *Chloroflexi*, and *Bacteroidetes* were relatively dominant among bacteria (Figure 5D). To conduct further investigation on the phylogenetic relationships and abundance at the genus level, representative sequences for the most abundant genera were aligned through multiple sequences, subsequently constructing a phylogenetic tree. The analysis of the bacterial and fungal genera in three groups, showed that the dominant genera were similar across all three groups, and all from the dominant phylum-level category. Specifically, *Mortierella*, *Cephalotrichum*, *Fusarium*, *Alternari*, and *Penicillium* were the main soil fungal genera (Figure 5E), while *Arthrobacter*, *Acidobacteria*, *Sphingomonas*, *Pseudomonas*, *Bacillus*, and *Skermanella* were the dominant soil bacterial genera (Figure 5F).

3.6 Screening of dominant species between different groups

To uncover variations in dominant species among the three sample groups at different taxonomic levels (phylum, genus, and species), we identified the top 10 most abundant species for each level. These species were used to create a ternary plot, which allows for clear visualization of the differences in dominant species across the three sample groups at each taxonomic level. *Olpidiomycota* and *Blastocladiomycota* were found to be dominant in NT, whereas *Glomeromycota* and *Kickxellomycota* were more abundant in L1 and L2 (Figure 6A). In contrast, there were no significant differences in bacterial phyla abundance among the three groups (Figure 6B). At the fungal genus level, *Neocamarosporium* and *Lecanicillium* were found to be dominant in NT, while *Leptosphaeria*, *Gibberella*, and *Chaetomium* were found to be dominated in L1 and L2 (Figure 6C). The bacterial genera *Lachnospiraceae* and *Lactobacillus* showed dominance in relative abundance in NT, whereas *Intestinimonas* dominated in relative abundance in L1 and L2 (Figure 6D). Notably, there were no significant differences in *Bacillus* abundance among the three groups (Figure 6D). Fungal species including *Mortierella alpina*, *Fusarium*

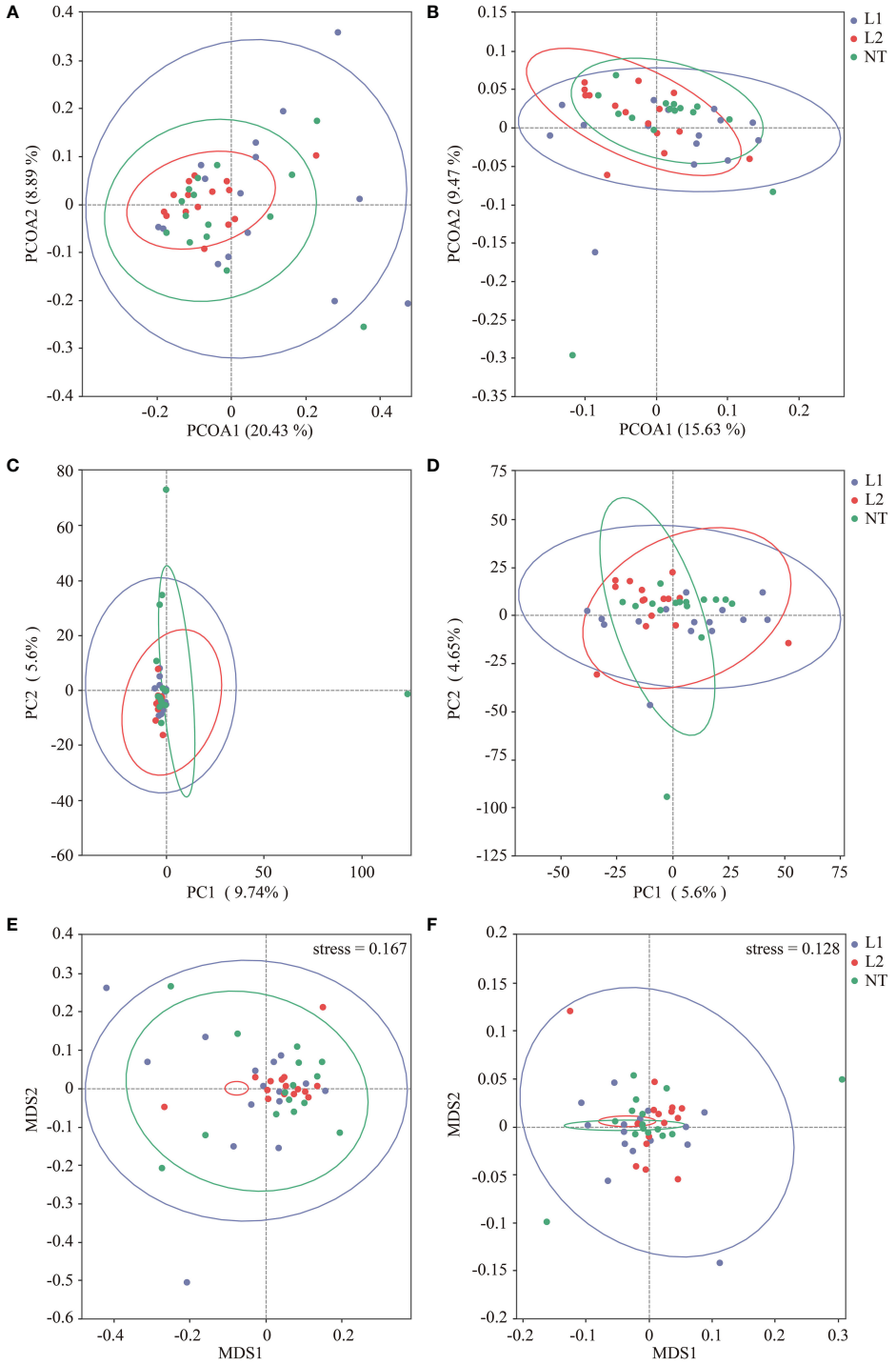


FIGURE 3
Multidimensional analysis of samples using PCOA, PCA, and NMDS. **(A)** fungi PCoA analysis. **(B)** Bacteria PCoA analysis. **(C)** fungi PCA analysis. **(D)** Bacteria PCA analysis. **(E)** fungi PCA analysis. **(F)** Bacteria PCA analysis.

delphinoides, and *Cephalotrichum nanum* were identified as dominant among the three groups. Furthermore, *Neocamarosporium chichastianum* was found to have relative dominance in the group categorized as NT, while *Chaetomium longicolleum*, *Albifimbria verrucaria*, *Gibberella intricans*, and *Leptosphaeria sclerotioides* were found to have relative dominance L1 and L2 groups (Figure S2A). In addition, the bacterial communities were dominated by *Pseudomonas frederiksbergensis* shared among all three groups. However, *Lachnospiraceae bacterium* A4 showed dominant relative abundance in the NT group. Similarly, L1 and L2 groups, the *Lachnospiraceae bacterium* DW17, *Pedobacter duraquae*, and *Stenotrophomonas chelatiphaga* were also found to be dominant (Figure S2B).

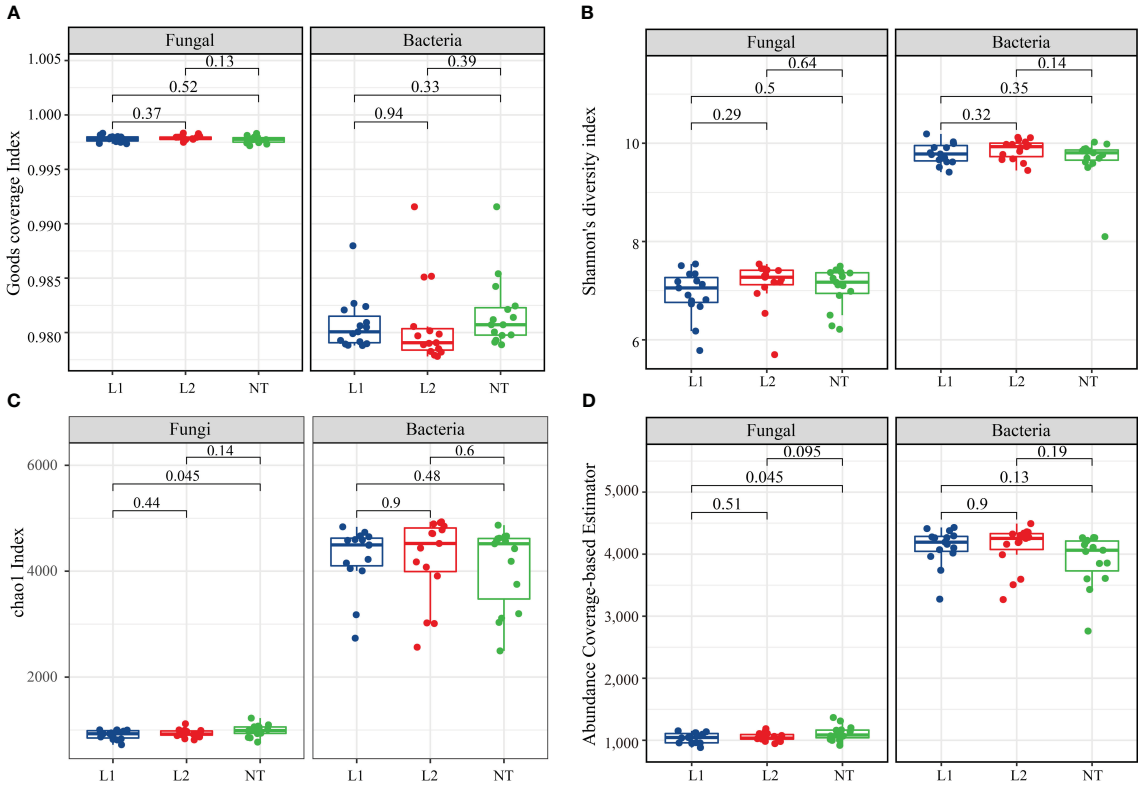


FIGURE 4
Alpha diversity index. **(A)** Goods coverage index for fungi and bacteria. **(B)** Shannon's diversity index for fungi and bacteria. **(C)** chao1 Index for fungi and bacteria. **(D)** Abundance coverage-based estimator for fungi and bacteria. The data were tested using the t-test, and a mean ± standard deviation was reported for each group with a sample size of n = 15 in each group. Statistical significance was expressed numerically, with a significance level of $P < 0.05$ indicating a significant difference between the two groups.

3.7 Redundancy analysis of microbial and environmental factors

We investigated the correlation between microbial and environmental factors by conducting a redundancy analysis (RDA) of bacteria and fungi on a genus level. The RDA results indicated that 29.76% of the correlation was identified on axis 1 and 23.23% on axis 2 for the fungal genus level. The analysis indicates that available K had a prominent role in explaining axis 1, while pH had a crucial role in explaining axis 2. Among them, available K showed the highest impact on *Penicillium*, whereas pH had the most significant effect on *Alternaria* (Figure 7A). Our findings also revealed that the RDA analysis carried out at the bacterial genus level explained 28.92% on axis 1 and 19.24% on axis 2. The analysis identified that available P had the highest contribution in Axis 1, whereas total P had the greatest contribution to Axis 2. Available P had a more profound impact on *Acidobacteria*, whereas total P significantly impacted *Skermanella* (Figure 7B). Spearman correlation analysis and coefficients were employed in this study to determine the correlation between environmental factors and microbial species richness (alpha diversity). Spearman rank correlation was used to examine the inter-variation relationship between environmental factors and species. The purpose of this analysis was to obtain a correlation and determine significant differed values between the two variables. The abundance of fungal genera was significantly influenced by various environmental factors. Specifically, *Dactylonectria* was significantly affected by available P, organic matter, total N, and total P, whereas *Conocybe* abundance was significantly affected by available K. *Leptosphaeria* abundance was significantly influenced by total K, while *Gibberella* abundance was significantly affected by total P. In the case of *Lecanicillium*, available P had a significant impact on its abundance, while the abundance of *Solicoccozyma* was significantly influenced by pH (Figure 7C). At the bacterial genus level, total K had a significant impact on *Solirubrobacter* abundance, while *Rhizobiaceae* abundance was significantly affected by available P. Available P and pH had a significant impact on *Lactobacillus* abundance, whereas these environmental factors had no significant effect on *Bacillus* (Figure 7D).

3.8 Functional prediction of soil microbial community

We used the FUNGuild functional prediction software for ITS, and its prediction of guilds allowed us to study fungal functions from an additional ecological perspective. The results showed that the software could not predict the majority of fungal functions. The most abundant predicted functions were plant pathogen, soil saprotroph, and wood saprotroph, while animal pathogen was

10.3389/frmbi.2023.1248384

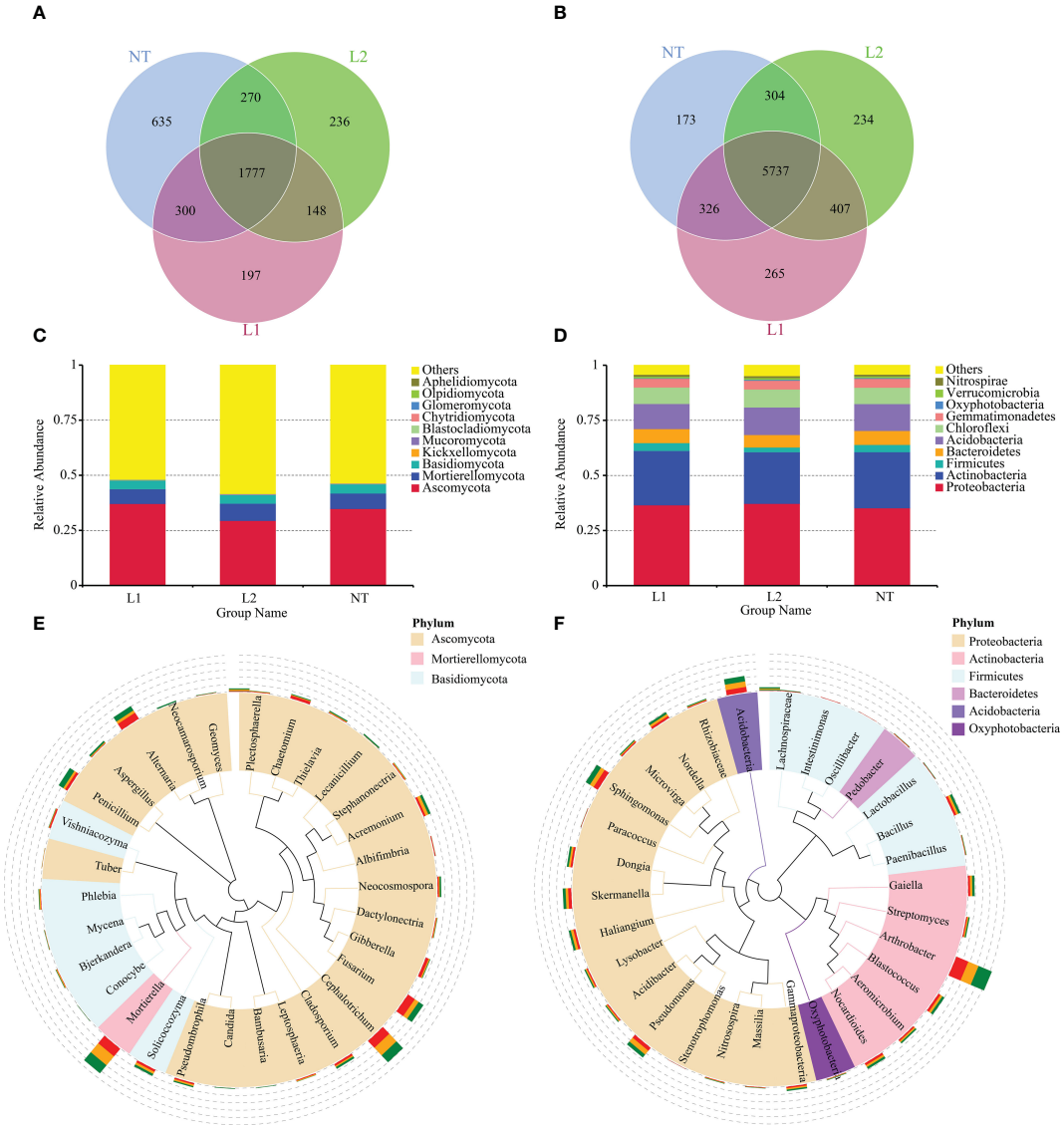


FIGURE 5
Species composition of the sample microbial community. **(A)** Venn diagram of samples based on fungal OTUs. **(B)** Venn diagram of samples based on bacteria OTUs. **(C)** Species composition at the fungal phylum level. **(D)** Species composition at the bacteria phylum level. **(E)** Species composition at the fungal genus level. **(F)** Species composition at the bacteria genus level.

also prevalent (Figure 8A). We predicted the effects of bacterial biochemical cycling processes for different environmental samples using the FAPROTAX functional classification database based on species information. The results indicated that chemoheterotrophy had the highest abundance, and fermentation, nitrate reduction, nitrogen respiration, and denitrification also accounted for a significant proportion of them (Figure 8B). In conclusion, our functional predictions of soil microbial communities were quite similar between transgenic and non-transgenic cotton, with no significant differences in microbial community functions observed.

4 Discussion

The goal of this research was to evaluate the influence of *ScALDH21* transgenic cotton on soil microbial communities' environments. Using high-throughput sequencing and data analysis, we investigated the impacts of transgenic and non-transgenic cotton on the soil microbes. Our analysis revealed no significant differences between the microbial composition of transgenic and non-transgenic cotton. These findings imply that there is no statistical distinction between microbial community

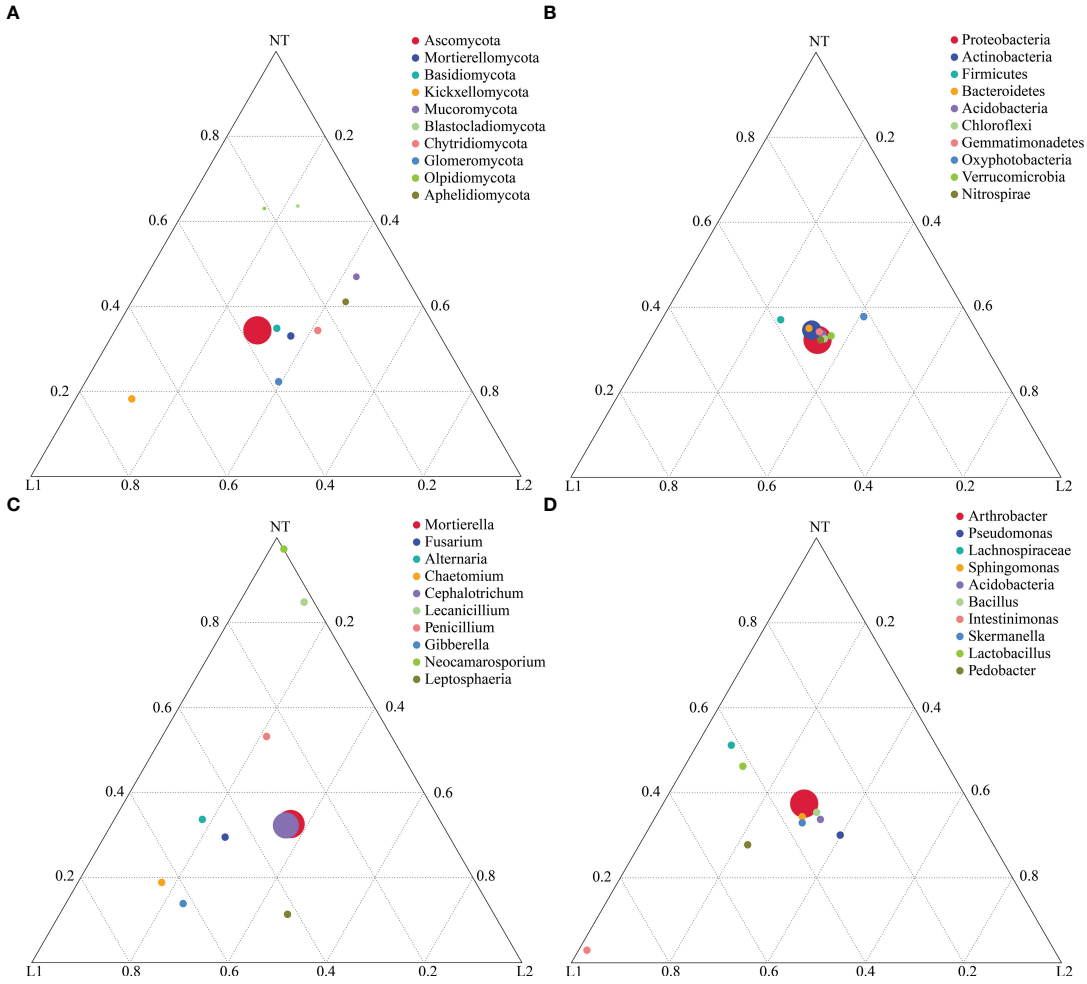


FIGURE 6
Highlighting dominant species on a ternary plot. **(A)** Fungal phylum level. **(B)** Bacteria phylum level. **(C)** Fungal genus level. **(D)** Bacteria genus level.

compositions on soil of transgenic and non-transgenic cotton. The results provide insights that promote major environmental challenges associated with transgenic cotton and affirm the safety of *ScALDH21* transgenic cotton.

Different plant species tend to cultivate their own root microflora comprised of diverse microorganisms, which reflects the soil's agricultural ecosystem. The soil surrounding the plant's root system is usually divided into two parts: rhizosphere soil and bulk soil. The rhizosphere soil is a narrow area (~5 mm) that surrounds the plant's root system, while the bulk soil extends beyond this distance. During the plant's vegetative period, the soil microorganisms in this area are greatly affected by the "rhizosphere effect" compared with those in the bulk soil (Siciliano et al., 1998; Han et al., 2017; Shahmoradi et al., 2019; Vuong et al., 2020). Therefore, it is not suitable to evaluate the safety of transgenic plants within the rhizosphere soil. In this study, the environmental safety of transgenic crops was evaluated by selecting the bulk soil within 5 cm beyond the cotton root after the completion of the cotton-growing period.

The pH of the soil in all cotton fields exceeded eight, suggesting that the soil was alkaline and must have had a limiting effect on certain microbial populations. Previous studies have also shown that changes in bacterial communities are significantly correlated with pH (Li et al., 2018). Soil microorganisms are significantly affected by the effect of organic matter content in soil on saprophytic fungi (Setala and McLean, 2004). Soil chemical property analysis revealed that the soil in the cotton field exhibited alkaline properties. After one year, the content of organic matter in L1 and L2 was slightly higher compared to the non-transgenic cultivars. Significant differences were found only in the levels of available K and total P, both of which also showed significant increases. These findings suggest that the cultivation of all cotton cultivars did not have a significant adverse impact on the soil's chemical properties. Furthermore, in comparison to other studies, these results are affected by species, plant breed, and geographic variation (Shen et al., 2006; Liang et al., 2018), which means that more research on these factors is needed.

Our sequencing results showed that species diversity of both fungi and bacteria was similar in both transgenic and non-transgenic cotton, which is consistent with previous studies indicating that transgenic *Bt* cotton has no significant effect on soil microbial communities (Shen et al., 2006; Hu et al., 2009; Kapur

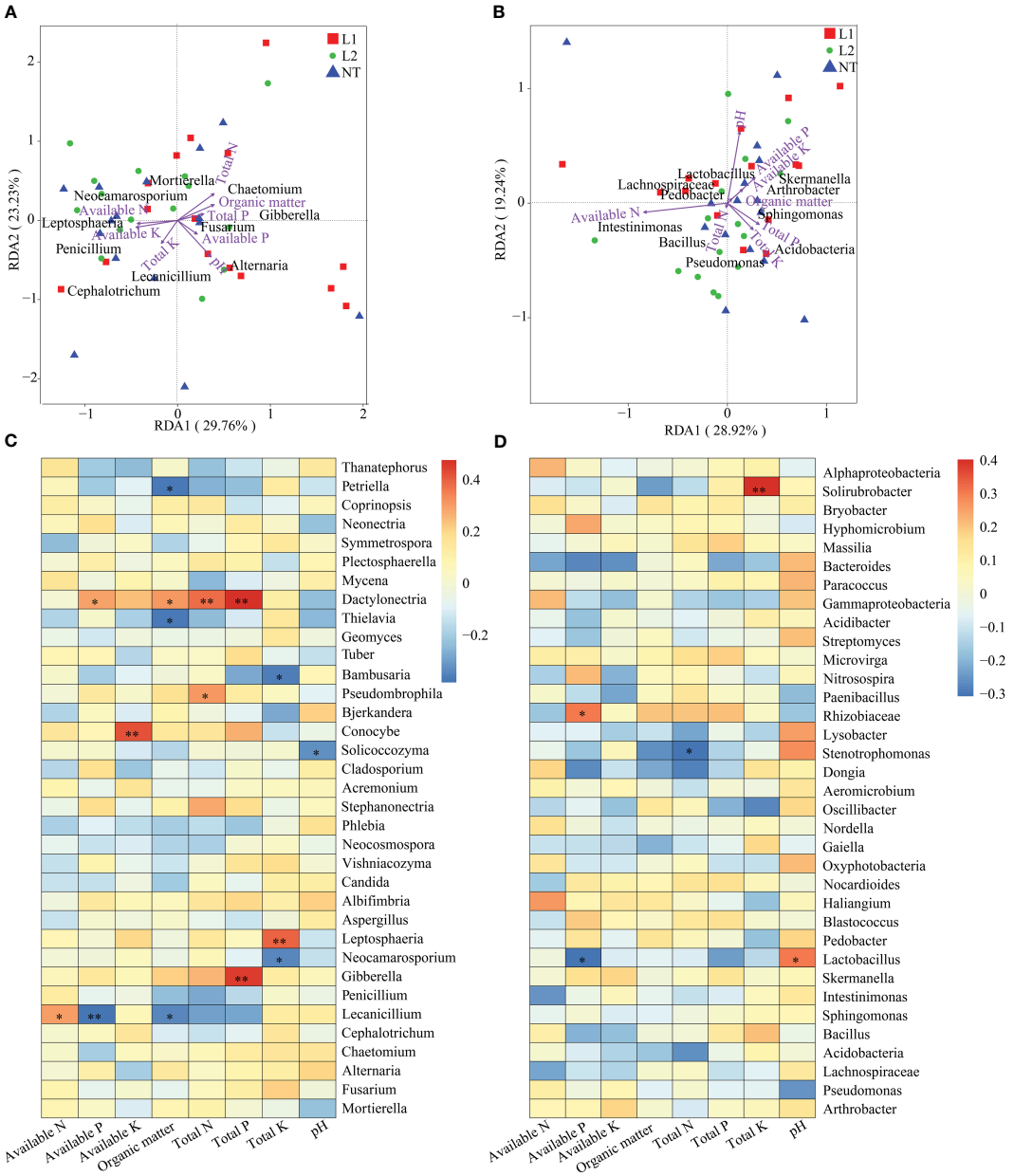


FIGURE 7
Environmental factor correlation analysis. **(A)** RDA of fungal genus-level diversity and environmental factors. **(B)** RDA of bacteria genus-level diversity and environmental factors. **(C)** Spearman analysis of fungal genus-level diversity and environmental factors. **(D)** Spearman analysis of bacteria genus-level diversity and environmental factors. In the heatmap, each cell's color represents the correlation between two variables. * In the heatmap indicates a significance level of 0.05, and ** indicates a significance level of 0.01.

et al., 2010; Zhang et al., 2015b). We also performed a comparative analysis of multiple samples to compare the diversity and composition of soil microorganisms between transgenic and non-transgenic cotton samples. Our analysis of alpha and beta diversity indices revealed that there were no significant variations in the diversity of soil microorganisms between the two groups, supporting the notion that transgenic cotton expressing multi-resistant *ScALDH21* has negligible effects on soil microbial communities.

Previous studies have reported conflicting results on the impact of transgenic crops on soil microbial communities. Transgenic *Bt* rice lines produce a significant effect on the diversity and abundance of endophytic bacteria in leaves and roots, while only minor differences were found in the bacterial communities of the inter-root soils between *Bt* and non-*Bt* rice lines (Wu et al., 2021). Our results indicate that transgenic cotton expressing multi-resistant *ScALDH21* had negligible effects on the diversity and composition of soil microorganisms, which is consistent with previous results that transgenic crop cultivation had no effect on soil microorganisms (Liang et al., 2018; Guan et al., 2021; Chen et al., 2022; Xu et al., 2023). Interestingly, the transgenic ath-miR399d tomato line has an effect on microbial communities and diversity in

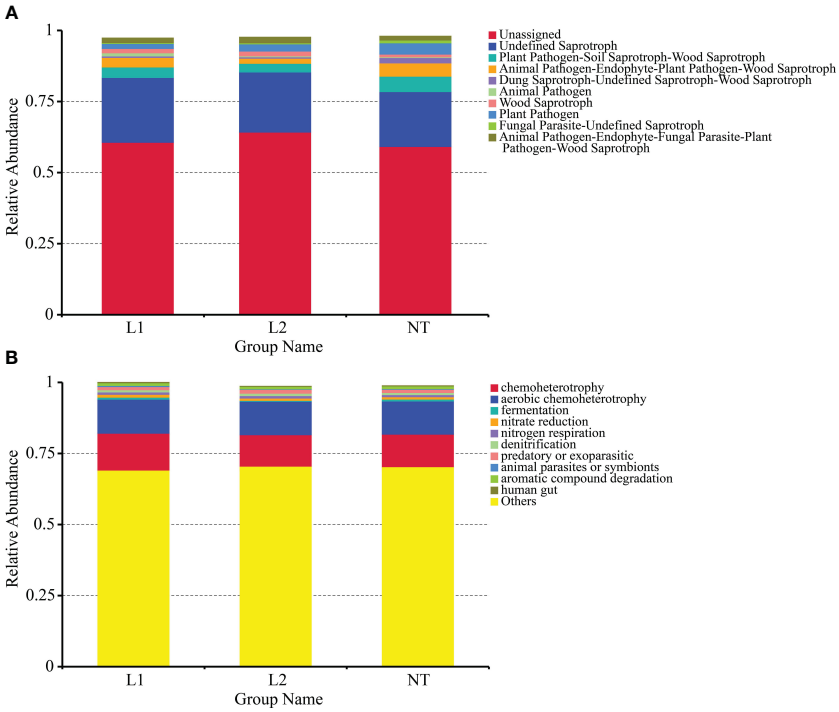


FIGURE 8
Functional prediction of soil microbial community based on metagenomic analysis. **(A)** Functional classification of fungi funguild in soils. **(B)** Functional classification of bacteria faprotax in soils.

soil over a short period of time (Gao et al., 2015), and this difference may be due to the different transgenic plants and microbial communities investigated in these studies. We also used ternary plots to visualize the differences in the abundance of dominant genera between the groups. We found that the difference in abundance of dominant genera between transgenic and non-transgenic cotton was small but not negligible. There were some genera in relative abundance in both transgenic and non-transgenic cotton samples that were different, such as *Leptosphaeria*, *Intestinimonas*, and *Pedobacter*, suggesting that these genera may be more sensitive to the transgenic cotton. Overall, our findings indicate that transgenic cotton expressing multi-resistant *ScALDH21* has minimal impact on the diversity and composition of soil microorganisms, which is consistent with previous studies on *Bt* cotton. However, our study also provides further insights into the abundance of dominant genera in transgenic and non-transgenic cotton, which may have important implications for monitoring potential ecological risks associated with the widespread use of transgenic crops.

To assess the environmental safety of transgenic cotton expressing *ScALDH21* gene, we analyzed the diversity and composition of soil microorganisms and their relationship with environmental factors. Our RDA analysis showed a significant association between soil chemical properties and soil microbial community composition, which is consistent with previous studies (Xu et al., 2023). However, we found that no significant differences were seen in different cotton lines. Although numerous studies have investigated the effects of environmental factors on soil microbial communities, few have explored the effects of transgenic cotton on these communities. Our results suggest that transgenic cotton expressing multi-resistant *ScALDH21* did not affect the correlation between microbial and environmental factors compared to non-transgenic cotton. Similarly, previous studies found no significant difference in the relationship between microbial composition and environmental factors in soils planted with transgenic *Bt* cotton and non-transgenic cotton (Hu et al., 2009). This finding suggests that the nutrient availability in soil has a significant impact on the fungal community. In conclusion, our RDA analysis demonstrated that there was no significant difference in the correlation between microbial composition and environmental factors between transgenic and non-transgenic cotton. These findings provide essential insights into the environmental safety of transgenic cotton expressing multi-resistant *ScALDH21* and support the safe implementation of transgenic cotton. Nonetheless, more studies are required to determine the long-term effects of transgenic crops on soil microbial communities and their relationships with environmental factors.

The functional prediction results demonstrated that there were no notable differences in microbial community function between the three groups. Our study provides further evidence that the use of transgenic crops expressing multi-resistant *ScALDH21* has minimal effects on soil microbial communities, similar to *Bt* cotton (Shen et al., 2006). Fungi and bacteria play critical roles in soil nutrient cycling and plant growth. As such, we used functional prediction analysis to investigate the functional diversity of the soil microbial community based on the abundance of genus-level OTUs. Our results revealed no significant differences in microbial community

function between transgenic and non-transgenic cotton samples. The consistency between the microbial functional profiles suggests that transgenic cotton expressing multi-resistant *ScALDH21* does not adversely affect the activities of soil microorganisms.

Our study results provide insights into the environmental safety of transgenic cotton expressing multi-resistant *ScALDH21* and support the safe implementation of these crops. However, the long-term effects of transgenic crops on soil microbial communities remain unclear. Therefore, long-term studies are required to ensure the environmental sustainability of transgenic crops.

5 Conclusion

In this study, the impact of ScALDH21 transgenic cotton on soil microbial communities was assessed using high-throughput sequencing technology. The results showed that there were no significant differences in the diversity and composition of soil microorganisms between transgenic and non-transgenic cotton. Environmental factors such as soil pH, available K, available P and total P had a significant impact on microbial community composition and species richness. Analysis also revealed that the dominant phylum, genus, and species were similar across all three groups. Overall, these findings demonstrate the feasibility of *ScALDH21* transgenic cotton from an environmental safety perspective and provide important insights for the safety evaluation of transgenic crops.

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

All authors contributed to the study's conception and design. DYZ and HY conceived of and designed the experiments. QY, JW, and HF performed the experiments and wrote the articles. DWZ, HF and HY helped to perform the experiments and collected the data. QY and JW participated in the statistical analysis. TB helped with the chart processing. QY, HY, TB and DYZ contributed to manuscript discussion and revision. All authors contributed to the article and approved the submitted version.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from the West Light Talents Cultivation Program of Chinese Academy of Sciences (2021-XBQNXZ-015); The National Natural Science Foundation of China (Grant No. 31700289); Key Research Program of Frontier Sciences, Chinese Academy of Sciences (ZDBS-LY-SM009); The Natural Science Foundation of Xinjiang Uygur Autonomous Region (2022D01A354); Leading Talents in Technological Innovation programme, 2022TSYCLJ0049.

Acknowledgments

We are grateful to Prof. Jianhui Xu and Prof. Haijiang Xu of the Research Institute of Economic Crops, Xinjiang Academy of Agricultural Sciences, China, for cotton planting support. We are grateful to Ma Qingqian, Chen Lihua and Huo Chunxiao from Xinjiang key lab of Conservation and Utilization of Gene Resources, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences for their help in accomplishing writing the manuscript.

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.

The author(s) DZ declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1248384/full#supplementary-material

References

Ali, H., Khan, E., and Sajad, M. A. (2013). Phytoremediation of heavy metals-Concepts and applications. *Chemosphere.* 91 (7), 869–881. doi: 10.1016/j.chemosphere.2013.01.075

Angelo, V. R., Cristina, R. O., Esther, S. V., Arturo, H. T., Carmen, R. A., and Conrado, C. I. (2022). Myths and realities about genetically modified food: A risk-benefit analysis. *Appl. Sci.* 12 (6), 2861. doi: 10.3390/app12062861

Bacha, S. A. S., and Iqbal, B. (2023). Advancing agro-ecological sustainability through emerging genetic approaches in crop improvement for plants. *Funct. Integr. Genomics* 23 (2), 145. doi: 10.1007/s10142-023-01074-4

Balachandar, D., Raja, P., Nirmala, K., Rithyl, T., and Sundaram, S. (2008). Impact of transgenic *Bt*-cotton on the diversity of pink-pigmented facultative methylotrophs. *World J. Microbiol. Biotechnol.* 24, 2087–2095. doi: 10.1007/s11274-008-9713-7

Bro, R., and Smilde, A. K. (2014). Principal component analysis. *Anal. Methods* 6 (9), 2812–2831. doi: 10.1039/C3AY41907J

Brooks, J. P., Edwards, D. J., Harwich, M. D., Rivera, M. C., Fettweis, J. M., Serrano, M. G., et al. (2015). The truth about metagenomics: Quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol.* 15 (1), 1–14. doi: 10.1186/s12866-015-0351-6

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13 (7), 581–583. doi: 10.1038/nmeth.3869

Chen, Y., Pan, L., Ren, M., Li, J., Guan, X., and Tao, J. (2022). Comparison of genetically modified insect-resistant maize and non-transgenic maize revealed changes in soil metabolomes but not in rhizosphere bacterial community. *GM Crops Food* 13 (1), 1–14. doi: 10.1080/21645698.2022.2025725

Donegan, K., Palm, C., Fieland, V., Porteous, L., Ganio, L., Schaller, D., et al. (1995). Changes in levels, species and DNA fingerprints of soil microorganisms associated with cotton expressing the Bacillus thuringiensis var. kurstaki endotoxin. *Appl. Soil Ecol.* 2 (2), 111–124. doi: 10.1016/0929-1393(94)00043-7

Fang, H., Wang, H., Cai, L., and Yu, Y. (2015). Prevalence of antibiotic resistance genes and bacterial pathogens in long-term manured greenhouse soils as revealed by metagenomic survey. *Environ. Sci. Technol.* 49 (2), 1095–1104. doi: 10.1021/es504157v

Gao, N., Shen, W., Lin, X., and Shi, W. (2015). Influence of transgenic ath-miR399d tomato lines on microbial community and diversity in rhizosphere soil. *Soil Sci. Plant Nutr.* 61 (2), 259–268. doi: 10.1080/00380768.2014.970116

Griffiths, B. S., Caul, S., Thompson, J., Birch, A. N. E., Cortet, J., Andersen, M. N., et al. (2007). Microbial and microfaunal community structure in cropping systems with genetically modified plants. *Pedobiologia* 51 (3), 195–206. doi: 10.1016/j.pedobi.2007.04.002

Guan, Z. J., Wei, W., Stewart, J. C. N., and Tang, Z. X. (2021). Effects of transgenic oilseed rape harboring the Cry1Ac gene on microbial communities in the rhizosphere soil. *Eur. J. Soil Biol.* 103, 103277. doi: 10.1016/j.ejsobi.2021.103277

Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., et al. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* 21 (3), 494–504. doi: 10.1101/gr.112730.110

Han, G., Lan, J., Chen, Q., Yu, C., and Bie, S. (2017). Response of soil microbial community to application of biochar in cotton soils with different continuous cropping years. *Sci. Rep.* 7 (1), 1–11. doi: 10.1038/s41598-017-10427-6

Hu, H. Y., Liu, X. X., Zhao, Z. W., Sun, J. G., Zhang, Q. W., Liu, X. Z., et al. (2009). Effects of repeated cultivation of transgenic *Bt* cotton on functional bacterial populations in rhizosphere soil. *World J. Microbiol. Biotechnol.* 25, 357–366. doi: 10.1007/s11274-008-9899-8

James, C. (2000). "Transgenic crops worldwide: Current situation and future outlook," in *Agricultural biotechnology in developing countries: Towards optimizing the benefits for the poor*, (Boston, MA: Springer) 11–23. doi: 10.1007/978-1-4757-3178-1_2

Kapur, M., Bhatia, R., Pandey, G., Pandey, J., Paul, D., and Jain, R. K. (2010). A case study for assessment of microbial community dynamics in genetically modified *Bt* cotton crop fields. *Curr. Microbiol.* 61, 118–124. doi: 10.1016/j.scitotenv.2018.05.013

Li, M., Shao, D., Zhou, J., Gu, J., Qin, J., Chen, W., et al. (2020). Signatures within esophageal microbiota with progression of esophageal squamous cell carcinoma. *Chin. J. Cancer Res.* 32 (6), 755. doi: 10.21147/j.issn.1000-9604.2020.06.09

Li, P., Li, Y., Shi, J., Yu, Z., Pan, A., Tang, X., et al. (2018). Impact of transgenic Cry1Ac+ CpTI cotton on diversity and dynamics of rhizosphere bacterial community of different root environments. *Sci. Total Environ.* 637, 233–243. doi: 10.1016/j.scitotenv.2018.05.013

Liang, J., Jiao, Y., Luan, Y., Sun, S., Wu, C., Wu, H., et al. (2018). A 2-year field trial reveals no significant effects of GM high-methionine soybean on the rhizosphere bacterial communities. *World J. Microbiol. Biotechnol.* 34, 1–10. doi: 10.1007/s11274-018-2495-7

Louca, S., Parfrey, L. W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science* 353 (6305), 1272–1277. doi: 10.1126/science.aaf4507

Magoc, T., and Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27 (21), 2957–2963. doi: 10.1093/bioinformatics/btr507

Mandal, A., Sarkar, B., Owens, G., Thakur, J., Manna, M., Niazi, N. K., et al. (2020). Impact of genetically modified crops on rhizosphere microorganisms and processes: A review focusing on *Bt* cotton. *Appl. Soil Ecol.* 148, 103492. doi: 10.1016/j.apsoil.2019.103492

Marti, E., Jofre, J., and Jose, L. B. (2013). Prevalence of antibiotic resistance genes and bacterial community composition in a river influenced by a wastewater treatment plant. *PLoS One* 8 (10), e78906. doi: 10.1021/es504157v

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8 (4), e61217. doi: 10.1046/j.0960-7412.2002.001607.x

Nap, J. P., Metz, P. L., Escaler, M., and Conner, A. J. (2003). The release of genetically modified crops into the environment: Part I. Overview of current status and regulations. *Plant J.* 33 (1), 1–18. doi: 10.1046/j.0960-7412.2002.001607.x

Setala, H., and McLean, M. A. (2004). Decomposition rate of organic substrates in relation to the species diversity of soil saprophytic fungi. *Oecologia* 139, 98–107. doi: 10.1007/s00442-003-1478-y

Shahmoradi, Z. S., Tohidfar, M., Marashi, H., Malekzadeh, S. S., and Karimi, E. (2019). Cultivation effect of chitinase-transgenic cotton on functional bacteria and fungi in rhizosphere and bulk soil. *Iran. J. Biotechnol.* 17 (2), e1982. doi: 10.21859/ijb.1982

Shen, R. F., Cai, H., and Gong, W. H. (2006). Transgenic *Bt* cotton has no apparent effect on enzymatic activities or functional diversity of microbial communities in rhizosphere soil. *Plant Soil* 285, 149–159. doi: 10.1007/s11104-006-9000-z

Shen, Z., Ruan, Y., Chao, X., Zhang, J., Li, R., and Shen, Q. (2015). Rhizosphere microbial community manipulated by 2 years of consecutive biofertilizer application associated with banana Fusarium wilt disease suppression. *Biol. Fert. Soils* 51, 553–562. doi: 10.1007/s00374-015-1002-7

Siciliano, S., Theoret, C., De, F. J., Hucl, P., and Germida, J. (1998). Differences in the microbial communities associated with the roots of different cultivars of canola and wheat. *Can. J. Microbiol.* 44 (9), 844–851. doi: 10.1139/cjm-44-9-844

Smilauer, P., and Jan, L. (2014). *Multivariate analysis of ecological data using CANOCO 5* (New-York, USA: Cambridge university press).

Sultana, F., Dev, W., Xin, M., Han, Y., Feng, L., Lei, Y., et al. (2023). Competition for light interception in different plant canopy characteristics of diverse cotton cultivars. *Genes (Basel).* 14 (2), 364. doi: 10.3390/genes14020364

Sung, D. O., Jang, Y. J., Park, S. Y., Lee, K., Lee, S. K., Yun, D. W., et al. (2021). Assessment of the effect on soil microbial communities of genetically modified soybean and a hybrid from crossing with wild soybean. *Plant Biotechnol. Rep.* 15, 855–862. doi: 10.1007/s11816-021-00710-4

Tran, N. H. T., Oguchi, T., Matsunaga, E., Kawaoka, A., Watanabe, K. N., and Kikuchi, A. (2018). Environmental risk assessment of impacts of transgenic Eucalyptus camaldulensis events highly expressing bacterial Choline Oxidase A gene. *Plant Biotechnol.* 35 (4), 393–397. doi: 10.5511/plantbiotechnology.18.0831a

Velmourougane, K., and Sahu, A. (2013). Impact of transgenic cottons expressing cry1Ac on soil biological attributes. *Plant Soil Environ.* 59 (3), 108–114. doi: 10.17221/616/2012-Pse

Vuong, T. M. D., Zeng, J. Y., and Man, X. L. (2020). Soil fungal and bacterial communities in southern boreal forests of the Greater Khingan Mountains and their relationship with soil properties. *Sci. Rep.* 10 (1), 1–10. doi: 10.1038/s41598-020-79206-0

Walker, T. S., Bais, H. P., Grotewold, E., and Vivanco, J. M. (2003). Root exudation and rhizosphere biology. *Plant Physiol.* 132 (1), 44–51. doi: 10.1104/pp.102.019661

Wu, N., Shi, W., Liu, W., Gao, Z., Han, L., and Wang, X. (2021). Differential impact of *Bt*-transgenic rice plantings on bacterial community in three niches over consecutive years. *Ecotox. Environ. Safe.* 223, 112569. doi: 10.1016/j.ecoenv.2021.112569

Xu, X., Liu, X., Li, F., Hao, C., Sun, H., Yang, S., et al. (2023). Impact of insect-resistant transgenic maize 2A-7 on diversity and dynamics of bacterial communities in rhizosphere soil. *Plants* 12 (10), 2046. doi: 10.3390/plants12102046

Yang, H., Yang, Q., Zhang, D., Wang, J., Cao, T., Bozorov, T. A., et al. (2023). Transcriptome reveals the molecular mechanism of the *ScALDH21* gene from the desert moss syntrichia caninervis conferring resistance to salt stress in cotton. *Int. J. @ Mol. Sci.* 24 (6), 5822. doi: 10.3390/ijms24065822

Yang, H., Zhang, D., Li, X., Li, H., Zhang, D., Lan, H., et al. (2016). Overexpression of *ScALDH21* gene in cotton improves drought tolerance and growth in greenhouse and field conditions. *Mol. Breed.* 36, 1–13. doi: 10.1007/s11032-015-0422-2

Yang, H., Zhang, D., Zhang, D., Bozorov, T. A., Abdullaev, A. A., Wood, A. J., et al. (2019). Overexpression of ALDH21 from Syntrichia caninervis moss in upland cotton enhances fiber quality, boll component traits, and physiological parameters during deficit irrigation. *Crop Sci.* 59 (2), 553–564. doi: 10.2135/cropsci2018.08.0477

Youssef, N., Sheik, C. S., Krumholz, L. R., Najar, F. Z., Roe, B. A., and Elshahed, M. S. (2009). Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl. Environ. Microb.* 75 (16), 5227–5236. doi: 10.1128/Aem.00592-09

Zhang, X., Wang, X., Tang, Q., Li, N., Liu, P., Dong, Y., et al. (2015a). Effects of cultivation of OsrHSA transgenic rice on functional diversity of microbial communities in the soil rhizosphere. *Crop J.* 3 (2), 163–167. doi: 10.1016/j.cj.2014.11.001

Zhang, Y. J., Xie, M., Wu, G., Peng, D. L., and Yu, W. B. (2015b). A 3-year field investigation of impacts of Monsanto's transgenic Bt-cotton NC 33B on rhizosphere microbial communities in northern China. *Appl. Soil Ecol.* 89, 18–24. doi: 10.1016/j.apsoil.2015.01.003

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 11 January 2024
DOI 10.3389/frmbi.2023.1310790

OPEN ACCESS

EDITED BY
Daniel Muller,
UniversitéClaude Bernard Lyon 1,
France

REVIEWED BY
Julia W. Neilson,
University of Arizona, United States
Linnea Katherine Honeker,
University of Arizona, United States

*CORRESPONDENCE
Rebecca C. Mueller
rebecca.mueller@usda.gov

RECEIVED 09 October 2023
ACCEPTED 18 December 2023
PUBLISHED 11 January 2024

CITATION
Goemann HM, Ulrich DEM, Peyton BM, Gallegos-Graves LV and Mueller RC (2024) Severe and mild drought cause distinct phylogenetically linked shifts in the blue grama (*Bouteloua gracilis*) rhizobiome. *Front. Microbiomes* 2:1310790. doi: 10.3389/frmbi.2023.1310790

Severe and mild drought cause distinct phylogenetically linked shifts in the blue grama (*Bouteloua gracilis*) rhizobiome

Hannah M. Goemann[1,2], Danielle E. M. Ulrich[3], Brent M. Peyton[1,4,5], La Verne Gallegos-Graves[6] and Rebecca C. Mueller[1,7]*

[1]Center for Biofilm Engineering, Montana State University, Bozeman, MT, United States, [2]Department of Microbiology and Cell Biology, Montana State University, Bozeman, MT, United States, [3]Department of Ecology, Montana State University, Bozeman, MT, United States, [4]Department of Chemical and Biological Engineering, Montana State University, Bozeman, MT, United States, [5]Thermal Biology Institute, Montana State University, Bozeman, MT, United States, [6]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, United States, [7]United States Department of Agriculture (USDA) Agricultural Research Service, Western Regional Research Center, Albany, CA, United States

Plants rely on a diverse rhizobiome to regulate nutrient acquisition and plant health. With increasing severity and frequency of droughts worldwide due to climate change, untangling the relationships between plants and their rhizobiomes is vital to maintaining agricultural productivity and protecting ecosystem diversity. While some plant physiological responses to drought are generally conserved, patterns of root exudation (release of small metabolites shown to influence microbes) and the consequential effects on the plant rhizobiome can differ widely across plant species under drought. To address this knowledge gap, we conducted a greenhouse study using blue grama (*Bouteloua gracilis*), a drought-tolerant C4 grass native to shortgrass prairie across North American plains, as a model organism to study the effect of increasing drought severity (ambient, mild drought, severe drought) on root exudation and the rhizobiome. Our previous results demonstrated physiological effects of increasing drought severity including an increase in belowground carbon allocation through root exudation and shifts in root exudate composition concurrent with the gradient of drought severity. This work is focused on the rhizobiome community structure using targeted sequencing and found that mild and severe drought resulted in unique shifts in the bacterial + archaeal and fungal communities relative to ambient, non-droughted controls. Specifically, using the change in relative abundance between ambient and drought conditions for each ZOTU as a surrogate for population-scale drought tolerance (e.g., as a response trait), we found that rhizobiome response to drought was non-randomly distributed across the phylogenies of both communities, suggesting that *Planctomycetota*, *Thermoproteota* (formerly *Thaumarchaeota*), and the *Glomeromycota* were the primary clades driving these changes. Correlation analyses indicated weak correlations between droughted community composition and a select few root exudate compounds previously implicated in plant drought responses including pyruvic acid, D-

glucose, and myoinositol. This study demonstrates the variable impacts of drought severity on the composition of the blue grama rhizobiome and provides a platform for hypothesis generation for targeted functional studies of specific taxa involved in plant-microbe drought responses.

KEYWORDS

soil microbiome, next-generation sequencing (NGS), drought, plant-microbe interaction, root exudate

Introduction

Drought is one of the most important abiotic factors limiting plant growth and development (de Vries et al., 2016; Berdugo et al., 2020). Increasing precipitation variability and rising temperatures due to climate change are predicted to increase drought intensity and the rate at which droughts develop (Mukherjee and Mishra, 2021). These changes will likely have dramatic consequences on the productivity and distribution of plant life in both agricultural and natural ecosystems. Understanding plant-microbe interactions in a wide variety of contexts is critical to making informed management decisions in agriculture and natural resource conservation in the face of global climate change.

The plant rhizosphere, the region of soil adjacent to plant roots that is directly influenced by plant growth, respiration, and nutrient exchange, is the epicenter of a complex web of plant-microbe interactions. The rhizosphere hosts a wide diversity of microorganisms including bacteria, archaea, and fungi collectively referred to as the rhizobiome that can improve plant access to nutrients and aid in plant survival under abiotic stress. Mutualistic fungi such as many arbuscular mycorrhizal fungi (AMF) can increase water exchange and nutrient mining to plants under drought stress in exchange for plant C inputs (Mathur et al., 2019). Bacteria including *Streptomyces* sp. (Abbasi et al., 2020; Li et al., 2020) and *Burkholderia* sp. (Naveed et al., 2014; Huang et al., 2017) have been utilized as plant inoculants to protect against drought stress. These and other drought-tolerant bacteria offer benefits to plants including the solubilization of mineral phosphates, nitrogen fixation, and production of plant growth-promoting hormones such as indole acetic acid (Passari et al., 2016; Dahal et al., 2017).

However, it remains difficult to discern broad patterns in rhizobiome community responses to drought stress due to variation in plant species, soil type, and climate. This is partly because we do not fully understand the influence of plant carbon inputs into the soil, namely as root exudates, on the rhizosphere microbiome. Root exudates are small molecules released by plant roots that frequently include sugars, amino acids, and organic acids (Bais et al., 2006). Many of these compounds have been found to have osmoprotectant or chemical signaling properties in various plant systems (Vries et al., 2019; Williams and de Vries, 2020). With recent advances in metabolomics technology, we are only beginning to understand the effects of changes in the quantity and composition of root exudates on rhizobiome community composition and function under increasing drought severity.

The goal of the present study was to examine how increasing drought severity influences the rhizobiome community diversity and composition of *B. gracilis*. Blue grama is a widespread, drought-tolerant C4 grass native to North America that accounts for 75-90% of net primary productivity of grasslands it inhabits and is a major contributor to belowground C storage (Costello, 1944). Blue grama is utilized for conservation, rangeland seeding, landscaping, reclamation, and erosion control and is also valuable for livestock and wildlife because it provides highly palatable forage year-round (Costello, 1944). The level of precipitation in the driest month of the year has historically been a strong influence on the distribution of blue grama across the central and western Great Plains from Canada to northern Mexico (Avendaño-González and Siqueiros-Delgado, 2021). The importance of the blue grama rhizobiome in mediating its drought tolerance is not well understood, although Ulrich et al. (2019) demonstrated that growing blue grama without its microbiome (grown in sterilized soil) resulted in less germination, less biomass, and lower drought tolerance compared to rhizobiome-intact plants. Prior results from this study indicated that increasing drought severity on blue grama grass (*Bouteloua gracilis*) resulted in strong shifts in root exudate composition with compounds such as sucrose, myo-inositol, pyruvic acid, and L-threonine primarily driving differences between drought severity treatments (Ulrich et al., 2022).

Here we further explored the above greenhouse experiment that subjected blue grama seedlings to three levels of drought severity by measuring the diversity and composition of the rhizosphere bacterial + archaeal and fungal communities using next generation amplicon sequencing. We tested two hypotheses: 1. Severe drought will have a stronger effect on the rhizobiome diversity and composition than mild drought and 2. drought-induced changes in root exudate composition would correlate to changes in rhizobiome composition. Understanding how increasing drought severity alters plant-soil feedbacks in drought-tolerant plants will provide important insights into plant-microbe mechanisms that may be applied to less drought-tolerant species (Mathur and Roy, 2021). Disentangling these relationships will also improve our abilities to model global nutrient dynamics and develop microbially enhanced food production and natural resource management solutions in the face of climate change.

Methods

Study site and experimental design

Blue grama was grown from commercial seed (Wind River Seed Co., Manderson, WY, USA) in a greenhouse (Bozeman, MT, USA) in small trays in a 50:50 mix of sand:soil collected from the top soil layer (0-10 cm) from a pinyon-juniper woodland near Bozeman, MT, which is the general ecosystem where blue grama naturally occurs (Coffin and Lauenroth, 1992; Goemann et al., 2024.). This ensured microbes relevant to blue grama would be available (Naylor and Coleman-Derr, 2018). Seeds were planted on October 1, 2019. After 30 days of initial growth, 120 seedlings of similar size were selected and transplanted with the soil from the trays to larger pots (~0.75 L, 6.4 cm diameter, 25.4 cm height) with sterilized playground sand (Ulrich et al., 2019). Greenhouse conditions during the study (Oct-Jan 2019) consisted of a 16-hour photoperiod, daytime temperature of 23.8°C, nighttime temperature of 21.1°C, and average daytime PAR of 500 µmol m^2 s^{-1}. Plants were watered to field capacity and allowed to acclimate to greenhouse conditions for 30 days. Then, established plants were divided into three treatment groups: ambient, mild drought, and severe drought (n = 40 plants per treatment). The ambient treatment group received normal watering to field capacity. The mild drought treatment group was watered half as frequently to maintain 50% of field capacity. The severe drought treatment group was watered 25% as frequently to maintain 25% of field capacity. Treatments lasted 30 days. These drought severity levels were selected based on previous work where blue grama required 33 days of drought to desiccate when water was completely withheld (i.e. 100% reduction in water) (Ulrich et al., 2019). In this study, less severe drought treatment levels (50% and 75% *vs.* 100% reduction in water) were selected to induce plant stress response without killing the plants, allowing measurement of shifts in physiology and root exudate concentration and composition as a function of drought severity.

Plant and soil measurements and collection

Plant physiology analysis is described in full in (Ulrich et al., 2022). Briefly, to determine how drought severity treatment influenced plant physiology, predawn leaf water potential, leaf gas exchange (photosynthesis, stomatal conductance) and root and shoot biomass were measured to determine root to shoot biomass ratios (root:shoot) in each treatment group before (T1) and after (T2) 30 days of treatment (December 19-January 18). Predawn leaf water potential was measured using a Scholander type pressure chamber (model 1505D, PMS Instruments, Corvallis, OR, USA). Gas exchange was measured using a portable photosynthesis instrument equipped with an infrared gas analyzer (LI-6800, Licor, Lincoln, NE USA) on 5 plants randomly selected from each treatment group. Plant physiology results demonstrating the effectiveness the drought treatments are listed in Supplementary Figure S1.

At time of harvest 50 mL each of rhizosphere and bulk soil fractions were collected separately, flash-frozen in an ethanol-dry ice bath and transferred to long-term storage at -80°C for microbial analysis. For rhizosphere collection a soil core was taken as close to the plant as possible directly adjacent to the root crown while the bulk soil core was collected close to the edge of the pot (~3 cm away). Root exudates were collected via vacuum filtration at the time of harvest and transferred to long-term storage at -80°C. Root exudate collection procedures followed the method of Phillips et al. (2008). Briefly, plants were removed from pots and roots were rinsed of soil in DI water and dipped in Gibco Antibiotic-Antimyotic solution (ThermoFisher Scientific CA, USA) to halt the microbial production of C compounds. This ensured that we collected exudates from the plant and not microbes. Plants were subsequently transplanted to filter flasks of 60-120 mesh glass beads (100 mL) that do not provide a C source but still provide mechanical pressure on the root system, resembling natural soil conditions. Plant roots were flushed with 100 mL of sterile water using a vacuum pump connected to the filter flasks, another 100 mL was added, and plants were allowed to release exudates over a 41-hour period. The remaining solution was filtered (0.22 µm) and collected for root exudate concentration and composition analyses. Root exudates were frozen and stored at −80°C until analysis. Exudate samples were submitted to the Environmental Molecular Science Laboratory (EMSL) for gas chromatography-mass spectrometry (GCMS) analysis and data preprocessing. GCMS methods are described in Ulrich et al. (2022).

DNA extraction, amplification, and sequencing

To obtain high-resolution representation of the soil microbial community and its response to drought, we performed rRNA gene amplicon sequencing on all replicates of bulk and rhizosphere soil for all treatment and pretreatment (T1) samples (n=15 each for T1 rhizosphere and bulk soils, n=5 for post-drought (T2) rhizosphere and bulk samples across three treatments, total n=60). Genomic DNA was extracted from 0.5 g frozen soil samples using the Fast DNA SPIN Kit for soil (MP Biomedicals, Santa Ana, CA) according to the manufacturer's instructions. We targeted the V4 region of the 16S rRNA gene using 515F (Parada et al., 2016) and r806b (GTGYCAGCMGCCGCGGTAA and GGACTACVSG GGTATCTAAT, respectively) from the Earth Microbiome Project (Thompson et al., 2017), and the D2 hypervariable region of the fungal large subunit (LSU) using LR22R (Mueller et al., 2016) and LR3. PCR reactions were carried out in 20 µL volumes containing 1 U Phusion high-fidelity polymerase (ThermoFisher Scientific, MA, USA), a final concentration of 1x Phusion HF reaction buffer, 200 µM dNTPs, 0.5 µM each primer, 1 µM BSA, 1 mM MgCl, sterile molecular grade water, and 5 µL (10 ng) genomic DNA as a template. PCR conditions consisted of initial denaturation at 98°C for 30s, followed by 22 cycles of 98°C for 30 s, 55°C for 30 s, 72°C for 45 s and a final extension at 72°C for five minutes. Samples were purified using AMPure XP (Beckman-Coulter, Brea CA, USA) magnetic beads and barcoding was conducted using the Illumina Nextera Indexing Kit D,

with 10 cycles of indexing PCR in 20 µL volumes using the same concentrations as PCR1 with 10 µL template DNA. Samples were purified again with magnetic beads and PCR products were checked for quality and length in a 1% agarose gel and quantitated using the Quant-iT^{TM} dsDNA Kit (Invitrogen, Carlsbad CA, USA) with a BioTek H2 plate reader. Individual samples for each community were combined at equimolar concentrations and the two libraries were pooled at 4 nM and loaded onto an in-house Illumina MiSeq (Illumina, San Diego CA, USA) and sequencing using the V3 600 cycle kit.

Bioinformatics

Reads were merged, trimmed, and dereplicated with USEARCH, quality filtered with an ee-value of 1.0 and zero-radius operational taxonomic units (ZOTUs) were identified with UNOISE3 (v.11.0.667) (Edgar, 2010). A total of 4,668,982 16S and 4,872,815 LSU high quality reads were clustered into 15,541 and 6,893 ZOTUs respectively. Representative 16S and LSU ZOTUs were classified against their respective databases using a subset of the Genome Taxonomy Database (GTDB, Chaumeil et al., 2022) for 16S and the Ribosomal Database Project (RDP, Cole et al., 2014) v11 database for LSU using SINTAX (Edgar, 2016). All reads classified to chloroplast (16S), or animalia, protozoa, or viridiplantae (LSU) were removed prior to downstream analysis. For statistical analyses, datasets were randomly subsampled according to the sample with the lowest read number. To construct phylogenetic trees for environmental ZOTUs, we used reference phylogenies constructed using full length and near full length sequences of isolates and metagenome assemblies downloaded from GenBank, focusing on GTDB sequences for 16S and Assembling the Fungal Tree of Life sequences for LSU (Lutzoni et al., 2004). Reference sequences were aligned using mafft (v.7.250, Katoh and Standley, 2013), and a maximum likelihood tree was constructed with the general time reversible (GTR) + gamma model using RAxML (v7.2.7 Stamatakis, 2014). Environmental sequences were aligned to the reference using mafft with the –add flag and then mapped onto the reference tree using placer (v1.1.alpha19-0-g807fbf3, Matsen et al., 2010) with reference-aligned sequences. Trees were visualized and annotated using the interactive Tree of Life software (v6.8.1, Letunic and Bork, 2019).

Statistics and data analysis

Community analyses were conducted using R software (v4.3.0, R, 2023) in RStudio (RStudio Team, 2020). Data management was handled primarily with the tidyverse package (Wickham et al., 2019). Community ordination and alpha diversity analyses were performed using the packages vegan (Oksanen et al., 2019), picante (Kembel et al., 2010) and phyloseq (McMurdie and Holmes, 2013). We used PERMANOVA (Anderson, 2017) to test for treatment, timepoint, and bulk *vs.* rhizosphere effects on community beta diversity computed using unweighted UniFrac (Lozupone and Knight, 2005). To explore the effect of time on community composition we conducted Wilcoxen Signed Rank Tests to test differences in relative abundances between T1 (pretreatment) and ambient samples at T2 (post-drought).

To examine whether ZOTU responses to drought were non-randomly distributed across the bacterial + archaeal and fungal phylogenies in the blue grama rhizosphere, we performed phylosignal analysis using response ratios as a response trait on T2 (post-drought) rhizosphere samples. Response ratios were computed by dividing the difference between the read counts in the treatment (severe or mild drought) and ambient conditions by the total read count for a ZOTU. Response ratios were permuted 999 times for each ZOTU and averaged across 5 replicate samples in each treatment. Phylosignal analyses were performed using phytools (Revell, 2012) and geiger (Pennell et al., 2014) for Pagel's lambda and Blomberg's K (Diniz-Filho et al., 2012). Packages phylobase (Hackathon et al., 2020) and adephylo (Jombart et al., 2010) were utilized for Abouheif's C_{mean}. One-way t-tests were then conducted across taxonomic levels to determine which clades had response ratios significantly different from 0. All t-tests were adjusted with Benjamini-Hochberg (BH) multiple comparisons corrections, and we report the adjusted p-values. To aid in visualization of the phylogenetic signal analysis, the response ratio for each ZOTU and for clades with significantly positive or negative clade-wide response ratios were mapped back to the phylogenetic trees using the color strip tree annotation template provided by the interactive Tree of Life software (Letunic and Bork, 2019).

Having identified several clades of bacteria, archaea, and fungi which had significant clade-wide response to severe drought, we performed network analysis to determine whether there were any detectable inter-kingdom (i.e. bacterial-fungal) associations in relative abundances between drought and ambient conditions. Sparcc networks were constructed and analyzed using the NetCoMi package (Peschel et al., 2021) in RStudio with specific parameters listed in the Supplementary Information. We utilized the differential network analysis module in the NetCoMi package to determine the effects of mild and severe drought on the top 50 of these potential associations. Differential network analysis uses permutation tests to assess whether associations between taxa are different between two user-defined treatment groups, in our case defined as mild drought *vs.* ambient and severe drought *vs.* ambient conditions (Peschel et al., 2021). All figures were edited for publishing quality in Adobe Illustrator 2023 27.5.0.

To calculate correlations between root exudation and community composition we selected root exudate compounds identified as driving the differences between treatment groups (ambient, mild drought, and severe drought) reported Ulrich et al. (2022). Correlation analysis was performed using the envfit function in the vegan package. Envfit correlation p-values were also BH-corrected for multiple comparisons.

Results

Drought effects on rhizobiome composition

We found severe drought resulted in significantly altered beta diversity of the bacterial + archaeal and fungal communities compared to the ambient treatment in T2 (Bacteria + archaea: $F_2 = 1.41$, $R^2 = 0.0466$, $p = 0.01$, Fungi: $F_2 = 1.34$, $R^2 = 0.183$, $p = 0.025$, Figure 1). In comparison, mild drought conditions had minimal effects on the rhizobiome beta diversity. We also detected decreased dispersion in the bacterial + archaeal community ($F_2 = 4.81$, $p = 0.029$) in the severe drought condition compared to ambient. We did not find evidence for differences in alpha diversity between drought conditions (mild or severe) and ambient conditions or between mild and severe drought, in either the bacterial + archaeal or fungal communities across any of the three metrics tested: species richness, Shannon's diversity, or Faith's phylogenetic diversity (Figure 2; Supplementary Table S1), although we did detect a positive effect of time (T2 > T1) across all three metrics in the fungal community (Supplementary Table S1).

Beta diversity also differed between T1 and ambient conditions at T2 in both the bacterial + archaeal ($F_2 = 1.64$, $R^2 = 0.0271$, $p = 0.007$, Table 1) and fungal communities ($F_2 = 1.59$, $R^2 = 0.0264$, $p = 0.059$, Table 1), although the difference between T1 and T2 ambient conditions in both communities was slight. The differences between T1 and ambient-T2 in the bacteria + archaea were likely driven by the detection of *Armatimonadota* and *Nitrospirota* above 1% abundance only at T1, and detection of *Fibrobacterota* only in ambient-T2 samples (Supplementary Figure S2A). We did not detect any significant differences in relative abundances of fungal classes between T1 and ambient-T2 (Supplementary Figure S2B). We also did not detect differences between T1 and T2 in beta diversity between bulk and rhizosphere samples, or effects of any interactions between the combinations of soil compartment, time, and treatment (Table 1).

Distribution of drought response across the rhizobiome phylogeny

Phylosignal analysis using response ratios as a response trait on the bacterial + archaeal and fungal trees indicated that changes in relative abundance between the mild drought and ambient treatments as well as the severe drought and ambient treatments were not randomly distributed across either tree (Table 2). In the bacterial + archaeal community, all three phylosignal tests (Pagel's Lambda, Blomberg's K, and Abouheif's C_{mean}) rejected the null hypothesis of random distribution of the response ratio trait. In the fungal community for both the mild drought-ambient and severe drought-ambient comparisons, Blomberg's K test did not reject the null hypothesis but the other two tests were found to be significant.

In the mild drought–ambient comparison, global analysis of response ratios detected one archaeal phylum, four bacterial phyla, and one fungal class which had clade-wide response ratios significantly different from zero (i.e., increased or decreased relative abundance in drought *vs.* ambient conditions). *Thermoproteota* (archaea also known as *Thaumarchaeota*), and bacterial *Chloroflexota* and *Actinobacteriota* had positive phylum-wide response ratios (greater relative abundance in mild drought *vs.* ambient conditions) while bacterial *Omnitrophota* and *Planctomycetota* had negative phylum-wide response ratios (Supplementary Table S2). The fungal class *Eurotiomycetes* within the *Ascomycota* was the only fungal clade detected and had a negative class-wide response ratio in the mild drought-ambient conditions comparison. At the family level for bacteria + archaea,

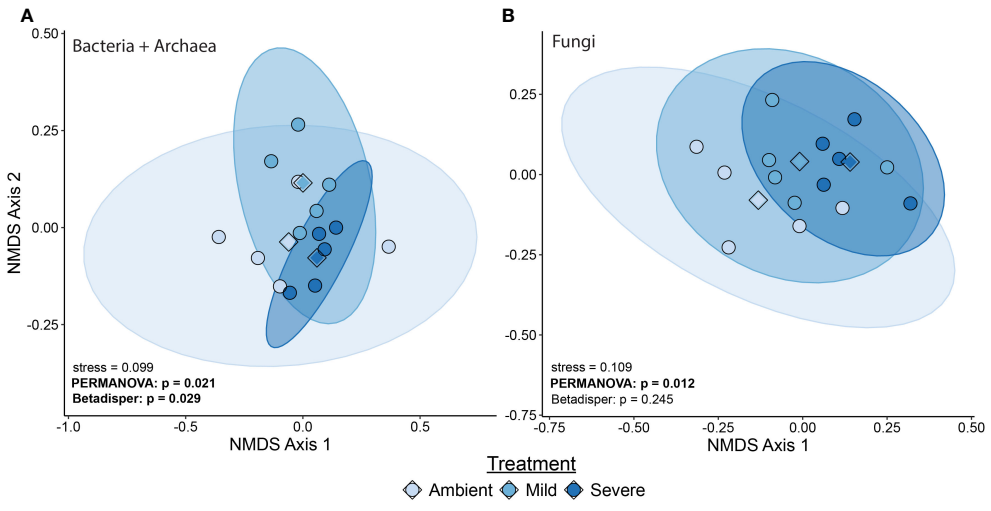


FIGURE 1
Beta diversity measured with unweighted Unifrac of bacterial/archaeal community **(A)** and fungal community **(B)** at T2. Points represent individual samples with a larger diamond representing the treatment centroid. Ellipses indicate the 95th percentile area of community distribution. Legend shows symbol and ellipsis color.

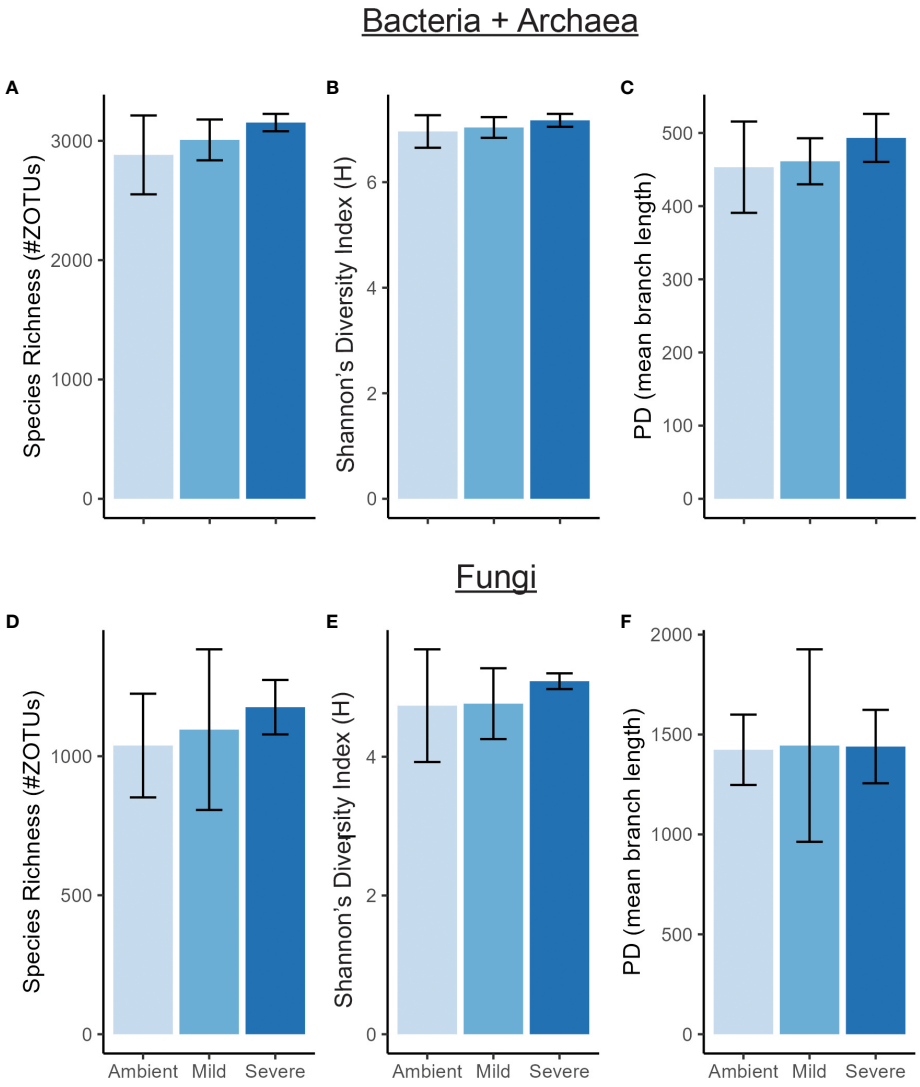


FIGURE 2
Alpha diversity of bacterial/archaeal community **(A–C)** and fungal community **(D–F)** measured by species richness, Shannon's diversity index (H), and Faith's Phylogenetic Diversity (PD) at T2. Error bars are ± SE.

we identified 15 families across the phylogeny with positive response ratios, and six with negative response ratios (Figure 3A inner circles, Supplementary Table S2). Most of the families with positive response ratios belonged to the *Actinobacteriota* phylum, including: *Rubrobacteraceae*, *Solirubrobacteraceae*, *Gaiellaceae*, *Illumatobacteraceae*, *Geodermatophilacaeae*, *Microbacteriaceae*, *Propionibacteriaceae*, and *UB11606* (order *Acidimicrobiales*). The other clades with positive response ratios included three families within the *Alphaproteobacteria* (*Devosiaceae*, *Xanthobacteraceae*, and *Beijerinckiacaeae*), *Roseiflexiaceae* (*Chloroflexota*), one *Acidobacteriota* family, *UBA5704 (Thermoanaerobaculia)*, and the only family present within the *Thermoproteota*, *Nitrososphaeraceae*. Negative family-wide responses included: *Bryobacteraceae* (*Acidobacteriota*), *Nevskiaceae* and *Burkholderiaceae* (*Gammaproteobacteria*), *SM1A02* and *PALSA_1355* (*Planctomycetota*) and *J027* (*Chloroflexota*).

In the severe drought-ambient comparison, we detected many more bacterial + archaeal clades with significantly positive or negative response ratios (11 phyla and 45 classes/families *vs.* 5 phyla and 22 classes/families in the mild drought-ambient comparison), with only a few overlapping between the two comparisons (Figure 3A outer circles, Supplementary Table S2). At the phylum level, the *Thermoproteota* and *Chloroflexota* maintained positive response ratios, while the *Planctomycetota* flipped from negative to positive. We also detected positive response ratios in the *Verrucomicrobiota* and *Acidobacteriota* which were not significant in the mild drought-response ratios. Likewise, the *Firmicutes*, *Gemmatimonadota*, and *Bdellovibrionota* were found to have negative response ratios only in the severe drought-ambient comparison. While the *Gemmatimonadota* have been associated with drought-resistance in some studies (Guo et al., 2023; Wang et al., 2022) there is evidence that members of this

TABLE 1 Summary of beta diversity test statistics (PERMANOVA).

	Variable	Df	SumofSqs	R^2	F	P-value
Bacteria + Archaea	Treatment (Drought)	2	0.471	0.0466	1.41	**0.01**
	Time	1	0.274	0.0271	1.64	**0.007**
	Microhabitat	1	0.173	0.0161	1.04	0.289
	Treatment : Time	2	0.383	0.0379	1.15	0.114
	Treatment : Microhabitat	2	0.333	0.0330	1.00	0.405
	Time : Microhabitat	1	0.160	0.0156	0.958	0.482
	Treatment : Time:Microhabitat	2	0.321	0.0318	0.965	0.553
	Beta dispersion	2	0.00336	NA	4.81	**0.029**
Fungi	Treatment (Drought)	2	0.370	0.1832	1.34	**0.025**
	Time	1	0.1706	0.02635	1.59	*0.059*
	Microhabitat	1	0.0803	0.01239	0.746	0.772
	Treatment : Time	2	0.2077	0.0321	0.965	0.477
	Treatment: Microhabitat	2	0.2412	0.0373	1.12	0.258
	Time: Microhabitat	1	0.078	0.0120	0.725	0.815
	Treatment : Time: Microhabitat	2	0.2633	0.0406	1.22	0.151
	Beta dispersion	2	0.001022	0.000511	0.766	0.486

Bold letters indicate significance at $p < 0.05$.

phylum are sensitive to dry-wet cycling (Qi et al., 2022) which the droughted samples experienced in this study. Although the *Actinobacteriota* did not have a significant phylum-wide response ratio in the severe drought-ambient comparison, we did detect positive response ratios in the *Rubrobacteraceae*, *Gaiellaceae*, and *IMCC26256* (order *Acidimicrobiia*) families within the *Actinobacteriota* as in the mild drought-ambient comparison. Unique to the severe drought-ambient comparison, actinobacteriotal family *Steptomycetaceae* had a negative response ratio. Within the *Proteobacteria* we observed positive response ratios in both drought treatments in the *Devosiaceae* and *Xanthobacteraceae*, and negative response ratios in both treatments in the *Burkholderiaceae* and *Nevskiaceae*. *Alphaproteobacteria* families *Sphingomonadaceae* and *Rhodobacteraceae* had negative response ratios unique to the severe drought-ambient comparison, and we observed a positive response ratio in the *Steroidobacteraceae* family in the severe drought-ambient comparison and a positive response in the *Beijerinckiaceae* family only in the mild drought-ambient comparison.

In the fungal community, severe drought also had a greater impact on response ratios than mild drought compared to ambient conditions. With the severe drought-ambient comparison we detected positive response ratios in the *Mucoromycotina* and

TABLE 2 Summary of phylosignal test results.

Test	Mild Drought		Severe Drought	
	Stat	P-value	Stat	P-value
	Bacteria + Archaea			
Pagel's Lambda	0.702	**2.21e-144**	0.701	**1.26e-144**
Blomberg's K	0.0270	**0.001**	0.0179	**0.011**
Abouheif's C_{mean}	0.107	**0.001**	0.103	**0.001**
	Fungi			
Pagel's Lambda	0.793	**1.51e-35**	0.808	**2.12e-51**
Blomberg's K	0.0267	0.157	0.0287	0.134
Abouheif's C_{mean}	0.132	**0.001**	0.198	**0.001**

Bold lettering indicates significance at $p < 0.05$.

Ascomycota phyla, and negative response ratio in the *Chytridiomycota*. We detected 4 classes which had positive response ratios including: *Glomeromycetes*, *Sordariomycetes*, *Dothideomycetes*, and the *Eurotiomycetes* (Figure 3B). The *Chytridiomycetes* were the only class with a clade-wide negative response ratio. Within the *Glomeromycetes* class there was a notable sub-clade, BLAST-identified as *Funneliformis* sp., which had negative response ratios although these did not influence the clade-wide positive response.

Impact of drought severity on rhizobiome associations

To determine the effects of drought severity on community associations, we performed network analyses using differential networks of mild drought-ambient and severe drought-ambient communities. In the mild drought-ambient differential network, we identified a number of bacterial-fungal associations affected by drought (Figure 4A). *Acidobacteriotal* class *Vicinamibacteria*,

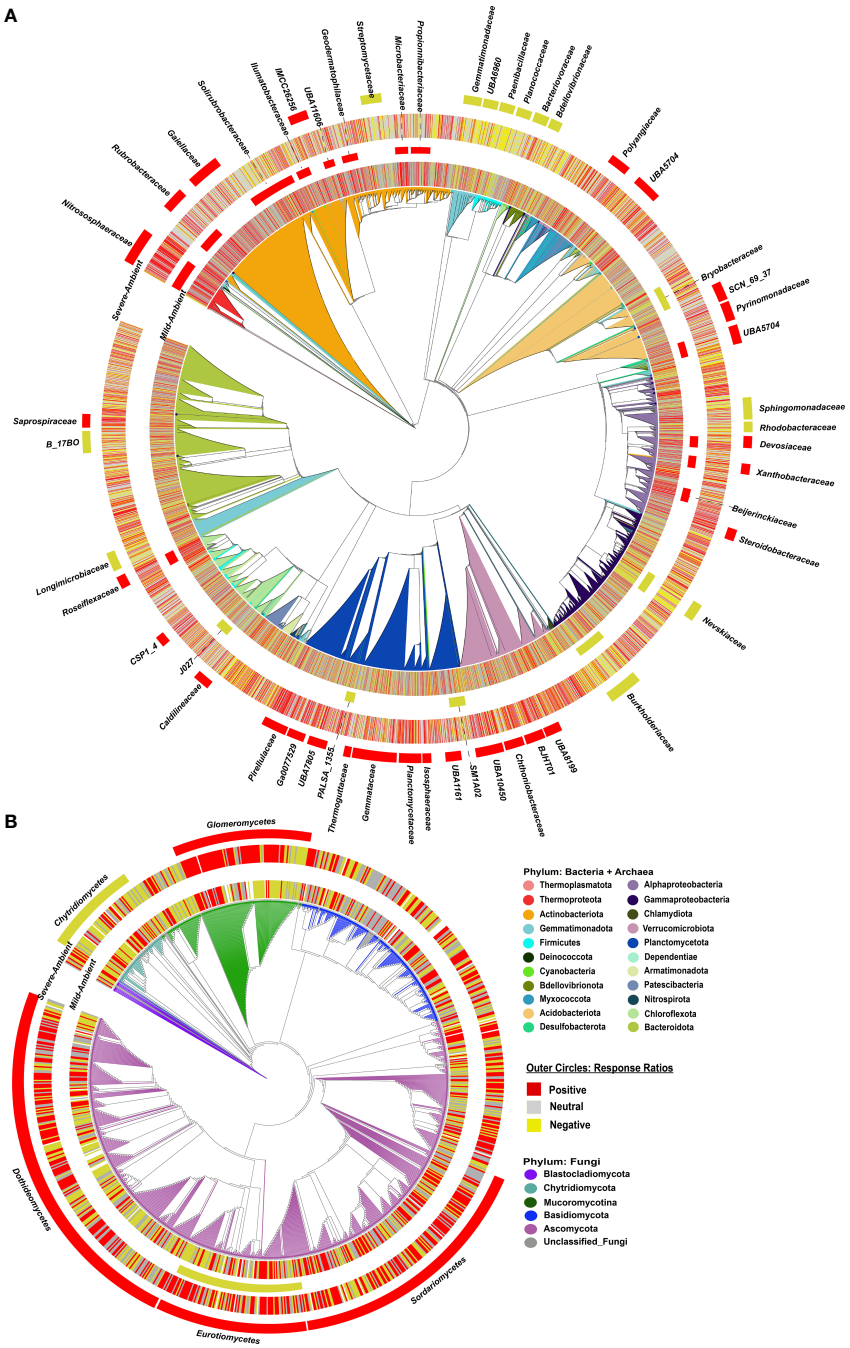


FIGURE 3
Phylogenetic trees of bacterial + archaeal **(A)** and fungal **(B)** communities, colored by phyla. The 1st ring for each the Mild Drought-Ambient comparison (inner circles) and Severe Drought-Ambient comparison (outer circles) indicates positive (red), neutral (gray) or negative (yellow) response ratios at the ZOTU level, then families for the bacteria + archaea and classes for fungi with significantly positive (red) or negative (yellow) response ratios are indicated by red (positive) or yellow (negative) colored strips for both drought-ambient comparisons.

Actinobacteriota class *Thermoleophilia*, and *Chloroflexota* classes *Ktdedonobacteria* and *Ellin6529* were the most connected classes in the network. The fungal class *Leotiomycetes* appeared to have several associations with bacterial classes that were negative under ambient conditions but positive under mild drought conditions (pink lines) including associations with *Ktedonobacteria*, *Thermoleophilia*, *Actinomycetia*, *Rubrobacteria*, and *Phycisphaerae*. In contrast, bacterial class *Actinomycetia* had several associations that were positive under ambient conditions but negative under mild drought (blue lines) including associations with *Acidimicrobia*, *Alphaproteobacteria*, *Gammaproteobacteria*, and *Saccharimonadia*. Some classes such as *Binatia*, *Planctomycetes*, and *Blastocatellia* had only positive associations with other taxa in both ambient conditions (green lines). Several fungal classes had primarily negative associations regardless of treatment group (black lines) including the *Glomeromycetes*, *Pezizomycetes*, *Dothideomycetes*, and *Sordariomycetes*.

In the severe drought-ambient differential network, *Planctomycetes* was the most connected class and appeared to have strong tradeoffs of positive associations in the ambient *vs*. severe drought communities. *Planctomycetes* associated negatively with ten classes in the ambient condition that were positive associations in the severe drought condition (purple lines) including *Actinomycetia*, *Blastocatellia*, *Ellin* 6529, *Gemmatimonadetes*, *Ktedonobacteria, Rubrobacteria*, Thermoleophilia, and fungal classes *Ascomycota incertae sedis, Eurotiomycetes*, and *Leotiomycetes*. In contrast, *Planctomycetes* were negatively associated with a different set of seven classes in the severe drought condition that were positive associations in ambient condition. These included *Alphaproteobacteria*, *Bacilli*, *Bdellovibrionata*, *Chlamydiia*, *Gammaproteobacteria*, *Thermoanaerobaculia*, and *Zygomycota incertae sedis Nitrososphaeraceae* was always positively associated with *Planctomycetes* while the majority of the fungal classes including the *Tremellomycetes*, *Agaricomycetes*, *Microbotryomycetes*, *Chytridiomycetes*, *Exobasidiomycetes*, and *Dothidiomycetes* were always negatively associated with *Planctomycetes* regardless of treatment.

Connections to plant physiology & root exudation

Previous results focusing on changes in blue grama root exudation in response to increasing drought severity found significantly increased total organic carbon in root exudate extracts under severe drought compared to ambient conditions (Ulrich et al., 2022). There were ten root exudate compounds driving the greatest differences in exudate composition across the three treatments (ambient, mild drought and severe drought). These compounds were sucrose, D-arabinose, pyruvic acid, D-mannose, sedoheptulose anhydride monohydrate, tagatose, D-glucose, L-threonine, and 4-guanidinobutyric acid (Ulrich et al., 2022). Here we detected weak evidence for a correlation between the composition of these root exudates across samples with the rhizosphere bacterial + archaeal (Mantel R^2 = 0.299, p = 0.084) and fungal (Mantel R^2 = 0.289, p = 0.087) community compositions (beta diversity). In the bacterial + archaeal community, the best possible model correlating root exudates to community composition included only myoinositol and D-arabinose (Mantel R^2 = 0.269, p=0.081). For the fungal community, the best possible model included pyruvic acid, D-mannose, tagatose, and D-glucose (Mantel R^2 = 0.296, p = 0.083). Pearson correlations between taxa with significant clade-wide response ratios (i.e., those indicated by Figure 3) and root exudates were also conducted but no significant correlations were detected.

Discussion

Severe drought had stronger effects on the blue grama rhizobiome than mild drought

In support of our first hypothesis, severe drought had a greater impact on the blue grama rhizobiome composition than mild drought compared to ambient, non-droughted controls. Shifts in overall composition of the bacterial + archaeal and fungal communities were detected, and a decrease in dispersion (e.g., reduced beta diversity) among bacterial + archaeal community members was observed during severe drought conditions compared to the ambient treatment (Figure 1). In contrast, mild drought did not induce a significant shift in beta diversity among either the bacteria + archaea or fungi (Figure 1). The greater impact of severe drought on the soil rhizobiome compared to mild (or moderate) drought has been hypothesized in modeling efforts such as demonstrated by Wang and Allison (2021). Simmons et al. (2020) also demonstrated that severe drought had a greater impact than moderate drought on rhizosphere bacterial community composition across multiple species of millet.

The lack of drought treatment effects (either mild or severe) on alpha diversity in either the bacterial + archaeal or fungal communities (Figure 2) indicated that the changes in community composition were primarily due to shifts in relative abundances of existing members. Importantly, greenhouse studies limit dispersal which can have significant effects on community diversity under abiotic stress (Evans et al., 2020). However, the controlled conditions of greenhouse studies allow for treatment effect isolation and deliberately limiting dispersal effects improves the ability to detect shifts more directly in existing community membership. Phylogenetic signal is considered a useful metric for studying the effects of drought on the soil microbiome as many clades at different taxonomic levels tend to respond roughly in unison (Naylor et al., 2017; Koyama et al., 2018; Xu et al., 2018). Accordingly, phylogenetic signal in comparisons of both mild drought-ambient and severe drought-ambient rhizobiomes were detected using response ratios of the differences in relative abundances of taxa between treatment groups as a response trait (Figure 3).

Several clades of microorganisms have been found to be resilient to drought, including the *Actinobacteriota*, *Firmicutes*, *Mucoromycotina* (which includes the *Glomeromycetes*), and various subclades among the *Ascomycota* (Xu et al., 2018; Santos-

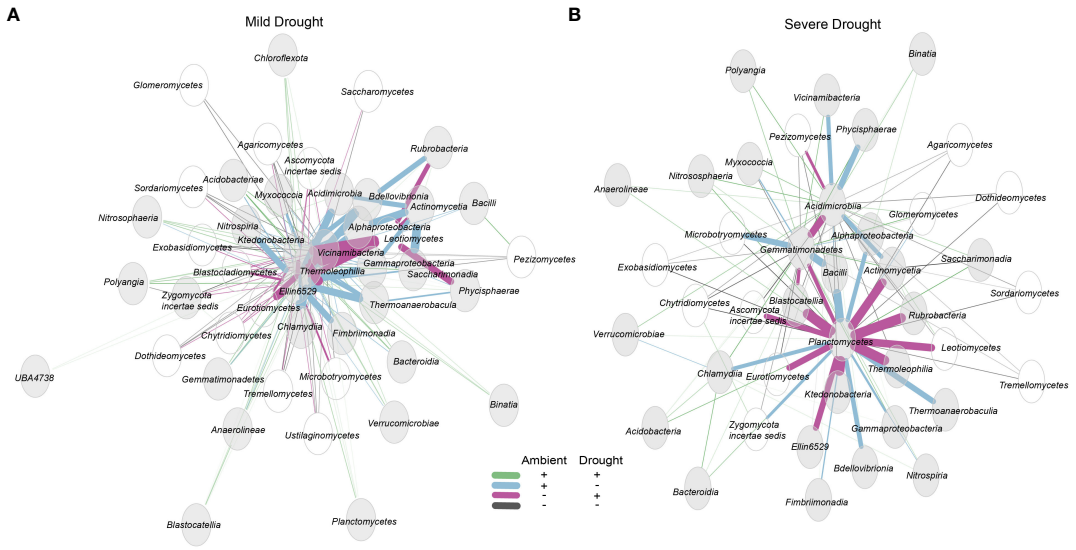


FIGURE 4
Sparcc differential network visualizing the top 50 associations of the total microbiome community at the class level under **(A)** mild drought vs. ambient conditions, and **(B)** severe drought vs. ambient conditions. Bacterial and archaeal nodes are gray while fungal nodes are unfilled. Line thickness indicates strength of association. All associations are significant with BH-adjusted $p < 0.05$.

Medellín et al., 2021). Consistently, we observed positive clade-wide responses (i.e. increased relative abundance in severe drought *vs.* ambient conditions) in all of these phyla with the exception of the *Firmicutes*, which in our study exhibited a clade-wide negative response in severe drought (i.e. decreased relative abundance *vs.* ambient conditions) but not mild drought. Overall, severe drought resulted in a greater number of significant response ratio changes (either positive or negative) across the bacterial + archaeal and fungal phylogenies than mild drought. For the fungal community in particular, the lack of significant response under mild drought conditions may reflect the greater water stress tolerance amongst mycelial community members (Santos-Medellín et al., 2021). Plant-microbe interactions are often the most beneficial when they help mediate access to the most limiting nutrient in an environment, such as water during drought stress (Ruth et al., 2011). Due to the lack of strong compositional effects in mild drought conditions, we hypothesize that blue grama may be able to self-regulate and tolerate mild drought. However, the greater investment in root exudation observed under severe drought stress (Ulrich et al., 2022) may be a mechanism to stimulate microbial mutualisms (i.e. the positive response ratios across the mycorrhizal *Glomeromycetes*) to help alleviate severe drought stress (Bécard and Piché, 1989; Vries et al., 2019).

Within the *Actinobacteriota*, more positive response ratios were detected in the mild drought than severe drought compared to ambient conditions, although some families such as *Rubrobacter*, which are known to be highly tolerant to desiccation (Naylor et al., 2017; Tóth et al., 2017), had positive response ratios in both drought severity treatments. Interestingly, while the actinobacteriotal genus *Streptomyces* has been found to alleviate drought stress in multiple agricultural plant species including tomatoes (Abbasi et al., 2020), maize (Warrad et al., 2020), and wheat (Li et al., 2020), here we observed a decrease in relative abundance of *Streptomycetaceae* in response to severe drought compared to ambient. Likewise, while we detected decreased relative abundances of *Burkholderiaceae (Alphaproteobacteria)* in both mild and severe drought compared to ambient, *Burkholderia* sp. have been used as plant inoculants to successfully protect against drought stress in *Arabidopsis*, maize, wheat, tomato, and bell peppers (Naveed et al., 2014; Tallapragada et al., 2016; Huang et al., 2017). Importantly, inoculants are most commonly applied to annual, domesticated crops with the exception of some trees (Umashankar et al., 2012; Bizos et al., 2020). Wild, perennial plant species such as blue grama typically have greater genetic diversity and often greater natural drought tolerance than cultivated species (Budak et al., 2013) which likely results in unique drought response patterns in the rhizobiome. The results presented here may also serve to broaden the known range of drought response patterns across plant species.

While the majority of the *Mucoromycotina* had positive response ratios in the severe drought-ambient comparison, we observed a distinct sub-group of *Funneliformis* sp. within the *Glomeromycetes* which had negative response ratios. This negative response is surprising as *Funneliformis* sp. are arbuscular mycorrhizal fungi which have been utilized as inoculants for several plants including thyme, soybeans and groundcherries to improve plant growth during water stress (Reyes et al., 2019; Amani Machiani et al., 2021). However, recent studies of bacterial colonization of mycorrhizal spores (Schrey et al., 2012; Lasudee et al., 2018) have identified *Streptomyces* sp. as among the most common endophytic mycorrhizal colonizers, and of *Funneliformis* sp. specifically. A tight association between *Funneliformis* sp. and *Streptomyces* sp. may explain the negative response ratio we observed in the *Streptomycetaceae* in the severe drought-ambient comparison. These results are also further evidence that drought-tolerant blue grama may not rely on the same drought response mechanisms that have been observed in other plant species.

Insights on inter-kingdom associations via network analysis

Network analysis enables the identification of co-occurrence patterns between microbial taxa which may provide clues to shared niche spaces, nutrient acquisition strategies, or even direct symbioses (Barberán et al., 2012). These analyses provide an important foundation for hypothesis generation through which future experiments can be designed to determine the extent and conditionality of associations between taxa. Here we observed several distinct network patterns between mild drought-ambient and severe drought-ambient networks (Figure 4).

In the severe drought-ambient differential network the *Planctomycetes* was the bacterial class most highly connected to other bacterial and fungal classes. There were ten negative associations with *Planctomycetes* in ambient conditions which were positive associations in severe drought conditions, seven associations which were positive in ambient conditions but negative in severe drought conditions, three that were always positive regardless of treatment, and seven which were always negative regardless of treatment conditions. *Planctomycetota* are among the most commonly detected clades in aquatic and terrestrial ecosystems (Delgado-Baquerizo et al., 2018; Wiegand et al., 2018), but due to their low cultivability, little is known about the extent of their metabolic capabilities and influence on nutrient cycling. Wang et al. (2015) reported the ability of *Planctomycetota* species to degrade complex heteropolysaccharides in soil. Elevated or maintained relative abundance of *Planctomycetota* in response to drought has also been observed in several studies (Sheik et al., 2011; Dai et al., 2019). In addition, some *Planctomycetota* have been found to display antifungal properties. Graça et al. (2016) explored 40 *Planctomycetota* isolates to identify bioactive molecules and found that 43% of their study group contained genetic markers for antifungal activity which aligns with our finding that many fungal classes were negatively associated with *Planctomycetes* regardless of treatment condition. Although this apparent relationship was not influenced by drought and was not detected at all in mild drought conditions, it is additional evidence of possible functional attributes of a largely uncharacterized clade of microorganisms that warrants further investigation.

The mild drought-ambient differential network was largely unique from the severe drought-ambient network. Here, *Planctomycetes* only exhibited positive associations with the most connected taxa. These most connected taxa included *Thermoleophilia* (*Actinobacteriota*), *Ktedonobacteria* (*Chloroflexota*), and *Vicinamibacteria* (*Acidobacteriota*). *Ktedonobacteria* have been identified as a keystone specialist group in tundra ecosystems with positive correlations to lower moisture content (Wong et al., 2023) and while much of their metabolic capability remains undetermined, Zheng et al. (2019) observed strong functionality related to plant decomposition through cellulolytic activity. Although *Vicinamibacteria* were not detected in the mild drought-ambient comparison of our response ratio analysis, closely related *Bryobacteraceae* exhibited a negative response, and family *SCN_69_37* of the *Vicinamibacteria* did have a positive response ratio in the severe drought-ambient condition. *Vicinamibacteria* have also previously been reported as highly drought tolerant (Huber et al., 2022) although their mechanism of survival is not yet understood.

The differential networks were further utilized to investigate possible inter-kingdom associations of interest. We hypothesized that the observed clade-wide positive response observed in the response ratios of the *Thermoproteota* (ammonia-oxidizing archaea) under severe drought may be linked to the concurrent increase in *Glomeromycetes* (arbuscular mycorrhizal fungi) relative abundance via turnover of organic N-rich root exudates such as urea, uracil and various amino acids that were detected under the severe drought conditions (Ulrich et al., 2022). Arbuscular mycorrhizal fungi (AMF) are well known plant growth promoters and can help alleviate drought stress through hyphal mining for water and nutrients in the bulk soil adjacent to plant roots (Wu et al., 2013; Mathur et al., 2019). However, due to low saprotrophic ability, AMF proliferation is highly reliant on other soil microbes to degrade organic compounds and release available nitrogen (Bukovská et al., 2016). Despite exhibiting strong phyla-wide positive response ratios in severe drought conditions, network analysis did not detect any direct associations between *Thermoproteota* and *Mucoromycotina*. Importantly, although sparcc network statistics are compositionally aware, users are required to manually set thresholds for taxa inclusion that may bias networks towards more highly abundant taxa (MatChado et al., 2021). Members of *Thermoproteota* were among the most abundant ZOTUs so associations with the relatively rarer *Mucoromycotina* may exist but were not detected in the sparcc analysis. Future work would also benefit from the inclusion of non-plant controls to disentangle plant-specific microbiome responses distinct from responses of the soil microbiome alone to drought (Bandopadhyay et al., 2023).

Lack of difference in rhizosphere *vs*. bulk samples suggests a wider scope of the rhizosphere zone of influence

Differentiating rhizosphere from bulk soil is common practice in soil microbial ecology studies to determine the influence of plants on the soil microbiome, and differences in community composition between rhizosphere and bulk soils have been well-documented (Castellano-Hinojosa et al., 2021; Ling et al., 2022). However, here we did not detect any significant differences between rhizosphere and bulk soil samples. We hypothesize that the small diameter of our pots (6.4 cm) did not allow for sufficient coring distance from plant roots to make this distinction as the roots largely filled the available space. This finding challenges a common working definition of rhizosphere soil as that which is specifically attached to roots (Chaparro et al., 2014; Zhu et al., 2020; Shami et al., 2022). Kuzyakov and Razavi (2019) reviewed the temporal dynamics and spatial stationarity of plant rhizospheres and found that the rhizosphere generally extends 0.5-4 mm from plant roots. Due to the 4-dimensional (3D-space + 1D-time), fibrous, high-density and

high turnover rate of root structure of most grasses including *B. gracilis* (Ravenek et al., 2016), it is very possible that the vast majority of the soil within our pots were within this zone of influence. In a recent global meta-analysis of rhizosphere *vs.* bulk soil microbiomes, Ling et al. (2022) also found minimal differences in alpha diversity between bulk and rhizosphere soils in grasslands. Since plant species, soil type and texture, nutrient availability and microbiome activity all influence plant-microbe interactions (Jiang et al., 2017; Ling et al., 2022), it follows that the size of the zone of influence of the rhizosphere would be similarly variable.

Root exudates weakly correlate to microbiome composition

In partial support of our second hypothesis, potential exudate compounds of interest which were primarily sugars and organic acids weakly correlated with the microbiome composition metrics. One possible compound of interest is pyruvic acid which is typically in high abundance for normal functioning photosynthesis, glycolysis, and the TCA cycle. Maurer et al. (2021) found high levels of pyruvate and amino acid exudation that had stimulatory effects on denitrification enzyme activity. Interestingly, while some compounds such as urea, arbutin, L-proline, and D-gluconic acid were only detected in the mild and severe drought treatment root exudates (Ulrich et al., 2022), no correlations between these compounds and shifts in the microbial community composition were detected here. As suggested by Bandopadhyay et al. (2023), perennial grass species may not be as reliant on stress signaling to microbes via root exudation due to their relatively greater investment in belowground root architecture compared to many annual plant species. Another possibility, as suggested by Karlowsky et al. (2018), is that root exudation during drought may be more important for the reinitiation of soil microbial activity following rewetting of the system. The authors suggest that further studies incorporating recovery from drought are needed to test this hypothesis across different soils, plants, and environments. In addition, utilizing metatranscriptomics to measure differential expression of metabolic pathways in the rhizobiome under drought will also be an important step in linking plant-microbe responses to drought. This would bypass the limitations of DNA-based analyses which cannot distinguish the contribution of active microbes to ecosystem processes (Carini et al., 2016). Metatranscriptomics would also provide evidence for potential utilization of root exudates. Experiments could then be designed to explore specific relationships between active microbial taxa and plant physiology.

Conclusion

Here we show that the blue grama rhizobiome underwent significant community shifts as a result of severe drought and was only marginally affected by mild drought. We detected phylogenetically linked tradeoffs in community composition as a result of severe drought conditions including increased relative abundances of arbuscular mycorrhizae (*Glomeromycetes*), ammonia oxidizing archaea (*Thermoproteota*), and *Planctomycetota*. The bacterial class *Planctomycetes* was also the most highly connected to other bacterial and fungal classes in severe drought conditions, suggesting possible importance in the blue grama drought response. We also observed negative response ratios in several typically drought-tolerant bacterial clades including the *Firmicutes, Gemmatimonadetes*, and actinobacterial *Streptomyces* that challenge the perception of their conserved response to drought across plant rhizobiomes and soil ecosystems. Although we detected weak evidence for correlations between root exudate composition and community beta diversity, these analyses provide a platform for future exploration of similar multi-omic datasets. Altogether this study provides an important step towards understanding plant-microbe feedbacks under drought conditions and provides a foundation for future targeted investigations of the potential protective effects of root exudates and the rhizobiome against drought stress.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA1006788.

Author contributions

HG: Data curation, Formal analysis, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing, Investigation. DU: Supervision, Validation, Visualization, Writing – review & editing, Project administration, Resources, Conceptualization, Funding acquisition, Investigation, Methodology. BP: Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing. LG: Conceptualization, Funding acquisition, Resources, Writing – review & editing, Methodology. RM: Methodology, Writing – review & editing, Investigation, Supervision, Validation, Visualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the U.S. Department of Energy Office of Science, Biological and Environmental Research Division, under award number F255LANL2018, and by the U.S. Department of Agriculture National Institute of Food and Agriculture under award number 2020-07194.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1310790/full#supplementary-material

References

Abbasi, S., Sadeghi, A., and Safaie, N. (2020). Streptomyces alleviate drought stress in tomato plants and modulate the expression of transcription factors ERF1 and WRKY70 genes. *Scientia Hortic.* 265, 109206. doi: 10.1016/j.scienta.2020.109206

Amani Machiani, M., Javanmard, A., Morshedloo, M. R., Aghaee, A., and Maggi, F. (2021). Funneliformis mosseae inoculation under water deficit stress improves the yield and phytochemical characteristics of thyme in intercropping with soybean. *Sci. Rep.* 11 (1). doi: 10.1038/s41598-021-94681-9

Anderson, M. J. (2017). "Permutational Multivariate Analysis of Variance (PERMANOVA)," in *Wiley StatsRef: Statistics Reference Online* (American Cancer Society), 1–15. doi: 10.1002/9781118445112.stat07841

Avendaño-González, M., and Siqueiros-Delgado, M. E. (2021). Past, present and future distribution of Bouteloua gracilis, a key species of North American grasslands, changes related to climate change. *J. Arid Environments* 186, 104417. doi: 10.1016/j.jaridenv.2020.104417

Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S., and Vivanco, J. M. (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu. Rev. Plant Biol.* 57, 233–266. doi: 10.1146/annurev.arplant.57.032905.105159

Bandopadhyay, S., Li, X., Bowsher, A. W., Last, R. L., and Shade, A. (2023). Disentangling plant- and environment-mediated drivers of active rhizosphere bacterial community dynamics during short-term drought. *BioRxiv*. doi: 10.1101/2023.06.06.543716

Barberán, A., Bates, S. T., Casamayor, E. O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6 (2), 343–351. doi: 10.1038/ismej.2011.119

Bécard, G., and Piché, Y. (1989). Fungal growth stimulation by CO2 and root exudates in vesicular-arbuscular mycorrhizal symbiosis. *Appl. Environ. Microbiol.* 55 (9), 2320–2325. doi: 10.1128/aem.55.9.2320-2325.1989

Berdugo, M., Delgado-Baquerizo, M., Soliveres, S., Hernández-Clemente, R., Zhao, Y., Gaitán, J. J., et al. (2020). Global ecosystem thresholds driven by aridity. *Science* 367 (6479), 787–790. doi: 10.1126/science.aay5958

Bizos, G., Papatheodorou, E. M., Chatzistathis, T., Ntalli, N., Aschonitis, V. G., and Monokrousos, N. (2020). The role of microbial inoculants on plant protection, growth stimulation, and crop productivity of the olive tree (Olea europea L.). *Plants* 9 (6), 743. doi: 10.3390/plants9060743

Budak, H., Kantar, M., and Yucebilgili Kurtoglu, K. (2013). Drought tolerance in modern and wild wheat. *Sci. World J.* 2013, e548246. doi: 10.1155/2013/548246

Bukovská, P., Gryndler, M., Gryndlerová, H., Püschel, D., and Jansa, J. (2016). Organic nitrogen-driven stimulation of arbuscular mycorrhizal fungal hyphae correlates with abundance of ammonia oxidizers. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.00711

Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S., and Fierer, N. (2016). Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nat. Microbiol.* 2 (3), 16242. doi: 10.1038/nmicrobiol.2016.242

Castellano-Hinojosa, A., Strauss, S. L., González-López, J., and Bedmar, E. J. (2021). Changes in the diversity and predicted functional composition of the bulk and rhizosphere soil bacterial microbiomes of tomato and common bean after inorganic N-fertilization. *Rhizosphere* 18, 100362. doi: 10.1016/j.rhisph.2021.100362

Chaparro, J. M., Badri, D. V., and Vivanco, J. M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* 8 (4), 790–803. doi: 10.1038/ismej.2013.196

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2022). GTDB-Tk v2: Memory friendly classification with the genome taxonomy database. *Bioinformatics* 38 (23), 5315–5316. doi: 10.1093/bioinformatics/btac672

Coffin, D. P., and Lauenroth, W. K. (1992). Spatial variability in seed production of the perennial bunchgrass Bouteloua gracilis (Gramineae). *Am. J. Bot.* 79 (3), 347–353. doi: 10.2307/2445025

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42 (Database issue), D633–D642. doi: 10.1093/nar/gkt1244

Costello, D. F. (1944). Important species of the major forage types in colorado and wyoming. *Ecol. Monogr.* 14 (1), 107–134. doi: 10.2307/1961633

Dahal, B., NandaKafle, G., Perkins, L., and Brözel, V. S. (2017). Diversity of free-Living nitrogen fixing Streptomyces in soils of the badlands of South Dakota. *Microbiological Res.* 195, 31–39. doi: 10.1016/j.micres.2016.11.004

Dai, L., Zhang, G., Yu, Z., Ding, H., Xu, Y., and Zhang, Z. (2019). Effect of drought stress and developmental stages on microbial community structure and diversity in peanut rhizosphere soil. *Int. J. Mol. Sci.* 20 (9), 2265. doi: 10.3390/ijms20092265

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., et al. (2018). A global atlas of the dominant bacteria found in soil. *Sci. (New York N.Y.)* 359 (6373), 320–325. doi: 10.1126/science.aap9516

de Vries, F. T., Griffiths, R. I., Knight, C. G., Nicolitch, O., and Williams, A. (2016). Harnessing rhizosphere microbiomes for drought-resilient crop production. *Science* 368 (6488), 270–274. doi: 10.1126/science.aaz5192

Diniz-Filho, J. A. F., Santos, T., Rangel, T. F., and Bini, L. M. (2012). A comparison of metrics for estimating phylogenetic signal under alternative evolutionary models. *Genet. Mol. Biol.* 35 (3), 673–679. doi: 10.1590/S1415-47572012005000053

Edgar, R. C. (2016). SINTAX: A simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *BioRxiv*, 074161. doi: 10.1101/074161

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26 (19), 2460–2461. doi: 10.1093/bioinformatics/btq461

Evans, S. E., Bell-Dereske, L. P., Dougherty, K. M., and Kittredge, H. A. (2020). Dispersal alters soil microbial community response to drought. *Environ. Microbiol.* 22 (3), 905–916. doi: 10.1111/1462-2920.14707

Goemann, H. M., Ulrich, D. E.M., Peyton, B. M., Gallegos-Graves, L. V., and Mueller, R. C. (2024). Severe and mild drought cause distinct phylogenetically linked shifts in the blue grama (*Bouteloua gracilis*) rhizobiome. *Bouteloua gracilis*. Available at: https://www.fs.fed.us/database/feis/plants/graminoid/bougra/all.html.

Graça, A. P., Calisto, R., and Lage, O. M. (2016). Planctomycetes as novel source of bioactive molecules. *Front. Microbiol.* 7. doi: 10.3389/fmicb.2016.01241

Guo, B., Zhang, H., Liu, Y., Chen, J., and Li, J. (2023). Drought-resistant trait of different crop genotypes determines assembly patterns of soil and phyllosphere microbial communities. *Microbiol. Spectr.* 11 (5), e00068–e00023. doi: 10.1128/spectrum.00068-23

Hackathon, R., Bolker, B., Butler, M., Cowan, P., de Vienne, D., Eddelbuettel, D., et al. (2020). *Phylobase*. Available at: http://github.com/fmichonneau/phylobase.

Huang, X.-F., Zhou, D., Lapsansky, E. R., Reardon, K. F., Guo, J., Andales, M. J., et al. (2017). Mitsuaria sp. And Burkholderia sp. From Arabidopsis rhizosphere enhance drought tolerance in Arabidopsis thaliana and maize (Zea mays L.). *Plant Soil* 419 (1), 523–539. doi: 10.1007/s11104-017-3360-4

Huber, K. J., Vieira, S., Sikorski, J., Wüst, P. K., Fösel, B. U., Gröngröft, A., et al. (2022). Differential response of acidobacteria to water content, soil type, and land use during an extended drought in african savannah soils. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.750456

Jiang, Y., Li, S., Li, R., Zhang, J., Liu, Y., Lv, L., et al. (2017). Plant cultivars imprint the rhizosphere bacterial community composition and association networks. *Soil Biol. Biochem.* 109, 145–155. doi: 10.1016/j.soilbio.2017.02.010

Jombart, T., Balloux, F., and Dray, S. (2010). *adephylo*: New tools for investigating the phylogenetic signal in biological traits. *Bioinformatics* 26 (15), 1907–1909. doi: 10.1093/bioinformatics/btq292

Karlowsky, S., Augusti, A., Ingrisch, J., Akanda, M. K. U., Bahn, M., and Gleixner, G. (2018). Drought-induced accumulation of root exudates supports post-drought

recovery of microbes in mountain grassland. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01593

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30 (4), 772–780. doi: 10.1093/molbev/mst010

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinf. (Oxford England)* 26 (11), 1463–1464. doi: 10.1093/bioinformatics/btq166

Koyama, A., Steinweg, J. M., Haddix, M. L., Dukes, J. S., and Wallenstein, M. D. (2018). Soil bacterial community responses to altered precipitation and temperature regimes in an old field grassland are mediated by plants. *FEMS Microbiol. Ecol.* 94 (1), fix156. doi: 10.1093/femsec/fix156

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

Lasudee, K., Tokuyama, S., Lumyong, S., and Pathom-aree, W. (2018). Actinobacteria Associated With Arbuscular Mycorrhizal Funneliformis mosseae Spores, Taxonomic Characterization and Their Beneficial Traits to Plants: Evidence Obtained From Mung Bean (Vigna radiata) and Thai Jasmine Rice (Oryza sativa). *Front. Microbiol.* 9. doi: 10.3389/fmicb.2018.01247

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* 47 (W1), W256–W259. doi: 10.1093/nar/gkz239

Li, H., Guo, Q., Jing, Y., Liu, Z., Zheng, Z., Sun, Y., et al. (2020). Application of streptomyces pactum act12 enhances drought resistance in wheat. *J. Plant Growth Regul.* 39 (1), 122–132. doi: 10.1007/s00344-019-09968-z

Ling, N., Wang, T., and Kuzyakov, Y. (2022). Rhizosphere bacteriome structure and functions. *Nat. Commun.* 13 (1). doi: 10.1038/s41467-022-28448-9

Lozupone, C., and Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 71 (12), 8228–8235. doi: 10.1128/AEM.71.12.8228-8235.2005

Lutzoni, F., Kauff, F., Cox, C. J., McLaughlin, D., Celio, G., Dentinger, B., et al. (2004). Assembling the fungal tree of life: Progress, classification, and evolution of subcellular traits. *Am. J. Bot.* 91 (10), 1446–1480. doi: 10.3732/ajb.91.10.1446

MatChado, M. S., Lauber, M., Reitmeier, S., Kacprowski, T., Baumbach, J., Haller, D., et al. (2021). Network analysis methods for studying microbial communities: A mini review. *Comput. Struct. Biotechnol. J.* 19, 2687–2698. doi: 10.1016/j.csbj.2021.05.001

Mathur, P., and Roy, S. (2021). Insights into the plant responses to drought and decoding the potential of root associated microbiome for inducing drought tolerance. *Physiologia Plantarum* 172 (2), 1016–1029. doi: 10.1111/ppl.13338

Mathur, S., Tomar, R. S., and Jajoo, A. (2019). Arbuscular Mycorrhizal fungi (AMF) protects photosynthetic apparatus of wheat under drought stress. *Photosynthesis Res.* 139 (1), 227–238. doi: 10.1007/s11120-018-0538-4

Matsen, F. A., Kodner, R. B., and Armbrust, E. V. (2010). pplacer: Linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinf.* 11 (1), 538. doi: 10.1186/1471-2105-11-538

Maurer, D., Malique, F., Alfarraj, S., Albasher, G., Horn, M. A., Butterbach-Bahl, K., et al. (2021). Interactive regulation of root exudation and rhizosphere denitrification by plant metabolite content and soil properties. *Plant Soil* 467 (1), 107–127. doi: 10.1007/s11104-021-05069-7

McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8 (4), e61217. doi: 10.1371/journal.pone.0061217

Mueller, R., Gallegos-Graves, L., and Kuske, C. (2016). A new fungal large subunit ribosomal RNA primer for high-throughput sequencing surveys. *FEMS Microbiol. Ecol.* 92, 1–11. doi: 10.1093/femsec/fiv153

Mukherjee, S., and Mishra, A. K. (2021). Increase in compound drought and heatwaves in a warming world. *Geophysical Res. Lett.* 48 (1), e2020GL090617. doi: 10.1029/2020GL090617

Naveed, M., Hussain, M. B., Zahir, Z. A., Mitter, B., and Sessitsch, A. (2014). Drought stress amelioration in wheat through inoculation with Burkholderia phytofirmans strain PsJN. *Plant Growth Regul.* 73 (2), 121–131. doi: 10.1007/s10725-013-9874-8

Naylor, D., and Coleman-Derr, D. (2018). Drought stress and root-associated bacterial communities. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.02223

Naylor, D., DeGraaf, S., Purdom, E., and Coleman-Derr, D. (2017). Drought and host selection influence bacterial community dynamics in the grass root microbiome. *ISME J.* 11 (12), 2691–2704. doi: 10.1038/ismej.2017.118

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *Community ecology package (version 2.5-6)* (CRAN). Available at: https://cran.r-project.orghttps://github.com/vegandevs/vegan.

Parada, A. E., Needham, D. M., and Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* 18 (5), 1403–1414. doi: 10.1111/1462-2920.13023

Passari, A. K., Chandra, P., Zothanpuia,, Mishra, V. K., Leo, V. V., Gupta, V. K., et al. (2016). Detection of biosynthetic gene and phytohormone production by endophytic actinobacteria associated with Solanum lycopersicum and their plant-growth-promoting effect. *Res. Microbiol.* 167 (8), 692–705. doi: 10.1016/j.resmic.2016.07.001

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., et al. (2014). geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30 (15), 2216–2218. doi: 10.1093/bioinformatics/btu181

Peschel, S., Müller, C. L., Von Mutius, E., Boulesteix, A.-L., and Depner, M. (2021). NetCoMi: Network construction and comparison for microbiome data in R. *Briefings Bioinf.* 22 (4), bbaa290. doi: 10.1093/bib/bbaa290

Phillips, R. P., Erlitz, Y., Bier, R., and Bernhardt, E. S. (2008). New approach for capturing soluble root exudates in forest soils. *Funct. Ecol.* 22 (6), 990–999. doi: 10.1111/j.1365-2435.2008.01495.x

Qi, J., Chen, B., Gao, J., Peng, Z., Jiao, S., and Wei, G. (2022). Responses of soil bacterial community structure and function to dry–wet cycles more stable in paddy than in dryland agricultural ecosystems. *Global Ecol. Biogeography* 31 (2), 362–377. doi: 10.1111/geb.13433

R, C. T. (2023). *R: A language and environment for statistical computing (4.3.0)* (R Foundation for Statistical Computing). Available at: https://www.R-project.org/.

Ravenek, J. M., Mommer, L., Visser, E. J. W., van Ruijven, J., van der Paauw, J. W., Smit-Tiekstra, A., et al. (2016). Linking root traits and competitive success in grassland species. *Plant Soil* 407 (1), 39–53. doi: 10.1007/s11104-016-2843-z

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3 (2), 217–223. doi: 10.1111/j.2041-210X.2011.00169.x

Reyes, S. M. R., Hoyos, G. R., Júnior, D., da, C. F., Filho, A. B. C., and Fonseca, L. P. M. (2019). Physiological response of Physalis Peruviana L. seedlings inoculated with Funneliformis mosseae under drought stress. *Rev. Ciências Agrárias* 42 (1), 836. doi: 10.19084/RCA18140

RStudio Team (2020). *RStudio: Integrated Development for R* (RStudio, PBC). Available at: http://www.rstudio.com/.

Ruth, B., Khalvati, M., and Schmidhalter, U. (2011). Quantification of mycorrhizal water uptake via high-resolution on-line water content sensors. *Plant Soil* 342 (1), 459–468. doi: 10.1007/s11104-010-0709-3

Santos-Medellín, C., Liechty, Z., Edwards, J., Nguyen, B., Huang, B., Weimer, B. C., et al. (2021). Prolonged drought imparts lasting compositional changes to the rice root microbiome. *Nat. Plants* 7 (8), 1065–1077. doi: 10.1038/s41477-021-00967-1

Schrey, S. D., Erkenbrack, E., Fruh, E., Fengler, S., Hommel, K., Horlacher, N., et al. (2012). Production of fungal and bacterial growth modulating secondary metabolites is widespread among mycorrhiza-associated streptomycetes. *BMC Microbiol.* 12, 164. doi: 10.1186/1471-2180-12-164

Shami, A., Jalal, R. S., Ashy, R. A., Abuauf, H. W., Baz, L., Refai, M. Y., et al. (2022). Use of metagenomic whole genome shotgun sequencing data in taxonomic assignment of dipterygium glaucum rhizosphere and surrounding bulk soil microbiomes, and their response to watering. *Sustainability* 14 (14). doi: 10.3390/su14148764

Sheik, C. S., Beasley, W. H., Elshahed, M. S., Zhou, X., Luo, Y., and Krumholz, L. R. (2011). Effect of warming and drought on grassland microbial communities. *ISME J.* 5 (1010), 1692–1700. doi: 10.1038/ismej.2011.32

Simmons, T., Styer, A. B., Pierroz, G., Gonçalves, A. P., Pasricha, R., Hazra, A. B., et al. (2020). Drought drives spatial variation in the millet root microbiome. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00599

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30 (9), 1312–1313. doi: 10.1093/bioinformatics/btu033

Tallapragada, P., Dikshit, R., and Seshagiri, S. (2016). Influence of Rhizophagus spp. And Burkholderia seminalis on the Growth of Tomato (Lycopersicon esculatum) and Bell Pepper (Capsicum annuum) under Drought Stress. *Commun. Soil Sci. Plant Anal.* 47 (17), 1975–1984. doi: 10.1080/00103624.2016.1216561

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., et al. (2017). A communal catalogue reveals earth's multiscale microbial diversity. *Nature* 551 (7681), 457–463. doi: 10.1038/nature24621

Tóth, Z., Táncsics, A., Kriszt, B., Kröel-Dulay, G., Ónodi, G., and Hornung, E. (2017). Extreme effects of drought on composition of the soil bacterial community and decomposition of plant tissue. *Eur. J. Soil Sci.* 68 (4), 504–513. doi: 10.1111/ejss.12429

Ulrich, D. E. M., Clendinen, C. S., Alongi, F., Mueller, R. C., Chu, R. K., Toyoda, J., et al. (2022). Root exudate composition reflects drought severity gradient in blue grama (Bouteloua gracilis). *Sci. Rep.* 12 (1). doi: 10.1038/s41598-022-16408-8

Ulrich, D. E. M., Sevanto, S., Ryan, M., Albright, M. B. N., Johansen, R. B., and Dunbar, J. M. (2019). Plant-microbe interactions before drought influence plant physiological responses to subsequent severe drought. *Sci. Rep.* 9 (1), 249. doi: 10.1038/s41598-018-36971-3

Umashankar, N., Venkateshamurthy, P., Krishnamurthy, R., Raveendra, H. R., and Satish, K. M. (2012). Effect of microbial inoculants on the growth of silver oak (Grevillea robusta) in nursery condition. *Int. J. Environ. Sci. Dev.* 3 (1), 72–76.

Vries, F. T. d, Williams, A., Stringer, F., Willcocks, R., McEwing, R., Langridge, H., et al. (2019). Changes in root-exudate-induced respiration reveal a novel mechanism through which drought affects ecosystem carbon cycling. *New Phytol.* 224 (1), 132–145. doi: 10.1111/nph.16001

Wang, B., and Allison, S. D. (2021). Drought legacies mediated by trait trade-offs in soil microbiomes. *Ecosphere* 12 (6), e03562. doi: 10.1002/ecs2.3562

Wang, X., Sharp, C., Jones, G., Grasby, S., Brady, A., and Dunfield, P. (2015). Stable-isotope probing identifies uncultured planctomycetes as primary degraders of a complex heteropolysaccharide in soil. *ASM Microbial Ecol.* 81 (14), 4607–4615. doi: 10.1128/AEM.00055-15

Wang, Y., Xie, Y., Ma, H., Zhang, Y., Zhang, J., Zhang, H., et al. (2022). Responses of soil microbial communities and networks to precipitation change in a typical steppe ecosystem of the loess plateau. *Microorganisms* 10 (4), 817. doi: 10.3390/microorganisms10040817

Warrad, M., Hassan, Y. M., Mohamed, M. S. M., Hagagy, N., Al-Maghrabi, O. A., Selim, S., et al. (2020). *A bioactive fraction from streptomyces* sp. *Enhances Maize Tolerance against Drought Stress.* 30 (8), 1156–1168. doi: 10.4014/jmb.2003.03034

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Software* 4 (43), 1686. doi: 10.21105/joss.01686

Wiegand, S., Jogler, M., and Jogler, C. (2018). On the maverick planctomycetes. *FEMS Microbiol. Rev.* 42 (6), 739–760. doi: 10.1093/femsre/fuy029

Williams, A., and de Vries, F. T. (2020). Plant root exudation under drought: Implications for ecosystem functioning. *New Phytol.* 225 (5), 1899–1905. doi: 10.1111/nph.16223

Wong, S.-K., Cui, Y., Chun, S.-J., Kaneko, R., Masumoto, S., Kitagawa, R., et al. (2023). Vegetation as a key driver of the distribution of microbial generalists that in turn shapes the overall microbial community structure in the low Arctic tundra. *Environ. Microbiome* 18 (1), 41. doi: 10.1186/s40793-023-00498-6

Wu, Q.-S., Srivastava, A. K., and Zou, Y.-N. (2013). AMF-induced tolerance to drought stress in citrus: A review. *Scientia Hortic.* 164, 77–87. doi: 10.1016/j.scienta.2013.09.010

Xu, L., Naylor, D., Dong, Z., Simmons, T., Pierroz, G., Hixson, K. K., et al. (2018). Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. *Proc. Natl. Acad. Sci.* 115 (18), E4284–E4293. doi: 10.1073/pnas.1717308115

Zheng, Y., Saitou, A., Wang, C.-M., Toyoda, A., Minakuchi, Y., Sekiguchi, Y., et al. (2019). Genome features and secondary metabolites biosynthetic potential of the class ktedonobacteria. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.00893

Zhu, X., Liu, M., Kou, Y., Liu, D., Liu, Q., Zhang, Z., et al. (2020). Differential effects of N addition on the stoichiometry of microbes and extracellular enzymes in the rhizosphere and bulk soils of an alpine shrubland. *Plant Soil* 449 (1), 285–301. doi: 10.1007/s11104-020-04468-6

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 08 February 2024
DOI 10.3389/frmbi.2024.1352566

OPEN ACCESS

EDITED BY
Catherine Gehring,
Northern Arizona University, United States

REVIEWED BY
Alain Isabwe,
University of Michigan, United States
Pengfan Zhang,
Innovative Genomics Institute (IGI),
United States

*CORRESPONDENCE
Julia Sacharow
Julia.Sacharow@en.uni-giessen.de

RECEIVED 08 December 2023
ACCEPTED 22 January 2024
PUBLISHED 08 February 2024

CITATION
Sacharow J, Ratering S, Quiroga S, Geißler-Plaum R, Schneider B, Österreicher Cunha-Dupont A and Schnell S (2024) Cercozoan diversity of spring barley grown in the field is strongly plant compartment specific. *Front. Microbiomes* 3:1352566. doi: 10.3389/frmbi.2024.1352566

Cercozoan diversity of spring barley grown in the field is strongly plant compartment specific

Julia Sacharow*, Stefan Ratering, Santiago Quiroga, Rita Geißler-Plaum, Bellinda Schneider, Alessandra Österreicher Cunha-Dupont and Sylvia Schnell

Professorship of General and Soil Microbiology, Institute of Applied Microbiology, Research Centre for Biosystems, Land Use and Nutrition, Justus-Liebig-University Giessen, Giessen, Germany

Protists are an important part of the plant holobiome and influence plant growth and pathogenic pressure as consumers. *Hordeum vulgare* is one of the most economically important crops worldwide, and its yield depends on optimal environmental plant-growth conditions and pathogen defense. This study aimed to analyse the natural compositions of the cercozoan diversity, one of the most important and dominant protist phyla, of spring barley at different developmental stages, from different plant compartments over two years. *Hordeum vulgare* bulk soil samples were taken before seeding and after harvest on an organic farming field. Bulk soil, rhizosphere soil, roots and leaves were sampled at the flowering and ripening stages, and analysed with cercozoan-specific primers. Results showed a clear dominance of the families Sandonidae, Allapsidae, Cercomonadidae, Rhogostomidae and the order Glissomonadida in all sample types. Separated analyses of root, leaf and soil samples showed that members of the family Sandonidae were strongly enriched in leaf samples, while members of the Allapsidae family were enriched in the roots. No compositional differences were detected between the different plant developmental stages, except for the beta diversity of the leaf samples at the flowering and ripening stages. It can be concluded that the cercozoan diversity of spring barley is primarily affected by the plant compartment and not by the plant developmental stage. Further studies are needed to analyze the cercozoan community in greater taxonomic depth and to target their ecological function.

KEYWORDS

***Hordeum vulgare*, soil cercozoa, leaf cercozoa, root cercozoa, holobiome**

Introduction

Protists are an essential component of the biodiversity and ecosystem functioning of soils. They have diverse feeding behaviours consisting of bacterivores, fungivores, omnivores, mixotrophs and phototrophs protists but with specific prey spectra (Geisen, 2016; Dumack et al., 2019; Asiloglu et al., 2020). Their influence on the soil microbial community via consumption also affects the performance of the surrounding plants through plant growth promotion and plant health improvement (Jentschke et al., 1995; Krome et al., 2010; Weidner et al., 2017; Bahroun et al., 2021). Phagotrophic protists, for example, were shown to act as top-down controller of plant pathogens, and the ciliate *Colpoda cucullus* increased the dry matter content of maize (Wu et al., 2022; Zhang et al., 2022). Despite their diversity, ecological importance as predators of the soil microbiome, and bio-indicators of soil quality (Zhao et al., 2019), they are under-researched in comparison to bacteria and fungi, although they are also good candidates for use in biological crop protection (Sacharow et al., 2023).

To understand the whole plant-microbe-soil system it is important to analyse protist patterns on plants: they were shown to be strongly shaped by plant biomass, soil pH-value and moisture (Öztoprak et al., 2020), whereas land use is controversial. While Glaser et al. (2015) showed only a small effect of land use on protist communities, Geisen et al. (2015) showed that the protist community was strongly influenced by land use: where forest and grassland soils were dominated by Rhizaria and Amoebozoa, peat soils were dominated by Alveolata. Santos et al. (2020) showed that protist trophic groups were also affected by the land use intensity. Analysis of the protist community composition of switch grass revealed a lower diversity in the rhizosphere soil than in the bulk soil. The protist composition of the rhizosphere soil was mainly controlled by dispersal constraints and plant selection (Guo et al., 2018; Ceja-Navarro et al., 2021). Similarly, richness decreased with increasing soil depth, as well as with the use of chemical fungicides (Guo et al., 2018; Degrune et al., 2019). Glyphosate treatment of barley leaves altered protist communities on roots and their interactions with the surrounding prey (Imparato et al., 2016). A comparison between rhizosphere and bulk soil after the application of chemical fungicides revealed an increased amount of Alveolata and Amoebozoa (Guo et al., 2018). Protist community composition is also strongly influenced by fertilisation. In comparison to other microorganisms in the soil, nitrogen fertilisation had a higher impact on protists than on bacteria and fungi (Zhao et al., 2019). Krashevska et al. (2014) also showed that amoeba communities are primarily structured by abiotic factors and antagonistic interactions, rather than by prey availability.

The phylum Cercozoa (Cavalier-Smith and Chao, 2003) is a large group of free-living protists, which are an important part of the soil ecosystem (Bass and Cavalier-Smith, 2004). A cloning-based analysis of Brassicaceae leaves showed a highly diverse leaf-associated Cercozoa community composed of bacterivores, plant pathogens and endophytes (Ploch et al., 2016). Flues et al. (2018) discovered that the phyllosphere and rhizosphere of various plants are dominated by the genera *Cercomonas*, *Neocercomonas* and *Paracercomonas*. They found differences in diversity between the phyllosphere and the rhizosphere, but no differences between plant species. In contrast, Walden et al. (2021) demonstrated that the Cercozoa communities of some plants were specific to plant species, and their diversity was also seasonally influenced, changing from spring to autumn. Furthermore, they were sensitive to invasions of the ecosystem engineer earthworm, changing to an earthworm-associated community of Cercozoa (Dumack et al., 2022). Different environments are dominated by different groups of Cercozoa. For example, grasslands were dominated by Sarcomonadea (69%) and dunes by Thecofilosea (43%). However, the families Sandonidae, Allapsidae, and Rhogostomidae prevailed in both environments (Roshan et al., 2021). An analysis of agricultural and forest soils used for wheat plants showed that the cercozoan rhizosphere community was influenced by soil type, as well as the genotype of the wheat plant, and the soils were dominated by Sarcomonadea (42.3%), followed by Thecofilosea (27.1%) and Imbricatea (19.2%). The agricultural soil was dominated by the families Limnofilidae, Protaspididae, Thaumatomonadidae and unclassified Cryomonadida, while the forest soil was dominated by Rhogostomidae, Mesofilidae, unclassified Cercozoa, unclassified Imbricatea and unclassified Tectofilosida (Rossmann et al., 2020).

Hordeum vulgare, first cultivated around 7800-7500 B.C., is currently one of the most economically important crops besides wheat, rice and maize and is used in the production of alcoholic beverages and animal feed (Nesbitt and Samuel, 1996; Chełkowski et al., 2003). The aim of this project was to analyse the Cercozoan communities' composition of this important crop from different sampling materials, at different growing stages, for two seasons. The results of the analysis will give an overview of the Cercozoan patterns of the plant and can help to understand the whole plant-microbe-soil system of barley.

Materials and methods

Seed preparation

Hordeum vulgare ODILIA (Öko Korn Nord, Germany) seeds were covered with a gum arabic and talc mixture before sowing. Gum arabic 25% (Roth, Germany) and talc powder (Roth, Germany) were autoclaved individually. The gum arabic was adjusted (Mettler-Toledo, Germany) to pH 7, mixed in a 1:1 ratio with $MgSO_4$ (Roth, Germany) and shaken at 125 RPM on an orbital shaker (PSU-20i, Bio San, Latvia) at 20°C for 20 min. The mixture was then slowly added to the seeds until they were covered completely. Finally, talc powder was added and the seeds mixed (Kloepper, 1981). This study was part of a larger experimental project (https://www.bonares.de/bread-and-beer) with different seed inoculations in which the mixture of gum arabic and talc was used for the control plants, while plant growth bacteria were added to other treatments not studied here.

Sampling fields on the Gladbacherhof

Fields on the Hessian State Domain Gladbacherhof (50° 23' N, 8° 15' E) of the Justus-Liebig-University Giessen were used for the

cultivation of the plants. Seeds were sown in 1.5 x 5 m plots at four randomly selected points in the field using a tractor and seed drill. The plots for this experiment were surrounded by plots from the other project experiments with different seed inoculations. *Hordeum vulgare* ODILIA was sown in April 2021 and March 2022 and weeds were removed by hand when the plants were small.

Sampling of plant material

Soil and plants (when applicable) were sampled before seeding, at the flowering and ripening stages, and after harvesting. Before seeding, 30 cm of the top soil were sampled with a Dutch auger. The samples were scratched into new plastic bags, transported to the laboratory on ice, sieved through a 2 mm size sieve, and stored in tubes (Sarstedt, Germany) at -20°C. At flowering, randomly selected plants were dug up, placed in new plastic bags and transported on ice to the laboratory. The bulk soil and the rhizosphere soil were separated by shaking. Bulk and rhizosphere soil were sieved to a size of 2 mm and stored in tubes at -20°C. Furthermore, soil-free parts of the roots and leaves were cut off with sterile scissors. The plant material was stored in tubes at -20°C. At the ripening, samples were taken in the same way as at flowering and after harvesting, the soil was sampled the same way as before the seeding.

Analysis of chemical soil parameters

The bulk soil samples from before the seeding were analysed for NH_4^+, NO_3^- and C:N. Ammonia content was determined by the method of Kandeler and Gerber (1988) after extraction with 1 M KCl. The nitrate was extracted from the soil according to the method in Cardinale et al. (2020) and nitrate content was determined with ion chromatography (Bak et al., 1991). For the C:N ratio measurement the soil was dried at 105°C for 24 h in a drying cabinet (ULE 500, Memmert, Germany) and milled in a Retsch mill (MM400, Retsch GmbH, Germany). In detail, the samples were placed in 2 ml tubes (Sarstedt, Germany), small iron balls were added and the tubes were shaken for 2 min at 30 s^{-1}. Milled soil (35 mg) was filled in small tin boats (4 x 4 x 11 mm, Elementar Analysensysteme, Germany), filled boats were pressed together and measured with an elemental analyser (Unicube, Elementar, Germany).

DNA extraction

The roots were washed by adding sterilized pure water in 2 ml tubes (Sarstedt, Germany) and shaken for 30 s. The washed roots were transferred into new 2 ml tubes and washed again until no further soil remains were observed in the water. Leaf and washed root material were cut into pieces with sterile scissors and crushed in liquid nitrogen using a sterile pestle and mortar. For the DNA extraction 100 - 150 mg leaf material, 150 – 250 mg root material and 300 – 350 mg soil material were used according to the protocol of Abdullaeva et al. (2021). After the extraction, the DNA was stored at -20°C.

Amplicon library preparation and ion torrent sequencing

The V4 variable region of the 18S rRNA gene from the leaf material, root material, rhizosphere and bulk soil was amplified in a semi-nested PCR to analyse the cercozoan diversity. In the first step, the primer pair S616F_Cerco (5'-TTAAAAAGCTCGTAGTTG-3') and S616F_Eocer (5'-TTAAAAAGCGCGTAGTTG-3') were used as forward primers and S963R_Cerco (5'-CAACTTTCGTTCT TGATTAAA-3') as reverse primer. In the second step of the semi-nested PCR the same forward primers were used as before with the S947R_Cerco (5'-AAGAAGACATCCTTGGTG-3') reverse primer and barcodes (Fiore-Donno et al., 2017). The PCR reaction for the amplification (MycyclerTM, Bio-Rad, USA) was prepared as described by Fiore-Donno et al. (2017) with 2 µl of DNA template for both steps of the semi-nested PCR and an optimised thermal program (Table 1). Further steps of the Ion Torrent metabarcoding were done as described by Abdullaeva et al. (2021) with a final pool concentration of 300 pM.

Data analysis

The raw Ion Torrent sequences were processed with the bioinformatic pipeline QIIME2 version 2021.2 (Bolyen et al., 2019). The QIIME2 cutadapt plugin (Martin, 2011) was used for demultiplexing of the gene sequences and QIIME2 plugin DADA2 (Callahan et al., 2017) was used for quality control, filtering, chimera identification, denoising, summarizing the amplicon sequence variation (ASV) table, which records the number of observations of each exact ASV in each sample. The taxonomy was assigned by the q2-feature-classifier (Bokulich et al., 2018) trained on the PR2 database 4.14.0 (Guillou et al., 2013). The ASVs that were identified as non-cercozoan were removed.

Statistical analysis

Data analysis of the ASV table was performed with QIIME2 version 2021.2, R-Studio version 4.2.3 (R Core Team, 2020), the phyloseq 1.28.0 package (McMurdie & Holmes, 2014) and the qiime2r 0.99.6 package (Bisanz, 2018). Alpha-diversity (R-Studio version 4.2.3) was determined by the mean value from the ASV table with observed richness, Shannon diversity index (Shannon and Weaver, 1964) and Fisher diversity index (Fisher et al., 1943) after rarefaction. The significant differences were determined with the Wilcox test (Wilcoxon, 1945) and the Holm correction method (Holm, 1979) through 999 permutations. For the plots ggplot2 3.4.2 (Wickham, 2016) was used. Beta-diversity (QIIME2 version 2021.2 and R-Studio version 4.2.3) was obtained with Aitchison principal component analysis (PCA) as well as Robust Aitchison principal component analysis (RPCA) with a centred log-ratio transform (CLR) and a robust centred log-ratio transform (RCLR) (DEICODE, Martino et al., 2019). The significant differences were determined by a Permutational Multivariate Analysis of Variance (PERMANOVA) using the Adonis method (Anderson, 2001) with

TABLE 1 Optimised touchdown PCR program according to Fiore-Donno et al. (2017) for the amplification of the V4 variable region of the 18S rRNA gene with the cercozoan specific primer pairs S616F_Cerco, S616F_Eocer and S963R_Cerco as well as S616F_Cerco, S616F_Eocer and S947R_Cerco.

Step – First PCR	Temperature – First PCR	Time – First PCR	
1.	95°C	3 min	Steps 2-4 repeated for 24 times.
2.	98°C	20 s	
3.	50°C	20 s	
4.	72°C	30 s	
5.	72°C	5 min	
Step – Second PCR	**Temperature – Second PCR**	**Time – Second PCR**	
1.	95°C	3 min	Steps 2–4 repeated for 5 times with a temperature decrement of 1°C per cycle in step 3. Steps 5-6 repeated for 19 times.
2.	98°C	20 s	
3.	70°C	30 s	
4.	72°C	30 s	
5.	98°C	20 s	
6.	65°C	30 s	
7.	72°C	30 s	
8.	72°C	5 min	

a Benjamini–Hochberg (Benjamini and Hochberg, 1995) correction and employing 999 permutations. The cercozoan ASVs from the underground samples (bulk soil, rhizosphere soil and roots) were pre-processed (R-Studio version 4.2.3) as to keep all ASVs present in a minimum of two samples and a minimum of five reads per ASV. The cercozoan ASVs from the leaf samples were not pre-processed. For the compositional data analysis R-Studio version 4.2.3 and the package ALDEx2 1.22.0 (Fernandes et al., 2013) were used. An ALDEx2 test was done by performing a centred log ratio (clr) transformation (Aitchison, 1982; Aitchison, 1986) using as denominator the geometric mean abundance of all features and 128 Monte-Carlo instances. Then a Welch's t-test with a Benjamini–Hochberg (Benjamini & Hochberg, 1995) correction was carried out. The phylogenetic tree was calculated with QIIME2 version 2021.2 based on the maximum likelihood method and adapted with R-Studio version 4.2.3, with the packages ape 5.7.1 (Paradis et al., 2004) and ggtree 3.8.2 (Guangchuang, 2022). For the summary figure illustrations of the University of Maryland Center for Environmental Science Integration and Application Network (https://ian.umces.edu/media-library/) were used and merged.

Results

Chemical analysis

The chemical analysis of the bulk soil before seeding of the two seasons of spring barley showed similar ammonium mean concentrations, similar mean C:N ratios and similar mean soil temperatures, but different nitrate mean concentrations (Table 2).

Ion torrent sequencing and pre-processing

A total of 640.438 and 3.840.352 raw sequences (two seasons collapsed, see Supplementary Figure 1 for comparison of the two seasons) were obtained for leaf samples and underground material (roots, bulk soil, rhizosphere soil), respectively. The sequence counts ranged from 12.521 to 103.274 for the leaf material and from 34.658 to 113.092 for the underground material. After quality control, denoising, sequence dereplication and chimera filtering, 108.622 sequences were removed from the leaf material and 851.482 sequences from the underground material. There were 531.816 remaining sequences for the leaf material and 2.988.870 remaining sequences for the underground material. They were grouped in 3619 cercozoan ASVs for the underground material, pre-processed (before performing differential abundance analysis to handle different types of zeros like structural zero, outlier zero, and sampling zero) to 1868 remaining ASVs. Sequences were grouped into 389 ASVs for the leaf material. Pre-processing the leaf material resulted in 63 remaining ASVs only, this step was therefore not carried out.

Taxonomic diversity

The ten most abundant ASVs in the leaf, root and soil samples could not be taxonomically determined further than the family level (hereafter: unidentified ASVs). These ASVs are marked with

TABLE 2 Mean values of the chemical analysis of the bulk soil before seeding of the two seasons of spring barley.

Spring barley season	NH_4^+ [µmol g^{-1} DW soil]	NO_3^- [µmol g^{-1} DW soil]	C:N ratio	Soil temperature [°C]
First season 2021	317.02	346.11	9.14	5.1
Second season 2022	331.36	869.25	8.06	6.3

DW, Dry weight.

numbers to distinguish them. They belonged to the families Sandonidae, Allapsidae, Cercomonadidae, Glissomonadida and Rhogostomidae (Figure 1). Leaf samples were dominated by several unidentified ASVs of the Sandonidae family (Figure 1A), whereas the root samples were dominated by several unidentified ASVs of the Allapsidae family, as well as Cercomonadidae (2) and Glissomonadida (3 and 4) (Figure 1B). At the ripening stage, unidentified ASVs of Paracercomonadidae (1) and Proleptomonadidae (1) families appeared. All soil samples were dominated by several unidentified ASVs (Figures 1C–E). The rhizosphere soil was composed of unidentified ASVs belonging to the families Sandonidae, Allapsidae, one Cercomonadidae (1) ASV and Rhogostomidae (Figure 1C), whereas the bulk soil was dominated by unidentified ASVs of the families Sandonidae, Cercomonadidae and Rhogostomidae. At the ripening stage, the unidentified ASV of the Euglyphidae (1) family appeared and before seeding unidentified ASV of the family Paracercomonadidae (2) was present (Figures 1D–E). A summary overview of these results can be seen in Supplementary Figure 2.

Alpha diversity

The observed ASV number, Shannon's and Fischer's diversity indices showed differences in the alpha diversity of the sampling material (Figure 2). There was a clear separation of the alpha diversity from the leaf material, from the soil material and the root material. There were no significant differences with the Wilcox test independent of the index considered between the leaf samples of the flowering and the ripening stages (p-value 1.00), as well as among the soil samples of the different sampling stages (p-values from 0.48 to 1.00), and the root samples of the different sampling stages (p-value 1.00). In contrast, significant differences between the different sampling materials could be detected. Alpha diversity of leaf and soil samples were significantly different independent of the index (p-values < 0.05). However, diversity indices of leaf material at flowering and ripening did not significantly differ from rhizosphere soil at ripening (p-values above 0.08). The same could be detected for the comparison of the root and soil material: no significant differences in alpha diversity could be detected between root material at flowering and ripening and rhizosphere soil at ripening (p-values above 0.08). Only significant differences could be detected, independent of the index (p-values < 0.05), of the two root samples and the other soil samples. Furthermore, only significant alpha diversity differences were found between the leaf samples and the root samples, independent of the index considered (p-values < 0.05). See the supplementary data for exact p-values (Supplementary Tables 1–3).

Beta diversity

The RPCA of the beta diversity (Aitchison distances) showed a clear separation of leaf, soil and root samples (Figure 3). Each of them formed its own group with only a few overlapping points in between. PERMANOVA showed significant differences (p-value = 0.001) between all groups, and the pairwise PERMANOVA also showed several differences (Supplementary Table 4). Leaves' community composition from the flowering stage was significantly different from that of leaves at the ripening stage (p-value 0.03). Furthermore, protist community composition of leaves from the flowering stage was significantly different in comparison to those from rhizosphere soil at flowering (p-value 0.04) and ripening (p-value 0.02), as well as those from root samples at both stages (p-values < 0.01). Leaf community composition at the flowering stage was not significantly different to the bulk soil community

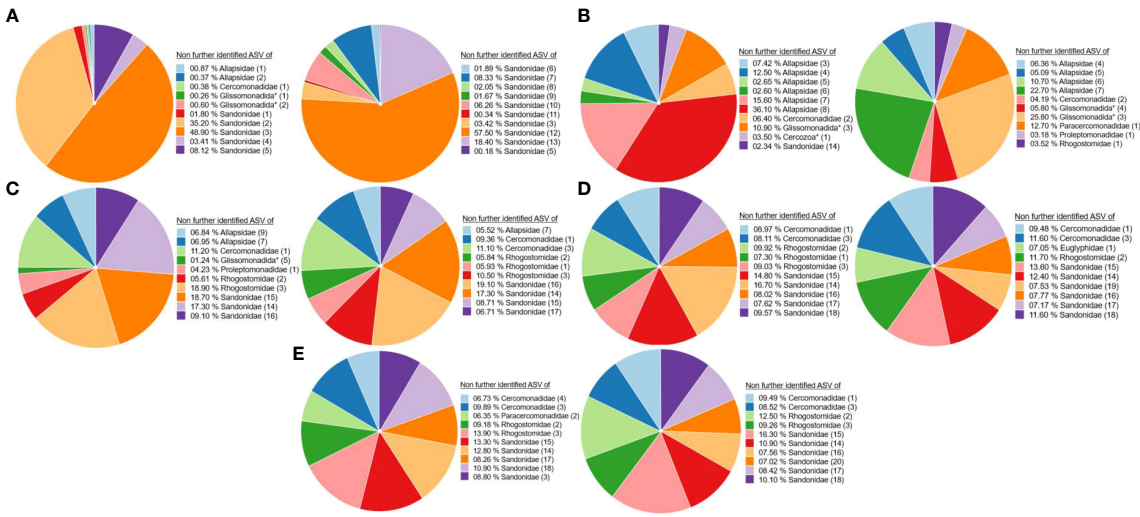


FIGURE 1
Ten most abundant (taxonomically not further determined than the family level) ASVs of cercozoan families. **(A)** Leaf samples at flowering (left) and ripening (right) stages. **(B)** Root samples at flowering (left) and ripening (right) stages. **(C)** Rhizosphere soil samples at flowering (left) and ripening (right) stages. **(D)** Bulk soil samples at flowering (left) and ripening (right) stages. **(E)** Bulk soil samples before seeding (left) and after harvesting (right). * Taxonomically not determined to family level. The number in brackets is the corresponding number of the ASV.

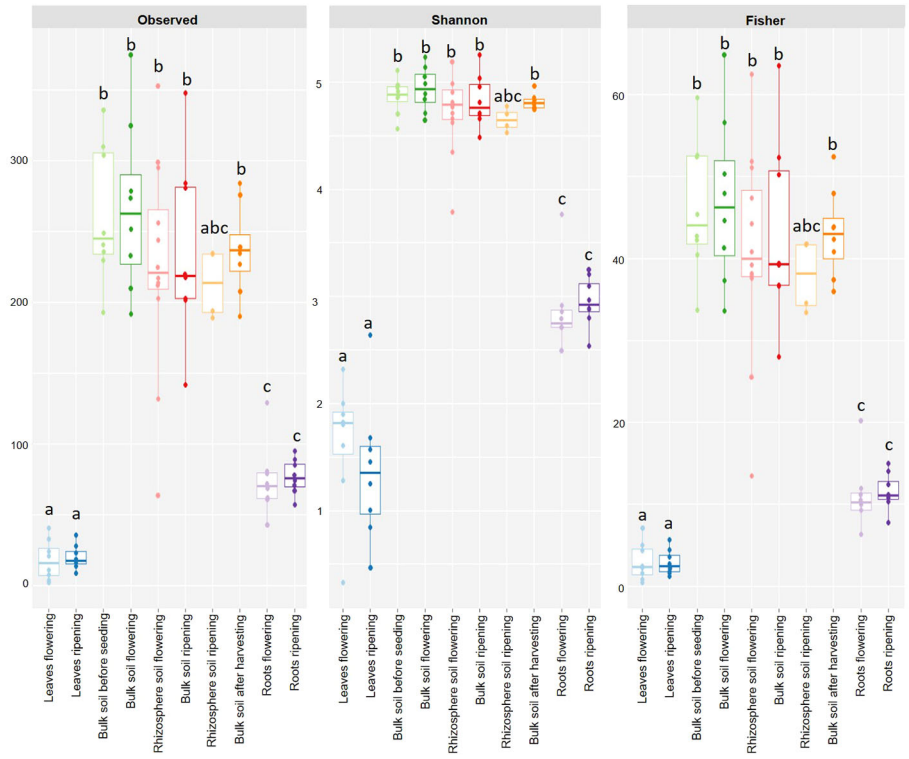


FIGURE 2
Alpha diversity measurements of leaf (blue), soil (green to orange) and root (purple) samples before seeding, at flowering, at ripening and after harvest with the observed ASVs, Shannon's and Fisher's diversity. The distinct letters delineate significantly diverse groups.

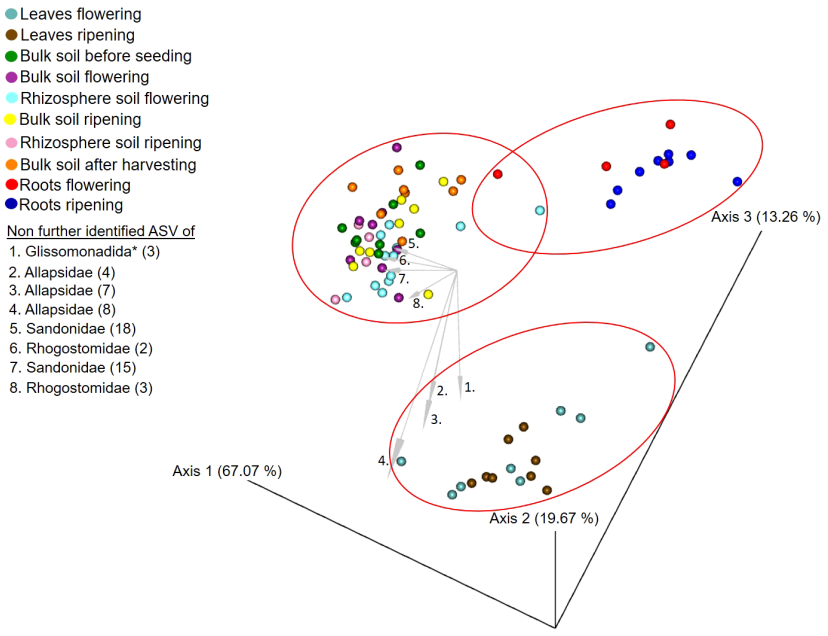


FIGURE 3
Robust principal component analysis plot of beta diversity measurement based on Aitchison distances of the leaf, soil and root samples before seeding, at flowering, at ripening and after the harvesting of the spring barley. Ellipses indicate groups of significantly different samples. Arrows indicate ASVs with the strongest influence on the beta diversity taxonomically not further determined than the family level. * Taxonomically not determined to family level. The number in brackets is the corresponding number of the ASV.

compositions before seeding, at the flowering stage, at the ripening stage and after harvest (p-values > 0.13). In contrast to this, the leaf material of the ripening stage was significantly different to all soil samples and root samples (p-values < 0.01). As for the alpha diversity, beta diversity was not significantly different between soil samples (p-values > 0.07), except for the rhizosphere soil ones. The rhizosphere soil cercozoan community composition at the flowering stage was significantly different from the one from rhizosphere soil at ripening (p-value 0.02). Furthermore, beta diversity of soil samples was significantly different from that of root samples (p-values < 0.01), but root samples' composition from the flowering and ripening stages was not significantly different from each other (p-value 0.53). See the supplementary data for exact p-values (Supplementary Table 4). In addition, Figure 3 also shows represented as arrows that unidentified ASVs belonging to the families Allapsidae, Sandonidae, Rhogostomidae and one ASV of the order Glissomonadida (3) have the biggest influence on the beta diversity.

Differential abundance

The MA plot and MW plot from root and leaf samples (Supplementary Figure 3) as well as the plots from soil and leaf samples (Supplementary Figure 5) showed an even distribution of differentially abundant ASVs in comparison to the mean (red dots), while the plots of the soil and root samples (Supplementary Figure 4) showed more red dots at the root than at the soil samples, indicating more abundant ASVs than the mean in the root samples.

Plotting of the clr value median difference between leaf and root samples (diff.btw plot of the ALDEx2 analysis) (Figure 4) showed an unidentified ASV of the family Sandonidae (3) as most abundant in leaf samples, and one of the Allapsidae (8) in root material. In parallel, the diff.btw plot of soil and root samples (Figure 5), showed an unidentified ASV of the family Tremulidae (1) as most abundant in the soil material, and Plasmodiophoridae (2) in the root material. The diff.btw plot of the ALDEx2 analysis of leaf and soil samples, showed the highest abundance in the leaf material, of an unidentified ASV of the family Sandonidae (3) and of an unidentified Allapsidae (8) in the soil material (Figure 6). The comparison of the ASVs with the biggest influence on the beta diversity from the arrows of Figure 3 and the unidentified ASVs of the families from Figures 4–6 showed two ASVs which were present in all graphics. The ASV from arrow 4, unidentified ASV of Allapsidae (8), was found in all ALDEx2 analyses as well as the ASV from arrow 8, unidentified ASV of Rhogostomidae (3).

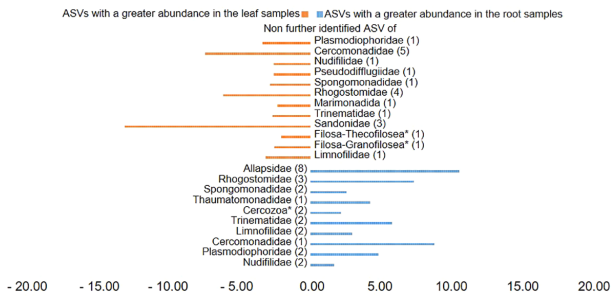


FIGURE 4
Divergent barplot with the ASVs (taxonomically not further determined than the family level) of the families that were significantly different in root and leaf samples. * Taxonomically not determined to family level. The number in brackets is the corresponding number of the ASV.

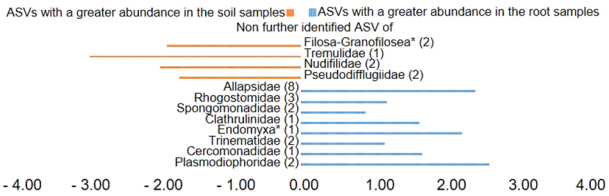


FIGURE 5
Divergent barplot with the ASVs (taxonomically not further determined than the family level) of the families that were significantly different in root and soil samples. * Taxonomically not determined to family level. The number in brackets is the corresponding number of the ASV.

Discussion

The chemical analysis of the soils where the first and second seasons of spring barley seeds were sown showed similar values for the soil parameters, except for nitrate concentration (Table 2). In the second season soil nitrate levels were higher than in the first season. Protists are very sensitive to the various changes that occur in the soil. Krashevska et al. (2014) showed that the density of living cells of protists increased with higher N inputs into the soil, and a higher amount of nitrate in the soil of the second season of barley could have also led to an altered Cercozoa community. However, the analysis of the differences in beta diversity between the two seasons using two different algorithms did not reveal clear differences. The analysis with R-Studio (phyloseq 1.28.0) of the sampling material of season one and season two showed no significant differences between the sampling material at the different sampling points for the two seasons (Supplementary Figure 1), the analysis with QIIME2 (DEICODE) of the sampling material of season one and season two showed only significant differences (besides the comparisons of the leaf samples) between the different sampling materials at the different sampling points for the two seasons. Nevertheless, this project aimed to provide a robust analysis of the natural cercozoan community of spring barley, besides the naturally occurring changes between seasons and therefore we combined both seasons and analysed them as one.

The spring barley plants of both seasons of this study were exposed to very hot and dry summers (Supplementary Table 5). Dry conditions, especially on the leaf surface, created difficult circumstances for the cercozoan protists. For this reason, only the sequencing results of the root and soil material, and not those of leaf material, were pre-processed, as the former environments were less directly affected by drought. Due to such problems, researchers

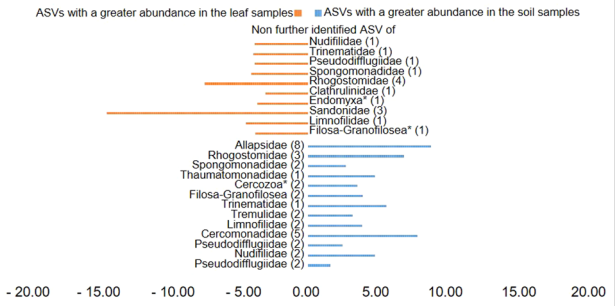


FIGURE 6
Divergent barplot with the ASVs (taxonomically not further determined than the family level) of the families that were significantly different in soil and leaf samples. * Taxonomically not determined to family level. The number in brackets is the corresponding number of the ASV.

focus, in addition to high-throughput sequencing of leaf-associated protists, on other methods like microorganisms' isolation, cloning or other plant compartments of important crops (Ploch et al., 2016; Flues et al., 2018; Rossmann et al., 2020). The unprocessed leaf material was analysed nonetheless, as it still best represents the leaves' natural cercozoan composition in this analysis, and it focuses on future problems linked to the resilience of the plant's natural cercozoan community during hot and dry summers resulting from man-made climate change.

Diversity analyses of the ten most dominant ASVs revealed a dominance of the family Sandonidae in all samples (Figure 1), at both flowering and ripening stages, except for the roots. Leaf surfaces in particular showed a high percentage of unidentified Sandonidae at the flowering (97.43%) and ripening stages (100%) (Figure 1A), while rhizosphere (Figure 1C) and bulk soils (Figures 1D, E) had lower percentages of those ASVs, although still dominant (30% for rhizosphere soil at flowering to 40% for bulk soils at the ripening stage and 60% after harvest). Interestingly, ASVs belonging to the Sandonidae family only remained on leaf surfaces by the ripening stage (Figure 1A). The lineages represented by these ASVs are likely highly adapted to drought stress on leaf surfaces, eventually forming cysts - a common trait among Cercozoa - giving them a significant advantage in changing climate conditions. Similarly, other Sandonidae ASVs were dominant in rhizosphere soils, within a particularly stable cercozoan community, with only a shift in abundance of the same ten most dominant ASVs from the flowering to the ripening stages (Figure 1C). Unfluctuating soil conditions, such as lower drought stress than on the leaves, and only a small concentration of exudates from older or dying roots at the ripening, allowed for a stable community, and selection of the most competitive lineages, further confirming the competitiveness of the Sandonidae ASVs. Common to biocrusts from grasslands and dunes (Roshan et al., 2021), Sandonidae exerted a strong dominance in bulk soils (Figures 1D, E). While there is a possibility of primer being more representative of Sandonidae 18S rRNA, Fiore-Donno et al. (2017) reported a 97% success rate in the analysis of a mock community with these primers.

In addition to Sandonidae, leaf samples also harboured unidentified ASVs of the families Allapsidae and Glissomonadida, confirming results by Ploch et al. (2016), who identified the genera *Sandona, Neoheteromita* and *Allapsa* in phyllosphere samples. The Allpasidae and Glissomondaida ASVs disappeared from leaf samples by the ripening stage, probably because those species were unable to endure the environmental conditions on the leaves (Figure 1A). Root samples were dominated by unidentified ASVs of the family Allapsidae (Figure 1B), known for inhabiting the roots and showing positive associations with plant genes involved in plant growth and development, and unidentified ASVs of the family Rhogostomidae were also associated with the cercozoan root community (Dumack et al., 2017; Rüger et al., 2023). Bulk soils before seeding, at flowering, at ripening and after harvest were also dominated by unidentified ASVs of the families Cercomonadidae, Rhogostomidae in addition to the aforementioned Sandonidae. Although slight variations in cercozoan composition were observed across different plant stages and sampling points, the top ten dominant ASVs remained mostly unchanged. Differences were raised through a strong prevalence of unidentified ASVs of the family Sandonidae in the leaf samples, unidentified ASVs of the family Allapsidae in the root samples, and mixed dominance in the soil samples. These findings were supported by the phylogenetic tree of the plant compartments - leaf, soil and roots (Supplementary Figure 6) – indicating the dominance of Sandonidae, Allapsidae and one Rhogostomidae (2) in the top ten unidentified ASVs, along with ASVs from Glissomonadida (3) and Cercomonadidae (1). In this context, it is important to note that protists possess varying quantities of 18S rRNA gene copies (Heywood et al., 2011). The variation in the number of 18S rRNA gene copies, potentially influenced by factors such as the environment, genome size, and cell size, may result in misrepresentation and overestimation of their presence, thereby compromising the accuracy of the results (Prokopowich et al., 2003; Zhu et al., 2005; Gong and Marchetti, 2019; Pan et al., 2022).

Alpha diversity analyses further confirmed compartment-specific cercozoan communities of leaf, root and soil, independent of the analysed index, with a clear separation between leaf, soil and root samples. No significant differences in alpha diversity were found between the different leaf, soil and root samples (Supplementary Tables 1-3). This means that the alpha diversity of the cercozoan community of spring barley is independent of the plant stage and was primarily influenced by the sampling material. The very low alpha diversity of the leaf samples can be explained by the very dry, hot summer and the harsh conditions aboveground, which reduced the richness of the cercozoan community from the leaves. In contrast, the very high alpha diversity observed in the soil can be explained by more favourable underground conditions concerning water availability and temperature. The alpha diversity of the root sampling material, which was far lower than the alpha diversity of the soil samples but a little bit higher than that of the leaf samples, can be explained by the influences and repelling exudates of the heat-suffering plant.

Beta diversity analyses initially confirmed previous results, with a clear separation of the diversity of the leaf, the root and the soil sampling material and only some overlaps (Figure 3). However, PERMANOVA analysis showed some interesting significant differences (Supplementary Table 4). The community of the leaves at the flowering stage was significantly different from the

community of the leaves at the ripening stage: harsh environmental conditions, such as drought and high solar radiation, and the dying leaves, which were yellow at the second sampling, created a niche for highly competitive Cercozoa. Given the previously observed strong shift of the ten most abundant phyllosphere cercozoan ASVs (Figure 1A), a significant difference in community compositions between these two stages, although not shown before, is not surprising. Furthermore, the leaf community composition at the flowering stage was significantly different in comparison to that of all the other communities of root and soil material, except for bulk soil. These significant differences support alpha diversity analysis results and the assumption of plant-compartment-specific cercozoan communities. The lack of significant differences in comparison to the bulk soil samples was more surprising. During germination, the first contact of the leaves with the bulk soil probably influenced the cercozoan composition of the leaf at flowering, explaining converging beta diversity. Moreover, wind dispersal and rainfalls also impact beta diversity (Genitsaris et al., 2014; Jauss et al., 2021). Nevertheless, environmental conditions changed the cercozoan phyllosphere community from the flowering to the ripening stage: at the ripening stage, the community was significantly different from all soil and root samples, showing a clear separation between the plant compartments again. Rhizosphere soil community was significantly different at flowering from that at ripening; community differences were most likely due to the plant's senescent state at ripening. Rhizosphere soil communities were not significantly different to bulk soil communities, and bulk soil communities did not show significant differences between each other but were significantly different to the root communities. This further confirms the plant's specific compartments, explaining why beta diversity did not differ between samples, except for really strong environmental (and subsequent plant) influences. The root communities showed no significant differences between the flowering and ripening stages, which was unexpected compared to the previous results mentioned in this paragraph and influence of the dying plant. Clearly, the exudates of the dying plant and the reduction of the dominating unidentified species of the family Allapsidae did not influence the beta diversity enough. The beta diversity was strongly driven by some single ASVs. These belonged to the unidentified species of the families Sandonidae, Allapsidae and Rhogostomidae. While unidentified ASVs of the family Allapsidae dominated the root samples, the unidentified ASVs of the family Sandonidae were present in a lot of samples and dominated the leaf communities. These drivers probably influenced the formation of the different plant compartments.

ALDEx2 analyses considered plant compartments and their specific cercozoan communities, and how these differed between leaf, root and soil. Comparison of leaf and root communities showed higher abundance than the mean of an unidentified Sandonidae (3) ASV in the leaf samples and an unidentified ASV of the Allapsidae (8) family in the root samples (Figure 4). These ASVs had the highest diff.btw values, and not unexpected given previously discussed results. The second highest diff.btw value in both compartments was shown by two different unidentified ASVs of the Cercomonadidae (5 and 1) family. They seem to be more generalist and occur in both compartments in higher abundances. There were fewer shifts in the comparison of soil and root communities than there were in the comparisons with the leaves (Figure 5). In the beta-diversity plot were also more overlapping points between those two compartments (Figure 3). The unidentified ASV of the family Tremulidae (1) had a higher abundance in the soil samples. Although not much is known about this family, the genus *Tremula* has been shown to survive dry soil conditions (Howe et al., 2011), and the protist *Tremula longifila*, usually cultured with eukaryotic prey, can also survive with bacterial prey in culture, indicating a functional flexibility under harsh environmental conditions. The unidentified ASVs of the families Plasmodiophoridae (2) and Allapsidae (8) were differentially abundant in the roots. Members of the family Plasmodiophoridae are not only plant parasites with strong differences within the species (Bulmana et al., 2001), but also well-known plant pathogens. Plasmodiophorids are also parasites and symbionts of stramenopiles (Neuhauser et al., 2014), commonly found in soils (Bates et al., 2013; Dupont et al., 2016). Their presence is therefore unsurprising. In line with the diversity analyses discussed above, unidentified ASVs of the Allapsidae family dominated the root compartment; when comparing the leaf and soil communities, the unidentified Sandonidae (3) ASV had higher abundance in the leaf samples, and the unidentified Allapsidae (8) family in the soil samples (Figure 6). The unidentified ASV of the Rhogostomidae (4) family also showed a high diff.btw value in the leaf samples. Previously found on leaves (Dumack et al., 2017), Rhogostomidae are believed to be able to withstand drier conditions thanks to their theca and resting stages (Belar, 1921). The unidentified ASV of the Cercomonadidae (5) family showed a high diff.btw value in the soil samples, similar to the comparison between leaf and root compartments. To better understand the ecological significance of these ASVs, which was not possible in this study due to the poor taxonomic resolution only up to the family level, future studies should also try to improve the taxonomic resolution. For this purpose, attempts should be made to isolate the Cercozoa species from this habitats or to find out more about their function by sequencing metagenomes, as has already been done by Thompson et al. (2020) in an Antarctic soil ecosystem or Oliverio et al. (2020) in below ground systems. However, differential abundance analysis showed greater overlaps between the compartments than previous analyses. These results were confirmed by the phylogenetic tree of the plant compartments (Supplementary Figure 6). The phylogenetic tree of the ten most dominant on family level identified ASVs showed that each of the ASVs occurred in at least one sample of the root compartment. Furthermore, the ASVs occurred in at least one sample from one of the other compartments or in samples of all compartments. Although the phylogenetic tree is dominated by the occurrence of the unidentified ASVs in the soil samples, several connecting points between the three plant compartments are pointed out.

Conclusion

Hordeum vulgare is one of the most economically important crops, and protists are an important part of the plant holobiome and influence plant growth and pathogen pressure. The natural composition of the cercozoan eukaryome is an important tool for

understanding plant performance. The sequencing results showed a clear separation between the compositions of the leaf, soil, and root samples, and a dominance of unidentified ASVs of the family Sandonidae in leaf samples, the family Allapsidae in root samples and a mixed dominance of unidentified ASVs of the families Sandonidae, Allapsidae, Cercomonadidae and Rhogostomidae in soil samples. It can be concluded that the cercozoan diversity of spring barley was highly determined by the plant compartment and not by the growing stage of the plant. Only the leaf material at flowering and ripening stages showed significant differences in the cercozoan composition, which was attributable to the strong environmental influences. However, the cercozoan community composition of rhizosphere soil, bulk soil and roots did not change significantly during plant growth. Based on those results, further analyses of the cercozoan community of other important crops could confirm the influence of the plant compartment. A more general understanding of the cercozoan community composition on the leaf, root and soil compartments of different crop plants together with analysis of the feeding behaviour of cercozoans on the crop plants could enable future recommendations for management actions in organic farming for improving plant growth and mitigating pathogen pressure.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/sra/PRJNA1043645, PRJNA1043645.

Author contributions

JS: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Writing – original draft. SR: Conceptualization, Investigation, Methodology, Supervision, Validation, Writing – review & editing. SQ-Q: Conceptualization, Investigation, Methodology, Writing – review & editing. RG-P: Conceptualization, Investigation, Methodology, Writing – review & editing. BS: Investigation, Methodology, Writing – review & editing. AÖ: Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing. SS: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Work was supported by a fully funded doctoral fellowship for JS from the German Federal Environment Foundation (Deutsche Bundesstiftung Umwelt) and by the Federal Ministry of Education and Research. The open access publication was funded by the open access publication fund of the Justus-Liebig-University.

Acknowledgments

The authors would like to thank Dr. Franz Schulz and the team of the Gladbacherhof for their assistance on the field. Furthermore, the authors would like to thank Renate Baumann for her technical assistance.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2024.1352566/full#supplementary-material

References

Abdullaeva, Y., Ambika Manirajan, B., Honermeier, B., Schnell, S., and Cardinale, M. (2021). Domestication affects the composition, diversity, and co-occurrence of the cereal seed microbiota. *J. Adv. Res.* 31, 75-86. doi: 10.1016/j.jare.2020.12.008

Aitchison, J. (1982). The statistical analysis of compositional data. *J. R. Statist. Soc* 44, 2.

Aitchison, J. (1986). "The statistical snalysis of compositional data," in *Monographs on statistics and applied probability*. Eds. N. V. Cox, N. Hinkley, N. Rubin and B. W. Silverman (New York: Chapman and Hall Ltd).

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*. 26, 1. doi: 10.1111/j.1442-9993.2001.tb00081.x

Asiloglu, R., Shiroishi, K., Suzuki, K., Turgay, O. C., Murase, J., and Harada, N. (2020). Protist-enhanced survival of a plant growth promoting rhizobacteria, Azospirillum sp. B510, and the growth of rice (*Oryza sativa* L.) plants. *Appl. Soil Ecology*. 154, 103599. doi: 10.1016/j.apsoil.2020.103599

Bahroun, A., Jousset, A., Mrabet, M., Mhamdi, R., and Mhadhbi, H. (2021). Protists modulate Fusarium root rot suppression by beneficial bacteria. *Appl. Soil Ecology*. 168, 104158. doi: 10.1016/j.apsoil.2021.104158

Bak, F., Scheff, G., and Jansen, K. H. (1991). A rapid and sensitive ion chromatographic technique for the determination of sulfate and sulfate reduction

rates in freshwater lake sediments. *FEMS Microbiol. Letters.* 85, 1. doi: 10.1111/j.1574-6968.1991.tb04694.x

Bass, D., and Cavalier-Smith, T. (2004). Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). *Int. J. Systematic Evolutionary Microbiol.* 54, 6. doi: 10.1099/ijs.0.63229-0

Bates, S. T., Clemente, J. C., Flores, G. E., Walters, W. A., Parfrey, L. W., Knight, R., et al. (2013). Global biogeography of highly diverse protistan communities in soil. *ISME J.* 7, 3. doi: 10.1038/ismej.2012.147

Belar, K. (1921). Untersuchungen über Thecamöben der Chlamydophrys-Gruppe. *Arch. Protistenkd.* 43, 287–354.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Society: Ser. B (Methodological).* 57, 1. doi: 10.1111/j.2517-6161.1995.tb02031.x

Bisanz, J. E. (2018). *qiime2R: Importing QIIME2 artifacts and associated data into R sessions (v0.99).* Available at: https://github.com/jbisanz/qiime2R

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome.* 6, 1. doi: 10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J., Dillon, M., Bokulich, N., Abnet, C., Al-Ghalith, G., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 8. doi: 10.1038/s41587-019-0209-9

Bulmana, S. R., Kühn, S. F., Marshalla, J. W., and Schnepf, E. (2001). A phylogenetic analysis of the SSU rRNA from members of the Plasmodiophorida and Phagomyxida. *Protist* 152.

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 12. doi: 10.1038/ismej.2017.119

Cardinale, M., Ratering, S., Sadeghi, A., Pokhrel, S., Honermeier, B., and Schnell, S. (2020). The response of the soil microbiota to long-term mineral and organic nitrogen fertilization is stronger in the bulk soil than in the rhizosphere. *Genes.* 11, 4. doi: 10.3390/genes11040456

Cavalier-Smith, T., and Chao, E. E. Y. (2003). Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist.* 154, 3–4. doi: 10.1078/143446103322454112

Ceja-Navarro, J. A., Wang, Y., Ning, D., Arellano, A., Ramanculova, L., Yuan, M. M., et al. (2021). Protist diversity and community complexity in the rhizosphere of switchgrass are dynamic as plants develop. *Microbiome.* 9, 1. doi: 10.1186/s40168-021-01042-9

Chełkowski, J., Tyrka, M., and Sobkiewicz, A. (2003). Resistance genes in barley (*Hordeum vulgare* L.) and their identification with molecular markers. *J. Appl. Genet.* 44, 3. 291–3. 309.

Degrune, F., Dumack, K., Fiore-Donno, A. M., Bonkowski, M., Sosa-Hernández, M. A., Schloter, M., et al. (2019). Distinct communities of Cercozoa at different soil depths in a temperate agricultural field. *FEMS Microbiol. Ecology.* 95, 4. doi: 10.1093/femsec/fiz041

Dumack, K., Ferlian, O., Morselli Gysi, D., Degrune, F., Jauss, R. T., Walden, S., et al. (2022). Contrasting protist communities (Cercozoa: Rhizaria) in pristine and earthworm-invaded North American deciduous forests. *Biol. Invasions*, 0123456789. doi: 10.1007/s10530-021-02726-x

Dumack, K., Flues, S., Hermanns, K., and Bonkowski, M. (2017). Rhogostomidae (Cercozoa) from soils, roots and plant leaves (*Arabidopsis thaliana*): Description of Rhogostoma epiphylla sp. nov. and *R. cylindrica* sp. nov. *Eur. J. Protistology* 60, 76-86. doi: 10.1016/j.ejop.2017.06.001

Dumack, K., Pundt, J., and Bonkowski, M. (2019). Food choice experiments indicate selective fungivorous predation in *Fisculla terrestris* (Thecofilosea, Cercozoa). *J. Eukaryotic Microbiol.* 66, 3. doi: 10.1111/jeu.12680

Dupont, A.Ö.C., Griffiths, R. I., Bell, T., and Bass, D. (2016). Differences in soil micro-eukaryotic communities over soil pH gradients are strongly driven by parasites and saprotrophs. *Environ. Microbiol.* 18, 6. doi: 10.1111/1462-2920.13220

Fernandes, A. D., Macklaim, J. M., Linn, T. G., Reid, G., and Gloor, G. B. (2013). ANOVA-Like Differential Expression (ALDEx) Analysis for mixed population RNA-Seq. *PloS One* 8, 7. doi: 10.1371/journal.pone.0067019

Fiore-Donno, A. M., Rixen, C., Rippin, M., Glaser, K., Samolov, E., Karsten, U., et al. (2017). New barcoded primers for efficient retrieval of cercozoan sequences in high-throughput environmental diversity surveys, with emphasis on worldwide biological soil crusts. *Mol. Ecol. Resources.* 18, 2. doi: 10.1111/1755-0998.12729

Fisher, R. A., Corbet, A. S., and Williams, C. B. (1943). The relation between the number of Species and the number of individuals in a random sample of an animal population. *J. Anim. Ecology.* 12, 1. doi: 10.2307/1411

Flues, S., Blokker, M., Dumack, K., and Bonkowski, M. (2018). Diversity of Cercomonad species in the phyllosphere and rhizosphere of different plant species with a description of *Neocercomonas epiphylla* (Cercozoa, Rhizaria) a leaf-associated protist. *J. Eukaryotic Microbiol.* 65, 5. doi: 10.1111/jeu.12503

Geisen, S. (2016). The bacterial-fungal energy channel concept challenged by enormous functional versatility of soil protists. *Soil Biol. Biochem.* 102, 1–4. doi: 10.1016/j.soilbio.2016.06.013

Geisen, S., Tveit, A. T., Clark, I. M., Richter, A., Svenning, M. M., Bonkowski, M., et al. (2015). Metatranscriptomic census of active protists in soils. *ISME J.* 9, 10. doi: 10.1038/ismej.2015.30

Genitsaris, S., Kormas, K. A., Christaki, U., Monchy, S., and Moustaka-Gouni, M. (2014). Molecular diversity reveals previously undetected air-dispersed protist colonists in a Mediterranean area. *Sci. Total Environ.* 478, 70–79. doi: 10.1016/j.scitotenv.2014.01.071

Glaser, K., Kuppardt, A., Boenigk, J., Harms, H., Fetzer, I., and Chatzinotas, A. (2015). The influence of environmental factors on protistan microorganisms in grassland soils along a land-use gradient. *Sci. Total Environ.* 537, 33–42. doi: 10.1016/j.scitotenv.2015.07.158

Gong, W., and Marchetti, A. (2019). Estimation of 18S gene copy number in marine eukaryotic plankton using a next-generation sequencing approach. *Front. Mar. Sci.* 6 (APR). doi: 10.3389/fmars.2019.00219

Guangchuang, Y. (2022). *Data integration, manipulation and visualization of phylogenetic trees. 1st edition* (Florida: Chapman and Hall/CRC).

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D1. doi: 10.1093/nar/gks1160

Guo, S., Xiong, W., Xu, H., Hang, X., Liu, H., Xun, W., et al. (2018). Continuous application of different fertilizers induces distinct bulk and rhizosphere soil protist communities. *Eur. J. Soil Biol.* doi: 10.1016/j.ejsobi.2018.05.007

Heywood, J. L., Sieracki, M. E., Bellows, W., Poulton, N. J., and Stepanauskas, R. (2011). Capturing diversity of marine heterotrophic protists: One cell at a time. *ISME J.* 5 (4), 674–684. doi: 10.1038/ismej.2010.155

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian J. Statistics.* 6, 2. 65–2. 70.

Howe, A. T., Bass, D., Scoble, J. M., Lewis, R., Vickerman, K., Arndt, H., et al. (2011). Novel cultured protists identify deep-branching environmental DNA clades of Cercozoa: New Genera Tremula, Micrometopion, Minimassisteria, Nudifila, Peregrinia. *Protist.* 162, 2. doi: 10.1016/j.protis.2010.10.002

Imparato, V., Santos, S. S., Johansen, A., Geisen, S., and Winding, A. (2016). Stimulation of bacteria and protists in rhizosphere of glyphosate-treated barley. *Appl. Soil Ecol.* 98, 47–55. doi: 10.1016/j.apsoil.2015.09.007

Jauss, R. T., Nowack, A., Walden, S., Wolf, R., Schaffer, S., Schellbach, B., et al. (2021). To the canopy and beyond: Air dispersal as a mechanism of ubiquitous protistan pathogen assembly in tree canopies. *Eur. J. Protistology* 80. doi: 10.1016/j.ejop.2021.125805

Jentschke, G., Bonkowski, M., Godbold, D. L., and Scheu, S. (1995). Soil protozoa and forest tree growth: non-nutritional effects and interaction with mycorrhizae. *Biol. Fertility Soils.* 20, 4. doi: 10.1007/BF00336088

Kandeler, E., and Gerber, H. (1988). Short-term assay of soil urease activity using colorimetric determination of ammonium. *Biol. Fertility Soils.* 6, 1. doi: 10.1007/BF00257924

Kloepper, J. W. (1981). Development of a powder formulation of rhizobacteria for inoculation of potato seed pieces. *Phytopathology.* 71, 6. doi: 10.1094/phyto-71-590

Krashevska, V., Sandmann, D., Maraun, M., and Scheu, S. (2014). Moderate changes in nutrient input alter tropical microbial and protist communities and belowground linkages. *ISME J.* 8, 5. doi: 10.1038/ismej.2013.209

Krome, K., Rosenberg, K., Dickler, C., Kreuzer, K., Ludwig-Müller, J., Ullrich-Eberius, C., et al. (2010). Soil bacteria and protozoa affect root branching via effects on the auxin and cytokinin balance in plants. *Plant Soil.* 328, 1. doi: 10.1007/s11104-009-0101-3

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 1. 10–1. 12.

Martino, C., Morton, J. T., Marotz, C. A., Thompson, L. R., Tripathi, A., Knight, R., et al. (2019). A novel sparse compositional technique reveals microbial perturbations. *Am. Soc. Microbiol. J.* 4. doi: 10.1128/mSystems.00016-19

McMurdie, P. J., and Holmes, S. (2014). Waste not, Want not: Why rarefying microbiome Data is inadmissible. *PloS Comput. Biol.* 10, 4. doi: 10.1371/journal.pcbi.1003531

Nesbitt, M., and Samuel, D. (1996). From staple crop to extinction? The archaeology and history of the hulled wheats. In: S. Padulosi, K. Hammer and J Heller (eds) *Hulled wheats. Proceedings of the first international workshop on hulled wheats, on 21–22 July 1995, Castelvecchio Pascoli, Tuscany, Italy.* (Rome: International Plant Genetic Resources Institute (IPGRI)), 41–100.

Neuhauser, S., Kirchmair, M., Bulman, S., and Bass, D. (2014). Cross-kingdom host shifts of phytomyxid parasites. *BMC Evolutionary Biol.* 14, 33.

Oliverio, A. M., Geisen, S., Delgado-Baquerizo, M., Maestre, F. T., Turner, B. L., and Fierer, N. (2020). The global-scale distributions of soil protists and their contributions to belowground systems. *Sci. Adv.* 6, eaax8787. doi: 10.1126/sciadv.aax8787

Öztoprak, H., Walden, S., Heger, T., Bonkowski, M., and Dumack, K. (2020). What drives the diversity of the most abundant terrestrial cercozoan family (Rhogostomidae, cercozoa, rhizaria)? *Microorganisms.* 8, 8. doi: 10.3390/microorganisms8081123

Pan, Y., Li, G., Su, L., Zheng, P., Wang, Y., Shen, Z., et al. (2022). Seagrass colonization alters diversity, abundance, taxonomic, and functional community structure of benthic microbial eukaryotes. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.901741

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics.* 20, 2. doi: 10.1093/bioinformatics/btg412

Ploch, S., Rose, L. E., Bass, D., and Bonkowski, M. (2016). High diversity revealed in leaf-associated protists (Rhizaria: Cercozoa) of Brassicaceae. *J. Eukaryotic Microbiol.* 63, 5. doi: 10.1111/jeu.12314

Prokopowich, C. D., Gregory, T. R., and Crease, T. J. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome* 46 (1), 48–50. doi: 10.1139/g02-103

R Core Team (2020). *R: A language and environment for statistical computingle (4.0.3)* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/

Roshan, S. K., Dumack, K., Bonkowski, M., Leinweber, P., Karsten, U., and Glaser, K. (2021). Taxonomic and functional diversity of heterotrophic protists (Cercozoa and endomyxa) from biological soil crusts. *Microorganisms*. 9, 2. doi: 10.3390/microorganisms9020205

Rossmann, M., Pérez-Jaramillo, J. E., Kavamura, V. N., Chiaramonte, J. B., Dumack, K., Fiore-Donno, A. M., et al. (2020). Multitrophic interactions in the rhizosphere microbiome of wheat: From bacteria and fungi to protists. *FEMS Microbiol. Ecol.* 96, 4. doi: 10.1093/femsec/fiaa032

Rüger, L., Ganther, M., Freudenthal, J., Jansa, J., Heintz-Buschart, A., Tarkka, M. T., et al. (2023). Root cap is an important determinant of rhizosphere microbiome assembly. *New Phytol.* 239, 1434–1448. doi: 10.1111/nph.19002

Sacharow, J., Salehi-Mobarakeh, E., Ratering, S., Imani, J., Österreicher Cunha-Dupont, A., and Schnell, S. (2023). Control of *Blumeria graminis* f. sp. *hordei* on barley leaves by treatment with fungi-Consuming protist isolates. *Curr. Microbiol.* 80, 12. doi: 10.1007/s00284-023-03497-5

Santos, S. S., Schöler, A., Nielsen, T. K., Hansen, L. H., Schloter, M., and Winding, A. (2020). Land use as a driver for protist community structure in soils under agricultural use across Europe. *Sci. Total Environ.* 717, 137228. doi: 10.1016/j.scitotenv.2020.137228

Shannon, C. E., and Weaver, W. (1964). *The theory of mathematical communication* (Urbana: International Business). Tenth printing.

Thompson, A. R., Geisen, S., and Adams, B. J. (2020). Shotgun metagenomics reveal a diverse assemblage of protists in a model Antarctic soil ecosystem. *Environ. Microbiol.* 22 (11), 4620–4632. doi: 10.1111/1462-2920.15198

Walden, S., Jauss, R. T., Feng, K., Fiore-Donno, A. M., Dumack, K., Schaffer, S., et al. (2021). On the phenology of protists: Recurrent patterns reveal seasonal variation of protistan (Rhizaria: Cercozoa and Endomyxa) communities in tree canopies. *FEMS Microbiol. Ecology*. 97, 7. doi: 10.1093/femsec/fiab081

Weidner, S., Latz, E., Agaras, B., Valverde, C., and Jousset, A. (2017). Protozoa stimulate the plant beneficial activity of rhizospheric pseudomonads. *Plant Soil.* 410, 1–2. doi: 10.1007/s11104-016-3094-8

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (Springer Verlag New York).

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biom. Bull.* 1, 80–83.

Wu, C., Ge, C., and Wang, F. (2022). Phagotrophic protist - mediated control of *Polymyxa graminis* in the wheat rhizosphere. *Plant Soil*, 0123456789. doi: 10.1007/s11104-022-05829-z

Zhang, W. L., Lin, Q. M., Li, G. T., and Zhao, X. R. (2022). The ciliate protozoan *Colpoda cucullus* can improve maize growth by transporting soil phosphates. *J. Integr. Agriculture*. 21, 3. doi: 10.1016/S2095-3119(21)63628-6

Zhao, Z. B., He, J. Z., Geisen, S., Han, L. L., Wang, J. T., Shen, J. P., et al. (2019). Protist communities are more sensitive to nitrogen fertilization than other microorganisms in diverse agricultural soils. *Microbiome*. 7, 1. doi: 10.1186/s40168-019-0647-0

Zhu, F., Massana, R., Not, F., Marie, D., and Vaulot, D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* 52 (1), 79–92. doi: 10.1016/j.femsec.2004.10.006

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 26 March 2024
DOI 10.3389/frmbi.2024.1331341

OPEN ACCESS

EDITED BY
Jun Zhao,
Nanjing Normal University, China

REVIEWED BY
Anthony Yannarell,
University of Illinois at Urbana-Champaign, United States
Andres Argüelles-Moyao,
University of São Paulo, Brazil

*CORRESPONDENCE
Lisa M. Markovchick
lisa_handforth@nps.gov

RECEIVED 31 October 2023
ACCEPTED 04 March 2024
PUBLISHED 26 March 2024

CITATION
Markovchick LM, Belgara-Andrew A, Richard D, Deringer T, Grady KC, Hultine KR, Allan GJ, Whitham TG, Querejeta JI and Gehring CA (2024) Utilizing symbiotic relationships and assisted migration in restoration to cope with multiple stressors, and the legacy of invasive species. *Front. Microbiomes* 3:1331341. doi: 10.3389/frmbi.2024.1331341

Utilizing symbiotic relationships and assisted migration in restoration to cope with multiple stressors, and the legacy of invasive species

Lisa M. Markovchick[1,2*], Abril Belgara-Andrew[1,2], Duncan Richard[1], Tessa Deringer[1,3], Kevin C. Grady[4], Kevin R. Hultine[5], Gerard J. Allan[1], Thomas G. Whitham[1,2], José Ignacio Querejeta[6] and Catherine A. Gehring[1,2]

[1]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, United States, [2]Center for Adaptable Western Landscapes, Northern Arizona University, Flagstaff, AZ, United States, [3]Rocky Mountain Research Station, United States Department of Agriculture (USDA) Forest Service, Flagstaff, AZ, United States, [4]School of Forestry, Northern Arizona University, Flagstaff, AZ, United States, [5]Department of Research, Conservation and Collections, Desert Botanical Garden, Phoenix, AZ, United States, [6]Soil and Water Conservation Group, Spanish National Research Council, Centro de Edafología y Biología Aplicada del Segura (CEBAS)-Consejo Superior de Investigaciones Científicas (CSIC), Murcia, Spain

Introduction: Climate change has increased the need for forest restoration, but low planting success and limited availability of planting materials hamper these efforts. Invasive plants and their soil legacies can further reduce restoration success. Thus, strategies that optimize restoration are crucial. Assisted migration and inoculation with native microbial symbiont communities have great potential to increase restoration success. However, assisted migrants can still show reduced survival compared to local provenances depending on transfer distance. Inoculation with mycorrhizal fungi, effective if well-matched to plants and site conditions, can have neutral to negative results with poor pairings. Few studies have examined the interaction between these two strategies in realistic field environments where native plants experience the combined effects of soil legacies left by invasive plants and the drought conditions that result from a warming, drying climate.

Methods: We planted two ecotypes (local climate and warmer climate) of *Populus fremontii* (Fremont cottonwoods), in soils with and without legacies of invasion by *Tamarix* spp. (tamarisk), and with and without addition of native mycorrhizal fungi and other soil biota from the warmer climate.

Results: Four main results emerged. 1) First year survival in soil legacies left behind after tamarisk invasion and removal was less than one tenth of survival in soil without a tamarisk legacy. 2) Actively restoring soil communities after tamarisk removal tripled first year cottonwood survival for both ecotypes, but only improved survival of the warmer, assisted migrant ecotype trees in year two. 3) Actively restoring soil communities in areas without a tamarisk history reduced first year survival for both ecotypes, but improved survival of the warmer, assisted

migrant ecotype trees in year two. 4) By the second year, inoculated assisted migrants survived at five times the rate of inoculated trees from the local ecotype.

Discussion: Results emphasize the detrimental effects of soil legacies left after tamarisk invasion and removal, the efficacy of assisted migration and restoring soil communities alongside plants, and the need to thoughtfully optimize pairings between plants, fungi, and site conditions.

KEYWORDS

assisted migration, mycorrhiza, restoration, ecotype, invasive species, physiology, *Populus fremontii*, translocation

1 Introduction

Climate change is increasing the need for, and importance of, natural regeneration, replanting, and restoration. In addition to the importance of sequestering carbon, replanting and restoration is increasingly needed following natural disasters that have increased in frequency, size and severity (National Academies of Sciences, Engineering, and Medicine, 2020; Parks and Abatzoglou, 2020; Fargione et al., 2021; Parks et al., 2023). However, planting material shortages, and climate change constrain both restoration efforts and their success (Wheeler et al., 2015; National Academies of Sciences, Engineering, and Medicine, 2020; Fargione et al., 2021; Simonson et al., 2021). For these reasons, optimizing the effectiveness of restoration is becoming increasingly important.

Climate change may outpace the ability of plants to adapt or migrate via natural dispersal, such that the practice of using local seed stock and planting material for restoration may not continue to be optimal (Whitham et al., 2020). Introducing additional genetic diversity into ecosystems using strategic assisted migration of plant provenances (local populations), or species, from warmer or drier locations, has been suggested as a strategy to help ecosystems adapt (Etterson et al., 2020; Gomory et al., 2020; Saenz-Romero et al., 2021). However, assisted migrant plants can be more susceptible to disease and can yield lower survival than local provenances due to adaptations to local temperature regimes including cold extremes (Tiscar et al., 2018; Cooper et al., 2019; Simler et al., 2019). Michalet et al. (2023) also warn of consequences for fire risk and climate buffering. Translocated Fremont cottonwood demonstrate survival correlated with transfer distance or temperature difference. Translocated Fremont cottonwoods (*Populus fremontii*) can have survival comparable to local provenances if the temperature differences between transplanted and home site are small, but survival reduced by up to 30% in trees transferred across greater temperature differences, especially those upwards of 3°C (Grady et al., 2011; 2015; Cooper et al., 2019). The local suitability of Fremont cottonwood for the shifting climate is affected by adaptations of their population and ecotype (Ikeda et al., 2017; Cooper et al., 2019; Blasini et al., 2021; 2022). Trees from the warmer Sonoran Desert (SD) ecotype have developed physiological traits which promote water movement through the plant for evaporative cooling (but may leave trees susceptible to drought-driven stress or frost), while trees from the cooler Mogollon Rim (MR) ecotype have developed traits which reduce their susceptibility to frost (which may help with drought-resilience, but result in more susceptibility to heat stress; Hultine et al., 2020; Blasini et al., 2021; 2022). Consequently, regional and population level adaptations can be an important consideration for sourcing planting material, but it is not always clear how to optimize provenance selection. There is also an urgent need for large-scale, multi-species experimental studies to provide evidence that assisted migration is advisable given local adaptation of plants to other aspects of the abiotic environment besides temperature, and to address implications for interspecies interactions including those between plants and soil microbes (Bucharova, 2017).

The need for experimental studies that assess the interactions between the assisted migration of plants and the microbiome intersects with another strategy suggested to improve restoration outcomes, stretch the impact of available planting material, and help assisted migrant plants adapt to new sites. Active restoration of diverse native mycorrhizal fungal communities symbiotic with native plant roots has shown great potential as a tool to increase revegetation success. Many kinds of disturbances negatively impact mycorrhizal fungi, such as pollution deposition, land use changes, invasion by exotic species, pesticide application, and climate change (e.g. Egerton-Warburton and Allen, 2000; Meinhardt and Gehring, 2012; Helander et al., 2018; León-Sánchez et al., 2018; Querejeta et al., 2021; IPBES, 2023). In particular, invasion by exotic plant species has been documented to simultaneously reduce native plant biomass, survival, and diversity in conjunction with reducing or shifting their mycorrhizal fungal partner communities. Invasive plants including *Cynodon dactylon* (Bermuda grass), *Calamagrostis epigejos* (bush grass), *Avena barbata* (slender wild oats), *Bromus hordeaceus* (soft chess), *Bothriochloa bladhii* (old world bluestem), *Tamarix* spp. (tamarisk), garlic mustard (*Alliaria petiolata*) and others (Hawkes et al., 2006; Wolfe et al., 2008; Wilson et al., 2012;

Endresz and Kalapos, 2013) have all been shown to alter mycorrhizal communities and impact native plant productivity and survival.

In the native riparian ecosystems of the southwestern United States, invasive tamarisk (*Tamrix* spp.) trees have repeatedly been shown to negatively impact the survival and biomass of Fremont cottonwood trees (*Populus fremontii*), and concurrently deplete and shift their active mycorrhizal fungal communities (e.g. Beauchamp et al., 2005; 2006; Gitlin et al., 2006; Meinhardt and Gehring, 2012; Hultine et al., 2015). Decreases in diversity and shifts in mycorrhizal fungal communities from invasive species seem to arise from two potential but not mutually exclusive causes, changes in soil chemistry caused by the invasive species, and reductions and shifts in the microbial and/or mycorrhizal fungal community due to reduced native plant hosts. For example, *Solidago canadensis* (Canada goldenrod) can shift the fungal community to one that promotes its own competitiveness over that of native plants (Zhang et al., 2010), and some invasive plant species such as tamarisk and garlic mustard are known to alter soil salinity, nitrate concentrations, or produce antifungal phytochemicals and allelochemicals (Lesica and DeLuca, 2004; Wolfe et al., 2008; Meinhardt and Gehring, 2012). However, it remains unclear if changes to soil chemistry and/or the impacts on mycorrhizal fungi and cottonwood survival and biomass persist as a legacy in the soil (soil legacy) once invasive tamarisk plants are removed, or the degree to which active additions of soil microbial communities could offset these effects.

Actively adding diverse, native soil microbial communities appropriate to plant host species, and provenances (often done via actively collecting, cultivating, and adding live soil inoculum from reference or non-degraded sites into degraded sites along with planting) following invasive plant removal could improve restoration outcomes. In particular, we use this technique to restore native mycorrhizal fungal communities, due to the body of literature demonstrating its success with mycorrhizal fungi (referred to hereafter as mycorrhizal restoration), and the research on Fremont cottonwoods showing that tamarisk negatively affects cottonwood survival, biomass, and mycorrhizal associations. Restoring mycorrhizal fungal communities in this way has been demonstrated to promote survival, growth, and diversity of plant communities, if thoughtfully incorporated into restoration planning in conjunction with plant provenances and communities, planting palettes and site conditions (e.g. Rua et al., 2016; Wubs et al., 2016; Koziol and Bever, 2017). At least one meta-analysis has shown that the benefits improve with time since planting (Neuenkamp et al., 2019). These findings concur with broader science reflecting that the ability of organisms to survive and adapt is likely dependent on the confluence of their own traits and adaptations with those of their microbiome (Zilber-Rosenberg and Rosenberg, 2008; Bordenstein and Theis, 2015 ; Whitham et al., 2020). In fact, because of shorter life cycles, microbiota like mycorrhizal fungi may be particularly crucial to promoting swift adaptation to climate change and rehabilitation of degraded systems (Wilkinson and Dickinson, 1995; Coban et al., 2022; Allsup et al., 2023; Argüelles-Moyao and Galicia, 2023).

However, consideration of factors including timing, soil properties, plant species and plant provenance have proven integral in achieving optimal results when implementing mycorrhizal restoration (e.g., Johnson et al., 1992; Mortimer et al., 2005; Maltz and Treseder, 2015). Neutral to negative effects occur when these factors are not adequately addressed or mass-produced products are utilized (e.g., Maltz and Treseder, 2015; Rua et al., 2016; Saloman et al., 2022), or when plant roots do not gain direct contact with active mycorrhizal fungi that are provided by mycorrhizal symbioses living on adjacent native plants (Jones et al., 2003; Grünfeld et al., 2022). Thus, it remains unclear how well optimized plant and fungal pairings must be, and how assisted migration and the tendency for co-adaptation between local soil, mycorrhizal fungi, and plants indicated by other studies (e.g. Johnson et al., 1992; 2010; 2014; Rua et al., 2016) will interact. Filling these knowledge gaps can inform the cost/benefit calculations for implementing active restoration of native mycorrhizal communities when it is needed (for example, see Hart et al., 2017), and assist in the design of inoculum and plant provenance combinations that best harness biotic co-adaptation and site conditions.

To address these knowledge gaps, we investigated the interactions between three experimental treatments under stressful drought conditions: assisted migration (from a warmer ecotype), inoculation with mycorrhizal fungi and associated soil organisms (collected from beneath a warmer ecotype), and soil legacies left by invasive tamarisk. We focus here on Fremont cottonwood due to its foundational nature in riparian communities of the arid southwest, extensive data on its ecotypes and ecophysiology, and the immense loss of riparian areas experienced in the United States. Riparian arteries of the southwestern U.S. comprise less than 0.5% and 2% of the land area in Arizona and the southwest, respectively, but support a disproportionate 60 to 75% of the wildlife (Poff et al., 2012). Best estimates placed the loss of these riparian habitats at 90% (Zaimes, 2007). Previous studies have also demonstrated that warmer, supraoptimal temperatures reduce canopy gas exchange via reduced stomatal conductance (Grady et al., 2013), increase plant stress (Hultine et al., 2013), and are expected to create predictable, linear losses in tree productivity (Grady et al., 2011). Importantly, with increasing aridity and record temperatures, recent heat waves have exceeded T_{crit}, the temperature at which Photosystem II is disrupted in cottonwoods, requiring evaporative leaf cooling to maintain photosynthesis. Combined with declining stream flows and water tables, these high temperatures threaten some cottonwood with catastrophic die offs (Moran et al., 2023). Because mycorrhizal associations are known to assist some plants in water uptake, movement, and water use efficiency (e.g. Querejeta et al., 2003; 2006; 2007; Egerton-Warburton et al., 2007), they could play an important role in affecting the survival of these trees and in restoration. We used provenance trials with assisted migration to replant a riparian corridor and floodplain after tamarisk invasion. We planted trees from six provenances, three each from two contrasting Fremont cottonwood ecotypes (warmer SD and cooler MR) at a cooler MR ecotype site.

Given prior research regarding the negative impacts of tamarisk neighbors on native cottonwoods and their mycorrhizal communities (Meinhardt and Gehring, 2012), the long-term effects of mycorrhizal disruptions and dispersal challenges (Peay et al., 2010; Pankova et al., 2018; Anthony et al., 2019), benefits of mycorrhizal inoculation (Neuenkamp et al., 2019), and challenges with assisted migration (Grady et al., 2015; Tiscar et al., 2018; Cooper et al., 2019; Simler et al., 2019), we hypothesized that: 1) the soil legacy left by prior tamarisk invasion would reduce plant survival, even after tamarisk removal; mycorrhizal restoration would improve the survival of trees from both ecotypes 2) in areas with a tamarisk soil legacy and 3) also in areas without a tamarisk history (since even non-tamarisk areas were degraded and inoculum was native and regionally appropriate); 4) assisted migrant trees from the warmer SD ecotype would have survival rates equal to or lower than those for trees from the local MR ecotype, regardless of tamarisk and inoculation treatments (given prior results from assisted migration trials with Fremont cottonwoods; e.g. Grady et al., 2011, 2015; Cooper et al., 2019).

2 Materials and methods

2.1 Common garden planting site

We chose a site along the Little Colorado River near Cameron, AZ, to establish an experimental common garden (35.71923, −111.3194). The site consists of 22 hectares in a riparian corridor and flood plain. The site falls within the cooler Mogollon Rim ecotype (MR) of Fremont cottonwoods (*Populus fremontii*). The site had been used as source for gravel since at least 2007. All areas included within the site were degraded. Native shrubs and trees that could provide the direct contact with mycorrhizal roots needed for *in situ* mycorrhizal inoculation of plantings were absent from the experimental areas. Areas without tamarisk often contained other invasive species such as *Alhagi maurorum* (camelthorn).

2.2 Cottonwoods and planting methods for the field experiment

Cottonwood cuttings were collected from trees at least 20 m apart during the 2014–2015 winter. Cuttings were collected from six provenances across two ecotypes (Mogollon Rim, MR; Sonoran Desert, SD; Blasini et al., 2021, 2022). The soil and climate conditions at the common garden and tree source provenances (populations) are shown in Table 1, along with the number of trees used in the experiment from each source population. The three source provenances from the SD ecotype have mean annual temperatures (MAT) that are warmer than the common garden planting site (Table 1; ranging from 0.1° to 1.8°C warmer, as measured in 2018). The three source provenances from the MR ecotype have mean annual temperatures (MAT) that are the same as, or cooler than, the common garden planting site (Table 1; ranging from 0° difference to 4.9°C cooler, based on 30 year averages). Cottonwoods from the cooler, Mogollon Rim (MR) ecotype express traits associated with frost and drought protection (e.g., later seasonal leaf flush, smaller xylem vessels, reduced stomatal density; Blasini et al., 2021, 2022). Alternatively, cottonwoods from the warmer, Sonoran Desert (SD) ecotype have traits associated with heat tolerance (e.g. hydraulic efficiency and evaporative cooling; Blasini et al., 2021, 2022), which may reduce tolerance to droughts that are increasing in frequency and severity in the western United States (Moran et al., 2023). Cottonwood cuttings for all treatments were propagated and grown in a soil mixture consisting of peat moss, perlite, and vermiculite in 10.16 cm by 76.20 cm pots to encourage deep root systems with minimal soil microbes. Cuttings were grown in the greenhouse under normal day and night cycles for two years and watered every third day. A few cuttings were screened for the presence of ectomycorrhizas prior to planting and none were observed.

Tamarisk was removed from the tamarisk-invaded areas of the site with bulldozers during the spring and summer of 2017. Planting holes were dug 150 cm deep to plant cottonwoods as close to the water table as possible and preclude the need for irrigation. Trees were planted during November and December 2017. Replicates of trees stratified across ecotype and provenance were planted in random order within blocks and replicated in five blocks for each treatment. Trees were all approximately 150 cm tall (plus 76 cm of roots and potting soil) at the time of planting and were planted approximately 150 cm deep. Paired inoculated and uninoculated blocks within the tamarisk and non-tamarisk areas were separated by a minimum of 10 m to maintain inoculation treatment integrity over the two-year experiment. Cottonwood trees were planted with (inoculated treatments) or without (not inoculated treatments) 120 ml of live soil inoculum. Live soil inoculum was added to the field soil during planting by mixing it with field soil when filling in planting holes, in the uppermost 20 cm of the hole containing fine roots. In conjunction with the deep planting, trees were watered when planted but received little water after planting aside from natural precipitation. In addition to the arid climate of the region, and lack of irrigation provided to the trees in the experiment, the study took place during the worst drought since 800 CE in the southwestern United States (Williams et al., 2022).

Locations of the common garden, and inoculum and tree sources are shown in Figure 1. Climate and soil data for the common garden and all source populations are presented in Table 1. The experimental design with the number of trees in each treatment can be found in Supplementary Table 1.

To create the live soil inoculum, living fine root fragments and rhizosphere soil were collected from Fremont cottonwoods in a healthy, tamarisk-free, Fremont cottonwood population from the SD ecotype (Bullpen; Table 1). This live soil was homogenized and then divided amongst ten 30 cm x 30cm x 30 cm plastic containers, mixed with 30% sterile clay particles. Plants known to associate with arbuscular and ectomycorrhizal fungi were cultivated in these containers in the greenhouse for ten months. Each container of the live soil was planted with arbuscular mycorrhizal hosts (leek [*Allium* sp.], corn [*Zea mays*], and marigolds [*Calendula* sp.]) and ectomycorrhizal hosts

TABLE 1 Climate and soil characteristics for the common garden and plant and inoculum ecotype source locations.

Source Population, River (Code)	Keams Canyon, Little Colorado River (KKH-OPI)	Jack Rabbit, Little Colorado River (JLA-JAK)	Citadel Wash, Little Colorado River (CLF-LCR)	Little Colorado River Common Garden, Little Colorado River (LCR)	Kingman, Willow Creek (KWF-WIL)	Bullpen, Clear Creek (BCE-BUL)	Horseshoe Ranch, Agua Fria (CAF-AUG)
Ecotype	**MR**	**MR**	**MR**	**MR**	**SD**	**SD**	**SD**
Latitude	35.81152	34.96	35.6088	35.71923	35.143	34.53972	34.25671
Longitude	-110.16958	-110.436	–111.31369	-111.3194	-113.54284	-111.6966	-112.06617
Elevation (m)	1982	1510	1349	1331	1210	1254	1061
Annual Precipitation (mm)	213.8	194.3	184.17	174.34	246.52	426.74	408.56
Min. Temperature (°C)	2.1	4.2	7.2	7.2	7.2	7.9	10.6
Max. Temperature (°C)	18.8	22.4	23.3	23.4	24.2	24.3	25.4
MAT (°C)	10.4	13.3	15.3	15.3	15.7	16.1	18.0
MAT Transfer (°C)	4.9	2.0	0	–	-0.4	-0.8	-2.7
Max VPD (hPa)	19.4	25.62	27.77	27.87	28.06	27.76	29.18
Soil Series	Tewa very fine sandy loam	Ives Soils	Jocity-Joraibi-Navajo-Riverwash complex	Jocity-Joraibi-Navajo-Riverwash complex/ Torrifluvents, saline	Arizo-Franconia/ Riverwash complex	Anthony fine sandy loam	Barkerville Cobbly Sandy Loam
Soil Content	59.3% sand, 23.2% silt, 17.5% clay	71% sand, 17% silt, 13% clay	40% sand, 25% silt, 35% clay	40% sand, 25% silt, 35% clay	65.7% sand, 22.8% silt, 11.5% clay	67.5% sand, 21% silt, 2.8% clay	68% sand, 20% silt, 13% clay
pH	7.6	8.2	8.2	8.2	7.9	8.2	7
Electrical Conductivity (dS/m)	1	20	0.7	0.7	1	1	0
World Reference Base (WRB) for Soil Resources/Food and Agriculture Organization of the United Nations (FAO) Soil Classifications	Yl-Luvic Yermosols	Yl-Luvic Yermosols	Yl-Luvic Yermosols	Yl- Luvic Yermosols	Yl-Luvic Yermosols	Yl-Luvic Yermosols	Yl-Luvic Yermosols
Number of Trees	159	157	156	–	158	157	157

1. Climate data represents the annual mean for each location, over 30 years, from 4 km surrounding the coordinates, from 1991 to 2020, based on data from the PRISM Climate Group (2023; https://prism.oregonstate.edu/explorer/bulk.php).
2. Soil data is from United States Department of Agriculture (USDA), Natural Resources Conservation Service (2020).
3. Climate abbreviations are as follows: mean annual temperature (MAT) vapor pressure deficit (VPD). MAT Transfer is the MAT of the planting site minus the MAT of the source population.
4. Ecotype/adaptive trait syndrome is as described in Blasini et al. (2021; 2022).
5. WRB and FAO soil classifications are as described in Food and Agriculture Organizaiton of the United Nations (1992).

(ponderosa pines [*Pinus ponderosa*]) and Fremont cottonwoods. Cottonwoods came from the same ecotypes and populations that were used in the field experiment (which are known to associate with both arbuscular and ectomycorrhizal fungi under these conditions, e.g. Meinhardt and Gehring, 2012). The greenhouse experienced natural light and dark cycles, and the live soil inoculum plantings were watered twice per week. Plants in the live soil inoculum were cut at the base prior to inoculum use. Roots of the plants used to cultivate mycorrhizal fungi within the inoculum during greenhouse cultivation were left in the soil, and the soil was mixed to increase homogeneity of the contents and break-up roots into fragments.

2.3 Survival measurements to address hypotheses in the field experiment

Survival of each tree was measured in the field during July 2018 (year 1) and September 2019 (year 2). Survival was primarily evaluated by examining whether trees had green leaves during the growing season. If trees lacked green leaves during the growing season, tissues at the base of the stem under the surface examined to identify if they were still green and living or dried and dead. In a small area at the base of the stem, the thin surface layer was scraped away with a fingernail to reveal the nature of the tissues underneath.

FIGURE 1
Map of cooler, Mogollon Rim (MR; blue) and warmer, Sonoran Desert (SD; red) ecotype source locations and Little Colorado River common garden (in the cooler, MR ecotype). Inoculum source is indicated with (I).

As a result of tree size, planting and rooting depths, and project constraints that included restoration as a goal, collection of the fine root fragments necessary to confirm mycorrhizal status across treatments was not possible.

2.4 Statistical analyses for survival data from the field experiment

All statistical analyses were conducted in R version 4.0.3 (R Core Team, 2020). Individual trees were used as the independent experimental unit since plants were randomly planted and environmental heterogeneity was relatively low (as in Grady et al., 2011, 2015).

To analyze binomial survival data by treatment variables in conjunction with the random effects of tree source population, we utilized a generalized linear mixed effects model using the glmer() function in the lme4 package (Bates et al., 2015) in R. To establish an overall model for use, a suite of mixed effects models were evaluated. All mixed effects models included year one field survival and year two field survival as the dependent variables, and plant ecotype, and tamarisk and inoculation treatments as the independent variables. Mixed-effects models were evaluated using AIC scores to test whether the random variables (which were not the focus of the experiment) of plant source population and field plot replicate meaningfully improved the model without jeopardizing its statistical power (indicated by the additional warnings noted in Supplementary Table 2). The model using source population as a random effect was selected based on AIC scores for both year 1 and year 2 survival and the ability of the model to perform analyses for both years without additional warnings (Supplementary Table 2). Eight treatments (2 tamarisk x 2 inoculation x 2 ecotype) yielded over fifty specific contrasts that could have been evaluated with post-hoc tests using the emmeans package. However, the objective of the experiment was to evaluate the specific contrasts addressing our hypotheses: 1) the soil legacy left by prior tamarisk invasion would generally reduce plant survival after tamarisk removal; 2) live microbial soil/mycorrhizal restoration would improve the survival of trees in areas with a tamarisk soil legacy and 3) live microbial soil/mycorrhizal restoration would improve the survival of trees in areas without a tamarisk history; 4) assisted migrant trees from the warmer SD ecotype would have survival rates equal to or lower than those for trees from the local MR ecotype. As a result, individual contrasts that addressed the specific hypotheses of interest were evaluated within the overall model, using vectors to specify the four contrasts desired (Muldoon, 2019) in conjunction with the emmeans() function in the emmeans package. The sidak adjustment for multiple comparisons was used to adjust post-hoc results for the number of post-hoc comparisons made (Length, 2020). Hypothesis test results are presented as the emmeans z ratio (z, which can be generally described as the contrast divided by its standard error), resulting p-value (p).

Measures of effect size were obtained utilizing the odds ratios from the emmeans() function in the emmeans package (Length, 2020), specifying type='response' to obtain results that were back-transformed from the logit scale (and can be seen in Supplementary Table 6). Probabilities of surviving and standard errors for graphical depictions are from these model results.

2.5 Greenhouse methods to assess inoculation efficacy

During May through August 2020, we conducted a greenhouse experiment that replicated the field study as closely as possible to confirm the efficacy of mycorrhizal inoculation utilizing the same methods and soils. In the greenhouse experiment, 80 small (~25 cm in length and 7 mm in diameter) Fremont cottonwoods (40 cuttings from the SD ecotype and 40 cuttings from the MR ecotype) were planted in D40 pots. Cottonwoods were planted in a mix of one-half potting soil and one-half live soil from the common garden. Half of the trees were planted in common garden soil from areas with a tamarisk legacy, and half in common garden soil from areas with no tamarisk legacy. The soil inoculum was added to cottonwoods during re-potting with common garden soil/potting soil mix. One half of the cottonwoods from each ecotype in each potting mix treatment received live, and one-half received sterilized, soil inoculum. Cottonwoods in each treatment were stratified across the same ecotypes as in the field experiment. Soil inoculum was sourced from a healthy, tamarisk-free, cottonwood population from the SD ecotype (Agua Fria; Table 1) and maintained in the greenhouse as described in Section 2.3 prior to being used for this experiment. Cottonwoods that died early in the experiment were replanted to maintain a balanced experimental design.

To confirm the efficacy of the inoculation methods used in the field, fine roots from the cottonwoods were collected 125 days after initiating treatments. Roots were cleaned and rinsed to remove soil, placed into plastic bags, and immediately frozen until evaluation. EMF have been shown to be more sensitive to tamarisk presence and assisted migration (Meinhardt and Gehring, 2012; Markovchick et al., 2023), but we measured both EMF and AMF colonization. Colonization was evaluated for trees in all treatments.

Root samples were divided into subsamples for separate evaluation of EMF and AMF colonization. Root tips (living, dead [as evidenced by their shriveled, dried, and darkened nature], and living colonized by an EMF morphotype) were counted under a dissecting microscope using the gridline intersect method (as described in Brundrett et al., 1996). To evaluate EMF colonization across treatments, 7500 gridline intersections were evaluated across 75 trees, and 258 EMF root tips were identified and counted. Separate root subsamples were cleared and stained, and evaluated for AMF hyphae, vesicles, and arbuscules using the gridline intersection method under a compound light microscope for a subsample of 14 trees using the methods of Brundrett et al. (1996). Care was taken to quantify dark septate endophytes (DSE; non-mycorrhizal root fungi that have no specialized exchange structure) separately, as indicated by melanized septate hyphae.

2.6 Statistical analyses for the greenhouse inoculation efficacy test

All statistical analyses were conducted in R version 4.0.3 (R Core Team, 2020). Individual trees were used as the independent experimental unit since plants were randomly planted and environmental heterogeneity was relatively low (as in Grady et al., 2011, 2015).

To investigate the efficacy of inoculation, EMF colonization and AMF colonization from the greenhouse inoculation experiment were assessed using mixed effects linear regression with the lmer() function in the lme4 package (Bates et al., 2015) in R. Inoculation and tamarisk treatment was used as the predictor (independent) variable, population was used as a random variable, and EMF and AMF colonization were dependent variables.

3 Results

3.1 Overall model results from the field experiment

Independent variables (tamarisk, inoculation and ecotype treatments) demonstrated significant effects on the dependent variable (survival) in the overall model. During year one, source ecotype and tamarisk soil legacy treatment were significant main effects ($Z=1.97$ and $p = 0.049$; $Z = -5.12$, $p < 0.001$), with significant interactions for inoculation by tamarisk, source by tamarisk, and inoculation by source by tamarisk ($Z = 1.99$, $p = 0.47$; $Z = -2.41$, $p < 0.05$; $Z = 2.46$, $p < 0.05$). During year two, the inoculation treatment was the significant main effect ($Z = -2.66$, $p < 0.01$), with significant interactions for inoculation by source, and inoculation by tamarisk ($Z = 2.206$, $p < 0.05$; $Z = 2.689$, $p < 0.01$). Overall model results can be found in Supplementary Table 4. Model results for each treatment in each year can be found in Supplementary Table 5. Specific post-hoc contrasts for each hypothesis also yielded significant results, are discussed individually below, and can be found in Supplementary Table 6.

3.2 H1: tamarisk soil legacy negatively impacts cottonwood survival in the field

Hypothesis 1 was supported for year one (Figure 2). Without inoculation treatments, trees planted in tamarisk soil were significantly less likely to survive year one than trees planted in soil without a tamarisk legacy ($z = -6.26$, $p < 0.001$). Survival of trees in tamarisk soil was 8.3% of that for trees in soil without a tamarisk legacy. Year 2 results were not significant (Supplementary Table 6).

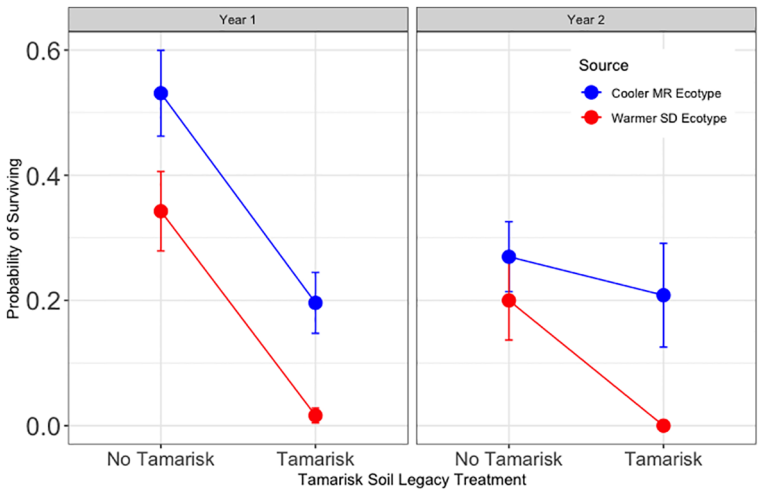


FIGURE 2
Results for Hypothesis 1. Probability of surviving for trees without inoculation sourced from the cooler, local Mogollon Rim (MR; blue) ecotype and from the warmer Sonoran Desert (SD; red) ecotype in each tamarisk legacy treatment. Points represent model means and error bars represent 1 SE. Across ecotypes and source, first year survival for trees in tamarisk soil was 8.3% of that for trees in soil with no tamarisk legacy ($Z = -6.26$, $p<0.001$).

3.3 H2: inoculation with live soil increases field survival of trees in tamarisk legacy soil

Hypothesis 2 was partially supported (Figure 3). For trees in tamarisk soil, year one survival was significantly higher with inoculation than without ($z = -3.05$, and $p < 0.01$). Survival in tamarisk soil if inoculated was 350% of that for trees that were not inoculated. During year two, inoculation results in tamarisk soil depended on tree source ecotype, as visualized in Figure 3, and demonstrated by the significant main effect of inoculation and significant interactions with source ecotype and tamarisk legacy soil in year two seen in the model results ($Z = -2.66$, $p<0.01$; $Z = 2.21$, $p<0.05$; $Z = 2.69$, $p <0.01$ respectively).

3.4 H3: inoculation with live soil increases field survival of trees in soil without a tamarisk legacy

Hypothesis 3 was not supported (Figure 4). For trees in soil without a tamarisk legacy, year one survival was significantly lower with live soil inoculum than without ($z = -3.36$, and $p <0.01$). Year one survival in soil without a tamarisk legacy if provided live inoculum was 50% of that for trees that did not receive inoculum.

For trees in soil without a tamarisk legacy, year two survival was not significantly different overall for cottonwoods that received inoculum and those that did not. (Supplementary Table 6). As can be seen in Figure 4, the overall model results reveal a year two interaction between the main effect of inoculation ($z = -2.66$, $p < 0.01$) and tree source ecotype ($z = 2.21$, $p < 0.05$).

3.5 H4: with inoculation using live soil from the same ecotype, assisted migrants will have survival rates in the field lower than or equal to those for trees from the local ecotype

Hypothesis 4 was supported during year one and not supported during year two (Figure 5). During year one, inoculated trees from the warmer (SD) ecotype had significantly lower survival across tamarisk soil legacy treatments ($z = -3.03$, $p < 0.05$). During year two, survival for inoculated assisted migrant trees was significantly higher for than for those from the local ecotype across tamarisk soil legacy status ($z = 3.16$, and $p < 0.01$). Survival of inoculated trees during year two for assisted migrants was 520% of that for trees from the local ecotype.

3.6 Greenhouse test of inoculation efficacy

Inoculation significantly increased both EMF ($F = 16.26$, $df = 1, 71$, and $p < 0.001$) and AMF ($F=10.20$, $df=1, 9.68$, $p <0.01$) colonization (Figure 6). Averaged across tamarisk treatments and tree source, inoculation increased the number of EMF root tips by nearly 500% ($df=67$, $t=4.06$, $p<0.001$). Averaged across tamarisk treatments and tree source, inoculation nearly doubled AMF colonization ($df=67$, $t=4.06$, $p<0.001$). Means and standard errors can be found in Supplementary Table 3.

4 Discussion

Due to the variety of benefits associated with native mycorrhizal symbioses, reductions and shifts in their communities due to

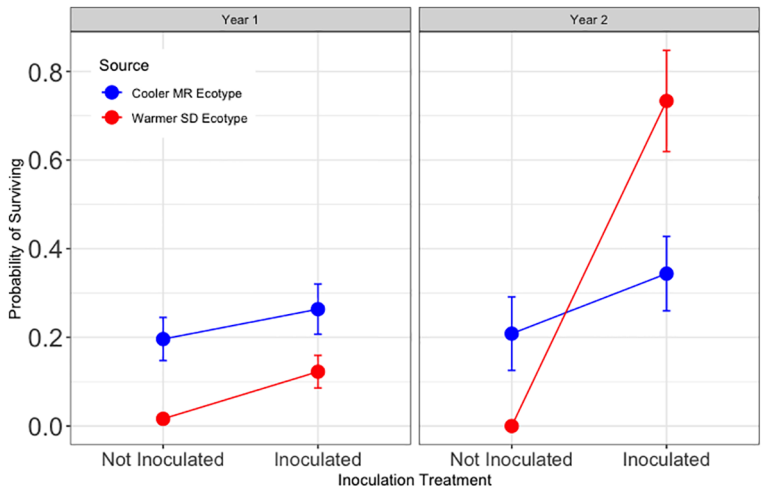


FIGURE 3
Results for Hypothesis 2. First and second year survival (left and right, respectively) for trees in tamarisk soil legacy from the cooler, local Mogollon Rim (MR; blue) ecotype and from the warmer Sonoran Desert (SD; red) ecotype. Points represent model means and error bars represent 1 SE. Inoculation more than tripled year one survival in tamarisk soil ($Z = -3.05$, and $p < 0.01$)). The significant effects of inoculation interacted with source and tamarisk soil legacy treatment during year two ($Z = -2.66$, $p<0.01$; $Z = 2.21$, $p<0.05$; $Z = 2.69$, $p <0.01$ respectively).

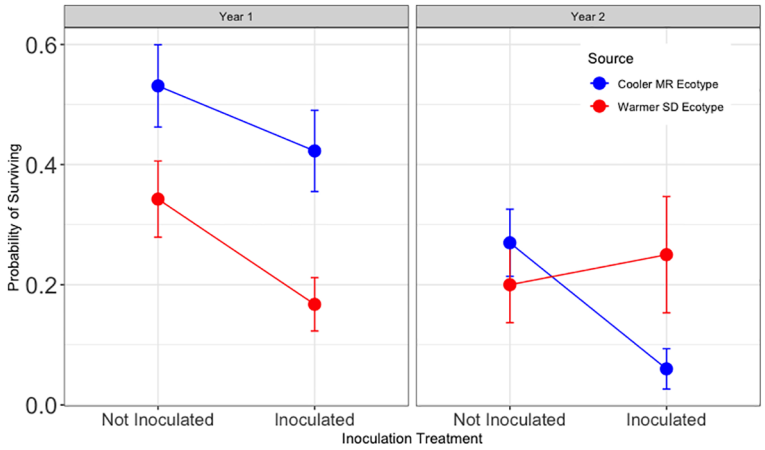


FIGURE 4
Results for Hypothesis 3. First and second year survival for trees in the common garden blocks without a tamarisk legacy. Shades of blue show results for trees sourced from the cooler Mogollon Rim (MR) local ecotype and shades of red show results for assisted trees sourced from the warmer Sonoran Desert (SD) ecotype. Points represent model means and error bars represent 1 SE. In soil without a tamarisk legacy, inoculated trees had one half the survival of non-inoculated trees ($p<0.01$). During the second year, an interaction between inoculation and tree ecotype ($Z = 2.21$, $p < 0.05$) resulted in no significant effect for inoculation overall.

invasive species and other disturbances are concerning, particularly as the stressors on plant communities (that mycorrhizae are known to help mitigate) increase under climate change. For example, mycorrhizal fungi increase plant access to nutrients, mediate responses to stressors, pests, and climate (Babikova et al., 2013; Kivlin et al., 2013; Wilkinson and Dickinson, 1995), improve plant water use efficiency (Querejeta et al., 2003; 2006), and provide access to deep water (Bornyasz et al., 2005; Querejeta et al., 2007). Although more challenging to prove definitively and the focus of a current debate (Karst et al., 2023; Klein et al., 2023), some evidence suggests that certain mycorrhizae may also share water and other resources among plants (Egerton-Warburton et al., 2007; Teste and Simard, 2008; Klein et al., 2016). While few studies have examined whether the eradication of invasive species permits recovery of the mycorrhizal fungal communities and their services, limited data suggests recovery times could be extensive. While fungal biomass began to recover following three years of garlic mustard (*Alliaria petiolota*) removal in eastern deciduous forests, fungal community composition and soil physical and chemical properties remained similar to those in invaded sites (Anthony et al., 2019). Three years after eradication of the invasive plant, fungal community composition, and soil chemical and physical properties, remained comparable to invaded communities. Other studies similarly suggest that disruptions to mycorrhizal fungal communities have

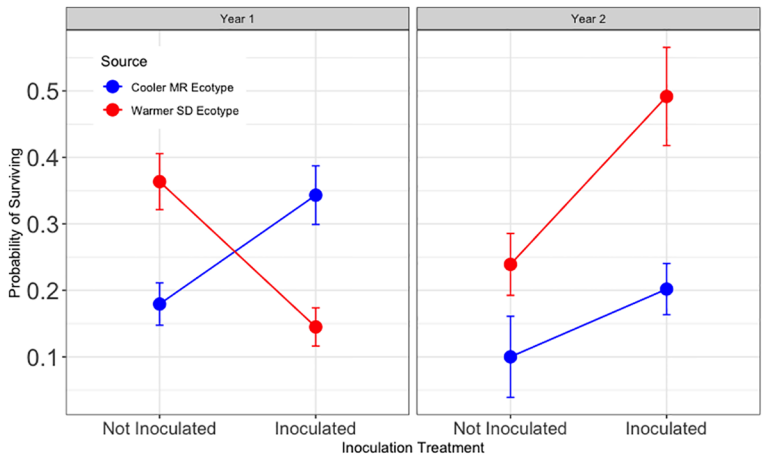


FIGURE 5
Results for Hypothesis 4. First and second year survival for trees in the common garden sourced from the cooler, local Mogollon Rim (MR; blue) ecotype and from the warmer Sonoran Desert (SD; red) ecotype in each inoculation treatment (combined data from Figures 4, 5). Points represent model means and error bars represent 1 SE. Inoculated assisted migrant trees from the warmer ecotype (SD) had a lower chance of survival during year one ($Z = -3.03$, $p = 0.010$) and a higher chance of surviving during the second year than trees from the cooler, local ecotype (MR) across tamarisk soil treatments ($Z = 3.16$, and $p < 0.006$).

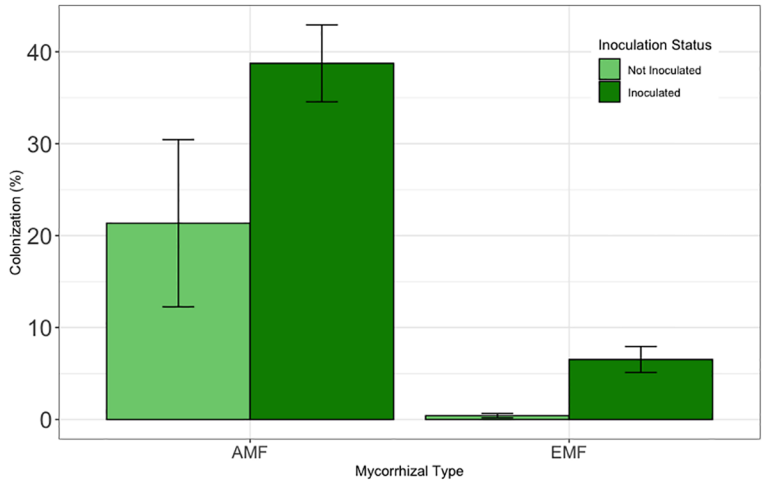


FIGURE 6
Results from greenhouse inoculation efficacy test. To be comparable to the field experiment, field and potting soil were not sterilized prior to inoculation test. Bars are means and error bars represent 1 SE. Inoculation increased EMF colonization by nearly 500% ($p<0.001$) and AMF colonization by nearly 100% ($p<0.05$).

long-term effects, which natural dispersal and planting materials and processes do not successfully mitigate (Sykorova et al., 2007; Southworth et al., 2009; Peay et al., 2010; Pankova et al., 2018).

In this study, the soil legacy left by exotic tamarisk invasion reduced first year survival of native cottonwood trees by over 90%, even after tamarisk trees were removed from the area. In areas with this tamarisk soil legacy, inoculation with native mycorrhizal fungi from native, uninvaded riparian areas significantly improved first-year cottonwood survival across trees from both ecotypes, more than tripling the survival rate. In soils without a tamarisk legacy, the effect of inoculation was more variable, interacting tree source ecotype and time since planting. Although assisted migrant SD trees inoculated with native, living soil inoculum from their home ecotype had a more challenging first year, they survived the second year in extremely harsh conditions at over five times the rate of inoculated local MR trees, revealing the real restoration implications for appropriate and inappropriate soil biome/plant pairings.

4.1 Greenhouse inoculation efficacy

Inoculation with live soil from non-invaded cottonwood riparian areas significantly increased EMF and AMF colonization of cottonwoods compared to non-inoculated controls in the greenhouse. These results provide support for our field experiment methods where it was not possible to sample roots over the two year observation period. Although colonization rates were fairly low, relatively small increases in colonization have been associated with relatively large effects on plant survival in cottonwoods. For example, in another recent intraspecific assisted migration study with Fremont cottonwoods (Markovchick et al., 2023), slight increases in colonization (from 0% to 5%) were associated with 46% of the variation in survival for the tree population with the lowest survival rates.

4.2 The effects of invasive vegetation

Our study provides evidence that soil legacies left by an invasive exotic species such as tamarisk can impact native plants negatively after the removal of the invasive species itself. This is consistent with previous studies showing that invasive species and shift mycorrhizal communities, reduce plant survival and/or biomass, and alter mycorrhizal colonization (Carey et al., 2004; Meinhardt and Gehring, 2012; Wilson et al., 2012). Direct measures of mycorrhizal communities (Pankova et al., 2018; Anthony et al., 2019; Duell et al., 2023) and the success of restoring soil and mycorrhizal communities along with vegetation (Wubs et al., 2016; Koziol and Bever, 2017) provide evidence that the effects of disturbances on mycorrhizal communities can last for years. The negative effects on native plants and on mycorrhizal mutualisms from invasion by exotic plant species have now been documented across multiple systems. Further research is needed to determine the exact limitations and mechanisms operating, and how ecosystem services provided by mycorrhizal fungi and the soil microbiome, such as soil aggregation and water access (e.g. Egerton-Warburton et al., 2007; Duell et al., 2023), are impacted.

4.3 Implications for assisted migration

Two types of assisted migration are generally recognized (Pedlar et al., 2012). First, species rescue assisted migration to avoid extinction of species threatened by climate change, and second, forestry assisted migration to identify populations and genotypes that would achieve higher survival and growth with

climate change than local stock, which is focused on widespread and commercially important species. Keith et al. (2023) identified a third type as community assisted migration, which is focused on the dependent or associated communities of plant species. This third type may be especially important with essential plant mutualists such as pollinators, dispersal agents and mycorrhizal fungi that affect plant performance. Such species have been referred to as the holobiome (usually a multicellular host and its microbial symbionts, as the primary unit of selection (e.g., Gilbert et al., 2012; Bordenstein and Theis, 2015). To the degree that this concept is correct, assisted migration that incorporates the holobiome should achieve higher success than assisted migration strategies that ignore these crucial interacting community members. Our findings showing that inclusion of mycorrhizal fungal mutualists results in much greater plant performance supports this concept.

Though it is challenging to address multiple stressors at once within field studies, it is nonetheless important to understand what results can be expected from many interacting factors. In this study, we have shown that small, incremental steps in assisted migration (Grady et al., 2015), even across ecotype or adaptive trait boundaries (Cooper et al., 2019; Blasini et al., 2021), can be similarly successful to local plant provenances under multiple stressors, if given appropriate mycorrhizal and live soil inoculum. This is consistent with results from Remke et al. (2020) suggesting that inoculation with microbiota from a plant provenance's home environment, while most effective in home soil, can help ameliorate the negative impacts of growing at hotter, drier sites. The results of the current experiment indicate that increasing intraspecific genetic diversity through assisted migration (Etterson et al., 2020; Gomory et al., 2020; Saenz-Romero et al., 2021) with appropriate mycorrhizal inoculation in restoration is a promising strategy for climate change adaptation for some locations. However, more research is needed to determine the most appropriate or beneficial plant/inoculum pairings Allsup et al. (2023) found that inoculating plants with microbiota adapted to climatic characteristics of the planting site enhanced plant survival. In other (e.g. Johnson et al., 2010, 2014) and the current study, inoculum from the same location or (in the current study) ecotype as the plants appears most effective. This suggests that the current study makes an important contribution to our understanding of how to best pair plant microbial source materials for restoration, replanting, and assisted migration. The current study suggests that even within the same plant species, ecosystems, and region, microbial pairings may need to match plant ecotype or physiological adaptations. For example, cottonwoods with physiological adaptations to frost (the local, cooler, MR ecotype; Hultine et al., 2020; Blasini et al., 2021, 2022) did not seem to benefit from microbial inoculations adapted to warmer, SD ecotype conditions (where trees in the same species are adapted to heat; Hultine et al., 2020; Blasini et al., 2021, 2022) unless there was a history of tamarisk.

The long-term effects of assisted migration on interactions between plants and their microbiomes still need to be investigated. For example, what might the effects of assisted migration be on the ability of trees to form common mycorrhizal networks that exchange nutrients, water, and pest signals (e.g. Bingham and Simard, 2011; Babikova et al., 2013; Klein et al., 2016), and how that is affected by including plants and mycorrhizal fungi from multiple provenances, for example? Additionally, given the different physiological adaptations of the two ecotypes, how will differences in their physiological adaptations (such as leaf size, leaf duration, ability to move water for evaporative cooling, etc.), and potentially their microbiomes, alter the albedo of the landscape, the ability of the canopy to provide microclimate buffering, and the carbon cycling of the ecosystem (Michalet et al., 2023)?

This reinforces the merits of considering the interaction of the organism and its microbiome in replanting, restoration, and climate change adaptation (Zilber-Rosenberg and Rosenberg, 2008; Bordenstein and Theis, 2015; Whitham et al., 2020; Allsup et al., 2023), and the urgent need for large-scale assisted migration studies that address inter-species interactions (Bucharova, 2017).

Research on community genetics also indicates that the interactions seen here between native plant ecotypes, soil microbes, and invasive plants could likely have broader implications including for insects, lichens, pathogens, endophytes and other organisms (e.g. Lamit et al., 2015; Keith et al., 2023). In fact, one recent paper (Argüelles-Moyao and Galicia, 2023) suggested that the soil microbiome and fungal species should be primary considerations in assisted migration and the selection of sites for afforestation, since they are as important as temperature and precipitation to outcomes, and regulate key ecosystem processes.

4.4 The impact of time and timing

In addition to interactions with tree ecotype and tamarisk soil legacies, this study found differences in the impact of inoculation treatments between the first and second growing seasons after planting. We are aware of at least two studies demonstrating that the timing of inoculation can impact its initial effectiveness, because establishing the mycorrhizal symbiosis, particularly under otherwise stressful conditions, imposes an initial cost on the plant (Mortimer et al., 2005; Maltz and Treseder, 2015). However, the clear impacts of timing on field survival for trees in non-tamarisk soil in the current study (negative the first year, and interacting with ecotype during the second year) demonstrate the importance of this consideration in a manner that we have not seen in the literature to date. Given the extremely harsh conditions at the experimental site and absence of irrigation or other types of support for new plantings, results suggest that inoculation at stressful times such as field planting might best be accompanied by adequate watering and support for new plantings. Although complicated by multiple interactions, the positive trend for SD trees appropriately paired with soil inoculum during the second year even in non-tamarisk soil legacies suggest that the beneficial effects from appropriately paired inoculum might grow with time, similar to findings from Neuenkamp et al. (2019).

4.5 Interactions between site conditions, and provenances of native plants and mycorrhizal fungi

The widespread success of mycorrhizal restoration is remarkable (Neuenkamp et al., 2019), especially considering

the continued prevalence of mass-produced mycorrhizal inoculums (Hart et al., 2017; Saloman et al., 2022). Given extensive evidence showing that mycorrhizal symbioses are impacted by and co-evolved with plant provenances and site conditions (Johnson et al., 1992, 2010, 2014; Rua et al., 2016), it is unsurprising that in non-tamarisk soil the effects of mycorrhizal inoculation depended on the pairings between the provenance of plants and the mycorrhizal fungi utilized. However, due to the presence of other invasive plant species and the lack of remaining cottonwood trees to provide fungal inoculum at the site, it seemed reasonable to hypothesize that inoculation with a diverse mix of native mycorrhizal fungi native to cottonwood trees in a riparian area in the same state would be broadly beneficial across plant provenances and ecotypes. Interestingly, even taking into consideration the effects of timing, our study did not support this. Further research is needed to reveal how pairing mycorrhizal inoculum sources and plant provenances can optimize results. For example, developing inoculum combinations that include microbiota from high salinity sites could be used to assist paired plant partners at a replanting site with high salinity, but it is not yet apparent whether plants of all intraspecies provenances would benefit equally, or what combination of mycorrhizal inoculation sources would best promote plant provenances of varying sources. As best practices for restoration begin to incorporate mycorrhizal restoration, there is an urgent need for additional research on mismatched verses optimal inoculum-plant pairings to avoid unintended, counterproductive negative impacts from a tool that has so much potential. Additionally, the physiological strategies of plants can vary by ecotype and population (e.g. Blasini et al., 2021; 2022). The outcomes of mycorrhizal symbiosis are known to vary with every change in partner and environment (Klironomos, 2003; Rillig and Mummey, 2006), and specific, limited subsets of mycorrhizal fungi within a community are known to provide certain services (e.g. Egerton-Warburton et al., 2007). Sevanto et al. (2023) recently demonstrated that within the same plant community, the outcomes of a pairing with a single fungal partner can vary based on tree genotype. These findings point to the importance of utilizing live, native soil inoculum, and of refining our understanding of optimal pairings between site conditions, plant source, and fungal inoculum source based on how plant physiological adaptations and fungal services co-vary. In some places, collection of live soil inoculum may not be possible (due to legal, archeological, and/or pathogenic concerns, for example), and/or scaling activities up to produce the needed volumes of live soil inoculum necessary may prove prohibitive. In our study, we collected relatively insignificant volumes of live soil inoculum (approximately two five gallon buckets or 38 liters) and propagated it in the greenhouse using methods to encourage growth of the mycorrhizal community. Such methods may reduce barriers to using live soil inoculum (making the practice more sustainable without requiring large volumes of living soil from natural areas), but could also potentially bias the resulting mycorrhizal community such that its overall makeup differs from exactly what it would be, or what might be optimal for the plants, in the field (e.g. Sykorova et al., 2007; Southworth et al., 2009).

4.6 Conclusions

This study demonstrated a large legacy effect of an invasive species on habitat restoration, and that field survival of assisted migrant plant provenances can be boosted by implementing intentional assisted migration of their soil biota and mycorrhizal fungi. This study provided evidence that appropriate, native soil inoculation could increase the efficacy of assisted migration from warmer areas as a climate change adaptation strategy, and that native soil inoculum should be sourced from the same ecotype as the plants being inoculated. These findings improve our understanding of fundamental ecological concepts about how invasive species and symbioses affect ecosystems and provide restoration best practice targets.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

LM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. AB-A: Investigation, Writing – review & editing. DR: Data curation, Investigation, Writing – review & editing. TD: Data curation, Investigation, Writing – review & editing. KG: Conceptualization, Investigation, Resources, Writing – review & editing. KH: Methodology, Resources, Writing – review & editing. GA: Conceptualization, Resources, Writing – review & editing. TW: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. JQ: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Supervision, Writing – review & editing. CG: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Visualization, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the National Science Foundation Macrosystems Biology program (DEB-1340852, DEB-1340856 and MRI-DBI-1126840), NAU's Presidential Fellowship Program, the Lucking Family Professorship, the ARCS Foundation and Mrs. John Van Denburgh, and the Arizona Mushroom Society's Dr. Chester Leathers Graduate Student Scholarship. The Little Colorado River

common garden is part of the NSF funded Southwest Experimental Garden Array (https://sega.nau.edu/).

Acknowledgments

We thank Babbitt Ranches, the Landsward Foundation, and the Pulliam Foundation, for initial funding and hosting of the Little Colorado River common garden. We gratefully acknowledge everyone who contributed to peer reviews of this paper, especially Nancy C. Johnson, the Gehring lab, the Cottonwood Ecology and Community Genetics Laboratory group, and attendees of the Biennial Conference of Science and Management on the Colorado Plateau and RiversEdge West annual conference for their constructive comments and feedback. We gratefully acknowledge the local Diné and Hopi communities, Blue Dot Education, High Tech High School students, Christopher Updike and many volunteers and field assistants for a variety of contributions that made this research a reality. We gratefully acknowledge WildEarth Guardians for staff and conference time that made discussion and revisions of this publication possible.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2024.1331341/full#supplementary-material

References

Allsup, C. M., George, I., and Lankau, R. A. (2023). Shifting microbial communities can enhance tree tolerance to changing climates. *Science.* 380, 835–840. doi: 10.1126/science.adf2027

Anthony, M. A., Stinson, K. A., Trautwig, A. N., Coates-Connor, E., and Frey, S. D. (2019). Fungal communities do not recover after removing invasive *Alliaria petiolata* (garlic mustard). *Biol. Invasions* 21, 3085–3099. doi: 10.1007/s10530-019-02031-8

Argüelles-Moyao, A., and Galicia, L. (2023). Assisted migration and plant invasion: importance of belowground ecology in conifer forest tree ecosystems. *Can. J. For. Res* 54 (1), 110-121. doi: 10.1139/cjfr-2023-0016

Babikova, Z., Gilbert, L., Bruce, T. J. A., Birkett, M., Caulfield, J. C., Woodcock, C., et al. (2013). Underground signals carried through common mycelial networks warn neighbouring plants of aphid attack. *Ecol. Lett.* 16, 835–843. doi: 10.1111/ele.12115

Bates, D., Machler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Beauchamp, V. B., Stromberg, J. C., and Stutz, J. C. (2005). Interactions between *Tamarix ramosissima* (saltcedar), Populus fremontii (cottonwood), and mycorrhizal fungi: effects on seedling growth and plant species coexistence. *Plant Soil* 275, 219–229. doi: 10.1007/s11104-005-1740-7

Beauchamp, V. B., Stromberg, J. C., and Stutz, J. C. (2006). Arbuscular mycorrhizal fungi associated with *Populus-Salix* stands in a semi-arid riparian ecosystem. *New Phytol.* 170, 369–380. doi: 10.1111/j.1469-8137.2006.01668.x

Bingham, M. A., and Simard, S. W. (2011). Do mycorrhizal network benefits to survival and growth of interior Douglas-fir seedlings increase with soil moisture stress? *Ecol. Evol.* 1, 306–316. doi: 10.1002/ece3.24

Blasini, D. E., Koepke, D. F., Bush, S. E., Allan, G. J., Gehring, C. A., Whitham, T. G., et al. (2022). Tradeoffs between leaf cooling and hydraulic safety in a dominant arid land riparian tree species. *Plant Cell Environ.* 45 (6), 1–18. doi: 10.1111/pce.14292

Blasini, D. E., Koepke, D. F., Grady, K. C., Allan, G. J., Gehring, C. A., Whitham, T. G., et al. (2021). Adaptive trait syndromes along multiple economic spectra define cold and warm adapted ecotypes in a widely distributed foundation tree species. *J. Ecol.* 109, 1298–1318. doi: 10.1111/1365-2745.13557

Bordenstein, S. R., and Theis, K. R. (2015). Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PloS Biol.* 13, e1002226. doi: 10.1371/journal.pbio.1002226

Bornyasz, M. A., Graham, R. C., and Allen, M. F. (2005). Ectomycorrhizae in a soil-weathered granitic bedrock regolith: linking matrix resources to plants. *Geoderma* 126, 141–160. doi: 10.1016/j.geoderma.2004.11.023

Brundrett, M., Bougher, N., Dell, B., Grove, T., and Malajczuk, N. (1996). "Working with mycorrhizas in forestry and agriculture," vol. 32. (Canberra: Australian Centre for International Agricultural Research), 374.

Bucharova, A. (2017). Assisted migration within species range ignores biotic interactions and lacks evidence. *Restor. Ecol.* 25, 14–18. doi: 10.1111/rec.12457

Carey, E. V., Marler, M. J., and Callaway, R. M. (2004). Mycorrhizae transfer carbon from a native grass to an invasive weed: evidence from stable isotopes and physiology. *Plant Ecol.* 172, 133–141.

Coban, O., de Deyn, G. B., and van der Ploeg, M. (2022). Soil microbiota as game-changers in restoration of degraded lands. *Science* 375, eabe0725. doi: 10.1126/science.abe0725

Cooper, H. F., Grady, K. C., Cowan, J. A., Best, R. J., Allan, G. J., and Whitham, T. G. (2019). Genotypic variation in phenological plasticity: reciprocal common gardens reveal adaptive responses to warmer springs but not to fall frost. *Global Change Biol.* 25, 187–200. doi: 10.1111/gcb.14494

Duell, E. B., O'Hare, A., and Wilson, G. W. T. (2023). Inoculation with native soil improves seedling survival and reduces non-native reinvasion in a grassland restoration. *Restor. Ecol.* 31, e13685. doi: 10.1111/rec.13685

Egerton-Warburton, L. M., and Allen, E. B. (2000). Shifts in arbuscular mycorrhizal communities along an anthropogenic nitrogen deposition gradient. *Ecol. Appl.* 10, 484–496. doi: 10.1890/1051-0761(2000)010[0484:SIAMCA]2.0.CO;2

Egerton-Warburton, L. M., Querejeta, J. I., and Allen, M. F. (2007). Common mycorrhizal networks provide a potential pathway for the transfer of hydraulically lifted water between plants. *J. Exp. Bot.* 58, 1473–1483. doi: 10.1093/jxb/erm009

Endresz, G., and Kalapos, T. (2013). Arbuscular mycorrhizal colonization of roots of grass species differing in invasiveness. *Community Ecol.* 14, 67–76. doi: 10.1556/ComEc.14.2013.1.8

Etterson, J. R., Cornett, M. W., White, M. A., and Kavajecz, L. C. (2020). Assisted migration across fixed seed zones detects adaptation lags in two major North American tree species. *Ecol. Appl.* 30, e02092. doi: 10.1002/eap.2092

Fargione, J., Haase, D. L., Burney, O. T., Kildisheva, O. A., Edge, G., Cook-Patton, S. C., et al. (2021). Challenges to the reforestation pipeline in the United States. *Front. Forests Global Change* 4. doi: 10.3389/ffgc.2021.629198

Food and Agriculture Organization of the United Nations (1992). Soil map of the world. Available online at: https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/faounesco-soil-map-of-the-world/en/.

Gilbert, S. F., Sapp, J., and Tauber, A. I. (2012). A symbiotic view of life: We have never been individuals. *Q. Rev. Biol.* 87, 325–341. doi: 10.1086/668166

Gitlin, A. R., Sthultz, C. R., Bowker, M. A., Stumpf, S., Kennedy, K., Muñoz, A., et al. (2006). Mortality gradients within and among dominant plant populations as barometers of ecosystem change during extreme drought. *Conserv. Biol.* 20, 1477–1486. doi: 10.1111/j.1523-1739.2006.00424.x

Gomory, E., Krajmerova, D., Hrivnak, M., and Longauer, R. (2020). Assisted migration vs. close-to-nature forestry: what are the prospects for trees under climate change? *Cent. Eur. Forestry J.* 66, 63–70. doi: 10.2478/forj-2020-0008

Grady, K. C., Ferrier, S. M., Kolb, T. E., Hart, S. C., Allan, G. J., and Whitham, T. G. (2011). Genetic variation in productivity of foundation riparian species at the edge of their distribution: implications for restoration and assisted migration in a warming climate. *Global Change Biol.* 17, 3724–3735. doi: 10.1111/j.1365-2486.2011.02524.x

Grady, K. C., Laughlin, D. C., Ferrier, S. M., Kolb, T. E., Hart, S. C., Allan, G. J., et al. (2013). Conservative leaf economic traits correlate with fast growth of genotypes of a foundation riparian species near the thermal maximum extent of its geographic range. *Funct. Ecol.* 27 (2), 428–438.

Grady, K. C., Kolb, T. E., Ikeda, D. H., and Whitham, T. G. (2015). A bridge too far: cold and pathogen constraints to assisted migration of riparian forests. *Restor. Ecol.* 23, 811–820. doi: 10.1111/rec.12245

Grünfeld, L., Skias, G., Rillig, M. C., and Veresoglou, S. D. (2022). Arbuscular mycorrhizal root colonization depends on the spatial distribution of the host plants. *Mycorrhiza* 32, 387–395. doi: 10.1007/s00572-022-01087-0

Hart, M., Chaudhary, B., Antunes, P. M., and Abbott, L. K. (2017). Fungal inoculants in the field: is the reward greater than the risk? *Funct. Ecol.* 32 (1), 1–10. doi: 10.1111/1365-2435.12976

Hawkes, C. V., Belnap, J., D'Antonio, C., and Firestone, M. K. (2006). Arbuscular mycorrhizal assemblages in native plant roots change in the presence of invasive exotic grasses. *Plant Soil* 218, 369–380. doi: 10.1007/s11104-005-4826-3

Helander, M., Saloniemi, I., Omacini, M., Druille, M., Salminen, J.-P., and Saikkonen, K. (2018). Glyphosate decreases mycorrhizal colonization and affects plant-soil feedback. *Sci. Total Environ.* 642, 285–291. doi: 10.1016/j.scitotenv.2018.05.377

Hultine, K. R., Allan, G. J., Blasini, D. E., Bothwell, H. M., Cadmus, A., Cooper, H. F., et al. (2020). Adaptive capacity in the foundation tree species *Populus fremontii*: implications for resilience to climate change and non-native species invasion in the American Southwest. *Conserv. Physiol.* 8, 1–16. doi: 10.1093/conphys/coaa061

Hultine, K. R., Bean, D. W., Dudley, T. L., and Gehring, C. A. (2015). Species introductions and their cascading impacts on biotic interactions in desert riparian ecosystems. *Integr. Comp. Biol.* 55, 587–601. doi: 10.1093/icb/icv019

Hultine, K. R., Burtch, K. G., and Ehleringer, J. R. (2013). Gender specific patterns of carbon uptake and water use in a dominant riparian tree species exposed to a warming climate. *Global Change Biol.* 19 (11), 3390–3405.

Ikeda, D. H., Max, T. L., Allan, G. J., Lau, M. K., Shuster, S. M., and Whitham, T. G. (2017). Genetically informed ecological niche models improve climate change predictions. *Global Change Biol.* 23, 164–176. doi: 10.1111/gcb.13470

IPBES (2023). *Assessment report on invasive alien species and their control of the intergovernmental science-policy platform on biodiversity and ecosystem services*. Eds. H. E. Roy, A. Pauchard, P. Stoett, T. Renard Truong, S. Bacher, B. S. Galil, P. E. Hulme, T. Ikeda, K. V. Sankaran, M. A. McGeoch, L. A. Meyerson, M. A. Nuñez, A. Ordonez, S. J. Rahlao, E. Schwindt, H. Seebens, A. W. Sheppard and V. Vandvik (Bonn, Germany: IPBES secretariat). doi: 10.5281/zenodo.7430692

Johnson, N. C., Tilman, D., and Wedin, D. (1992). Plant and soil controls on mycorrhizal fungal communities. *Ecology* 73, 2034–2042. doi: 10.2307/1941453

Johnson, N. C., Wilson, G. W. T., Bowker, M. A., Wilson, J. A., and Miller, R. M. (2010). Resource limitation is a driver of local adaptation in mycorrhizal symbioses. *PNAS* 107, 2093–2098. doi: 10.1073/pnas.0906710107

Johnson, N. C., Wilson, G. W. T., Wilson, J. A., Miller, R. M., and Bowker, M. A. (2014). Mycorrhizal phenotypes and the law of the minimum. *New Phytol.* 205, 1473–1484. doi: 10.1111/nph.13172

Jones, M. D., Durall, D. M., and Cairney, J. W. G. (2003). Ectomycorrhizal communities in young forest stands regenerating after clearcut logging. *New Phytol.* 157, 399–422. doi: 10.1046/j.1469-8137.2003.00698.x

Karst, J., Jones, M. D., and Hoeksema, J. D. (2023). Positive citation bias and overinterpreted results lead to misinformation on common mycorrhizal networks in forests. *Nat. Ecol. Evol.* 7 (4), 501–511. doi: 10.1038/s41559-023-01986-1

Keith, A. R., Bailey, J. K., and Whitham., T. G. (2023). Assisted migration experiments along a distance/elevation gradient show limits to supporting home site communities. *PloS Climate* 2, e0000137. doi: 10.1371/journal.pclm.0000137

Kivlin, S. N., Emery, S. M., and Rudgers, J. A. (2013). Fungal symbionts alter plant responses to global climate change. *Am. J. Bot.* 100, 1445–1457. doi: 10.3732/ajb.1200558

Klein, T., Rog, I., Livne-Luzon, S., and van der Heijden, M. G. A. (2023). Belowground carbon transfer across mycorrhizal networks among trees: Facts, not fantasy. *Open Res. Europe* 168, 1–16. doi: 10.12688/openreseurope.16594.1

Klein, T., Siegwolf, R. T. W., and Körner, C. (2016). Belowground carbon trade among tall trees in a temperate forest. *Science* 352, 342–344. doi: 10.1126/science.aad6188

Klironomos, J. N. (2003). Variations in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology* 84, 2292–2301. doi: 10.1890/02-0413

Koziol, L., and Bever, J. D. (2017). AMF, phylogeny, and succession: specificity of response to mycorrhizal fungi increases for late-successional plants. *Ecosphere* 7, e01555. doi: 10.1111/1365-2664.12843

Lamit, L. J., Busby, P. E., Lau, M. K., Compson, Z. G., Wojtowicz, T., Keith, A. R., et al. (2015). Tree genotype mediates covariance among communities from microbes to lichens and arthropods. *J. Ecol.* 103, 840–850. doi: 10.1111/1365-2745.12416

Length, R. V. (2020). emmeans: estimated marginal means, aka least-squares means. Available online at: https://CRAN.R-project.org/package=emmeans.

León-Sánchez, L., Nicolás, E., Goberna, M., Prieto, I., Maestre, F. T., and Querejeta, J. I. (2018). Poor plant performance under simulated climate change is linked to mycorrhizal responses in a semi-arid shrubland. *J. Ecol.* 106, 960–976. doi: 10.1111/1365-2745.12888

Lesica, P., and DeLuca, T. H. (2004). Is tamarisk allelopathic? *Plant Soil* 267, 357–365. doi: 10.1007/s11104-005-0153-y

Maltz, M. R., and Treseder, K. K. (2015). Sources of inocula influence mycorrhizal colonization of plants in restoration projects: a meta-analysis. *Restor. Ecol.* 23, 625–634. doi: 10.1111/rec.12231

Markovchick, L. M., Schaefer, E. A., Deringer, T., Kovacs, Z. I., Deckert, R. J., Yazzie, J., et al. (2023). Postrestoration colonization suggests slow regeneration, plant translocation barriers, and other host/symbiont lessons during the United Nations' Decade on Ecosystem Restoration. *Restor. Ecol.* 31, e13940. doi: 10.1111/rec.13940

Meinhardt, K. A., and Gehring, C. A. (2012). Disrupting mycorrhizal mutualisms: a potential mechanism by which exotic tamarisk outcompetes native cottonwoods. *Ecol. Appl.* 22, 532–549. doi: 10.1890/11-1247.1

Michalet, R., Carcaillet, C., Delerue, F., Domec, J. C., and Lenoir, J. (2023). Assisted migration in a warmer and drier climate: less climate buffering capacity, less facilitation and more fires at temperate latitudes? *Oikos*, e10248. doi: 10.1111/oik.10248

Moran, M. E., Aparecido, L. M. T., Koepke, D. F., Cooper, H. F., Doughty, C. E., Gehring, C. A., et al. (2023). Limits of thermal and hydrological tolerance in a foundation tree species (*Populus fremontii*) in the desert southwestern United States. *New Phytol* 240 (6), 2298-2311. doi: 10.1111/nph.19247

Mortimer, P. E., Archer, E., and Valentine, A. J. (2005). Mycorrhizal C costs and nutritional benefits in developing grapevines. *Mycorrhiza* 15, 159–165. doi: 10.1007/s00572-004-0317-2

Muldoon, A. (2019). Custom contrasts in emmeans. Available online at: https://www.r-bloggers.com/2019/04/custom-contrasts-in-emmeans/ (Accessed October 25, 2023).

National Academies of Sciences, Engineering, and Medicine (2020). *An Assessment of the need for native seeds and the capacity for their supply: interim report* (Washington, DC: The National Academies Press). doi: 10.17226/25859

Neuenkamp, L., Prober, S. M., Price, J. N., Zobel, M., and Standish, R. J. (2019). Benefits of mycorrhizal inoculation to ecological restoration depend on plant functional type, restoration context, and time. *Fungal Ecol.* 40, 140–149. doi: 10.1016/j.funeco.2018.05.004

Pankova, H., Dostalek, T., Vazacova, K., and Munzbergova, Z. (2018). Slow recovery of mycorrhizal fungi and plant community after fungicide application: an eight year experiment. *J. Vegetation Sci.* 29, 695–703. doi: 10.1111/jvs.12656

Parks, S. A., and Abatzoglou, J. T. (2020). Warmer and drier fire seasons contribute to increases in area burned at high severity in Western US forests from 1985 to 2017. *Geophysical Res. Lett.* 47, 1–10. doi: 10.1029/2020GL089858

Parks, S. A., Holsinger, L. M., Blankenship, K., Dillon, G. K., Goeking, S. A., and Swaty, R. (2023). Contemporary wildfires are more severe compared to the historical reference period in western US dry conifer forests. *For. Ecol. Manage.* 544, 121232. doi: 10.1016/j.foreco.2023.121232

Peay, K. G., Garbelotto, M., and Bruns, T. D. (2010). Evidence of dispersal limitation in soil microorganisms: isolation reduces species richness on mycorrhizal tree islands. *Ecology* 91, 3631–3640. doi: 10.1890/09-2237.1

Pedlar, J. H., McKenney, D. W., Aubin, I., Beardmore, T., Beaulieu, J., Iverson, L., et al. (2012). Placing forestry in the assisted migration debate. *BioScience* 62, 835–842. doi: 10.1525/bio.2012.62.9.10

Poff, B., Koestner, K. A., Neary, D. G., and Merritt, D. (2012). *Threats to western United States riparian ecosystems: a bibliography. General Technical Report RMRS-GTR-269* (Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station). Available at: https://www.fs.usda.gov/rm/pubs/rmrs_gtr269.pdf. doi: 10.2737/RMRS-GTR-269

PRISM Climate Group, Oregon State University (2023). PRISM climate data. Available online at: https://prism.oregonstate.edu/explorer/bulk.php (Accessed October 23, 2023).

Querejeta, J. I., Allen, M. F., Caravaca, F., and Roldan, A. (2006). Differential modulation of host plant δ13C and δ18O by native and nonnative arbuscular mycorrhizal fungi in a semiarid environment. *New Phyt* 169, 379–387. doi: 10.1111/j.1469-8137.2005.01599.x

Querejeta, J. I., Barea, J. M., Allen, M. F., Caravaca, F., and Roldan, A. (2003). Differential response of d13C and water use efficiency to arbuscular mycorrhizal infection in two aridland woody plant species. *Oecologia* 135, 510–515. doi: 10.1007/s00442-003-1209-4

Querejeta, J. I., Egerton-Warburton, L. M., and Allen, M. F. (2007). Hydraulic lift may buffer rhizosphere hyphae against the negative effects of severe soil drying in a California oak savanna. *Soil Biol. Biogeochemistry* 39, 409–417. doi: 10.1016/j.soilbio.2006.08.008

Querejeta, J. I., Schlaeppi, K., López-García, A., Ondoño, S., Prieto, I., van der Heijden, M. G. A., et al. (2021). Lower relative abundance of ectomycorrhizal fungi under a warmer and drier climate is linked to enhanced soil organic matter decomposition. *New Phytol.* 232, 1399–1413. doi: 10.1111/nph.17661

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

Remke, M. J., Hoang, T., Kolb, T., Gehring, C. A., Johnson, N. J., and Bowker, M. A. (2020). Familiar soil conditions help *Pinus ponderosa* seedlings cope with warming and drying climate. *Restor. Ecol.* 28, S344–S354. doi: 10.1111/rec.13144

Rillig, M. C., and Mummey, D. L. (2006). Mycorrhizae and soil structure. *New Phytol.* 171, 41–53. doi: 10.1111/j.1469-8137.2006.01750.x

Rua, M. A., Antoninka, A., Antunes, P. M., Chaudhary, V. B., Gehring, C. A., Lamit, L. J., et al. (2016). Home-field advantage? Evidence of local adaptation among plants, soil, and arbuscular mycorrhizal fungi through meta-analysis. *BMC Evolutionary Biol.* 16, 122. doi: 10.1186/s12862-016-0698-9

Saenz-Romero, C., O'Neill, G., Aitken, S. N., and Lindig-Cisneros, R. (2021). Assisted migration field tests in Canada and Mexico: lessons, limitations, and challenges. *Forests* 12, 9. doi: 10.3390/f12010009

Saloman, M. J., Demarmels, R., Watts-Williams, S. J., McLaughlin, M. J., Kafle, A., Ketelsen, C., et al. (2022). Global evaluation of commercial arbuscular mycorrhizal inoculants under greenhouse and field conditions. *Appl. Soil Ecol.* 169, 104225. doi: 10.1016/j.apsoil.2021.104225

Sevanto, S., Gehring, C. A., Ryan, M. G., Patterson, A., Losko, A. S., Vogel, S. C., et al. (2023). Benefits of symbiotic ectomycorrhizal fungi to plant water relations depend on plant genotype in pinyon pine. *Sci. Rep.* 13, 14424. doi: 10.1038/s41598-023-41191-5

Simler, A. B., Williamson, M. A., Schwartz, M. W., and Rizzo, D. M. (2019). Amplifying plant disease risk through assisted migration. *Conserv. Lett.* 12, e12605. doi: 10.1111/conl.12605

Simonson, W. D., Miller, E., Jones, A., García-Rangel, S., Thornton, H., and McOwen, C. (2021). Enhancing climate change resilience of ecological restoration — A framework for action. *Perspect. Ecol. Conserv.* 19, 300–310. doi: 10.1016/j.pecon.2021.05.002

Southworth, D., Carrington, E. M., Frank, J. L., Gould, P., Harrington, C. A., and Devine, W. D. (2009). Mycorrhizas on nursery and field seedlings of *Quercus garryana*. *Mycorrhiza* 19, 149–158. doi: 10.1007/s00572-008-0222-1

Sykorova, Z., Ineichen, K., Weimken, A., and Redecker, D. (2007). The cultivation bias: different communities of arbuscular mycorrhizal fungi detected in roots from the field, from bait plants transplanted to the field, and from a greenhouse trap experiment. *Mycorrhiza* 18, 1–14. doi: 10.1007/s00572-007-0147-0

Teste, F. P., and Simard, S. W. (2008). Mycorrhizal networks and distance from mature trees alter patterns of competition and facilitation in dry Douglas-fir forests. *Oecologia* 158, 193–203. doi: 10.1007/s00442-008-1136-5

Tiscar, P. A., Lucas-Borja, M. E., and Candel-Perez, D. (2018). Lack of local adaptation to the establishment conditions limits assisted migration to adapt drought-prone *Pinus nigra* to climate change. *For. Ecol. Manage.* 409, 719–728. doi: 10.1016/j.foreco.2017.12.014

United States Department of Agriculture (USDA) and Natural Resources Conservation Service (2020). *National soil survey handbook* (Accessed March 22, 2022 and February 25, 2024).

Wheeler, N. C., Steiner, K. C., Schlarbaum, S. E., and Neale, D. B. (2015). The evolution of forest genetics and tree improvement research in the United States. *J. Forestry* 113, 500–510. doi: 10.5849/jof.14-120

Whitham, T. G., Gehring, C. A., Bothwell, H. M., Cooper, H. F., Hull, J. B., Allan, G. J., et al. (2020). "Using the Southwest Experimental Garden Array to enhance riparian restoration in response to global change: identifying and deploying genotypes and for current and future environments (Chapter 4)," in *Riparian research and management: past, present, future*, vol. 2 . Eds. S. W. Carothers, R. R. Johnson, D. M. Finch, K. J. Kingsley and R. H. Hamre (Fort Collins, CO, U.S.: Department of Agriculture, Forest Service, Rocky Mountain Research Station), 63–79. Available at: https://www.fs.usda.gov/treesearch/pubs/60539.

Wilkinson, D. M., and Dickinson, N. M. (1995). Metal resistance in trees: the role of mycorrhizae. *Oikos* 72, 298–300. doi: 10.2307/3546233

Williams, A. P., Cook, B. I., and Smerdon, J. E. (2022). Rapid intensification of the emerging southwestern North American megadrought in 2020-2021. *Nat. Climate Change* 12, 232–234. doi: 10.1038/s41558-022-01290-z

Wilson, G. W. T., Hickman, K. R., and Williamson, M. W. (2012). Invasive warm-season grasses reduce mycorrhizal root colonization and biomass production of native prairie grasses. *Mycorrhiza* 22, 327–336. doi: 10.1007/s00572-011-0407-x

Wolfe, B. E., Rodgers, V. L., Stinson, K. A., and Pringle, A. (2008). The invasive plant *Alliaria petiolata* (garlic mustard) inhibits ectomycorrhizal fungi in its introduced range. *J. Ecol.* 96, 777–783. doi: 10.1111/j.1365-2745.2008.01389.x

Wubs, E. R. J., van der Putten, W. H., Bosch, M., and Bezemer, T. M. (2016). Soil inoculation steers restoration of terrestrial ecosystems. *Nat. Plants* 2, 16107. doi: 10.1038/nplants.2016.107

Zaimes, G. N. (2007). "Chapter 7: Human alterations to riparian areas," in *Understanding Arizona's riparian areas*. Ed. G. N. Zaimes (Tucson, AZ, USA: The University of Arizona, Arizona Cooperative Extension, College of Agriculture and Life Sciences), 83–109. Available at: http://arizona.openrepository.com/arizona/bitstream/10150/146921/1/az1432-2007.pdf.

Zhang, Q., Yang, R., Tang, J., Yang, H., Hu, S., and Chen, X. (2010). Positive feedback between mycorrhizal fungi and plants influences plant invasion success and resistance to invasion. *PloS One* 5, 1–10. doi: 10.1371/journal.pone.0012380

Zilber-Rosenberg, I., and Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol. Rev.* 32, 723–735. doi: 10.1111/j.1574-6976.2008.00123.x

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 26 March 2024
DOI 10.3389/frmbi.2023.1229723

OPEN ACCESS

EDITED BY
Rebecca C. Mueller,
Agricultural Research Service (USDA),
United States

REVIEWED BY
Yuanqiang Zou,
Beijing Genomics Institute (BGI), China
Shenghui Li,
Puensum Genetech, China

*CORRESPONDENCE
Senlin Zheng
zhengsenlin@msn.com

†These authors have contributed equally to this work

RECEIVED 26 May 2023
ACCEPTED 30 November 2023
PUBLISHED 26 March 2024

CITATION
Zheng S, Chen H, Yang H, Zheng X, Fu T, Qiu X and Wang M (2024) Differential enrichment of bacteria and phages in the vaginal microbiomes in PCOS and obesity: shotgun sequencing analysis.
Front. Microbiomes 2:1229723.
doi: 10.3389/frmbi.2023.1229723

Differential enrichment of bacteria and phages in the vaginal microbiomes in PCOS and obesity: shotgun sequencing analysis

Senlin Zheng[1*†], Huimin Chen[2], Hongyi Yang[2†], Xulan Zheng[3], Tengwei Fu[4], Xiaoyan Qiu[4] and Meiqin Wang[2]

[1]Third Institute of Oceanography, Ministry of Natural Resources, Xiamen, China, [2]Department of Gynecology and Obstetrics, The First Affiliated Hospital of Xiamen University, Xiamen, China, [3]Department of Psychology, University of British Columbia, Vancouver, BC, Canada, [4]Yanxuan Biotech Ltd., Xiamen, China

Introduction: Previous research has linked vaginal bacteria to polycystic ovary syndrome (PCOS) and obesity in women, yet the specific disparities in vaginal microbiota between these conditions remain unclear.

Methods: In this study, we aimed to elucidate the contribution of dysregulated vaginal microbiota to PCOS and obesity by analyzing the vaginal microbiota in reproductive-aged women with and without PCOS, as well as obese and non-obese women, using shotgun sequencing.

Results: Swab specimens were collected from four groups of subjects: PCOS and obese, PCOS and non-obese, non-PCOS and obese, and non-PCOS and non-obese. A total of 333 bacteria and 24 viruses/phages were identified to the species level. Clustering analysis revealed that non-PCOS and non-obese individuals exhibit a similar "healthy" vaginal microbiome, while both obesity and PCOS were associated with microbial dysbiosis. Significant differences in abundance were observed for 26 bacterial species and 6 phages/viruses between groups. Notably, pathobionts such as *Streptococcus pyogenes*, *Leptospira santarosai*, *Citrobacter amalonaticus*, *Listeria ivanovii*, and *Clostridium perfringens* were significantly less abundant or absent in the non-PCOS and non-obese group. Furthermore, the abundance of *Lactobacillus*, *Pseudomonas* bacteria, and their corresponding phages exhibited positive correlations. *Lactobacillus* bacteria, *lactobacillus* phage, and *pseudomonas* phage/virus were identified as indicators of a healthy vaginal microbiome. Importantly, the differentially enriched bacteria in the PCOS and obesity groups were distinct.

Discussion: This study confirms that PCOS and obesity are associated with differing enrichment of bacteria and viruses/phages, with both conditions linked to microbial dysbiosis. Moreover, our findings suggest that vaginal phage diversity is associated with a healthy vaginal microbiota, while dysbiosis is associated with a decrease in phages alongside increased bacterial diversity.

KEYWORDS

vaginal microbiome, polycystic ovary syndrome, obesity, bacteria, phages

1 Introduction

Polycystic ovary syndrome (PCOS) is an endocrine disorder affecting women of reproductive age with a worldwide prevalence of approximately 8–13% (Jobira et al., 2020). In patients with PCOS, numerous small cysts (fluid-filled sacs) form in the ovaries. Typical clinical features include abnormally high levels of androgens, irregular menstrual cycles, hirsutism, acne, obesity, and infertility (Jobira et al., 2020).

PCOS pathogenesis remains elusive, but emerging research highlights the role of human microbiome in its occurrence and progression (Duan et al., 2021). Studies have shown that PCOS is associated with dysbiosis of the gut microbiota (Lindheim et al., 2017; Liu et al., 2017; Insenser et al., 2018; Torres et al., 2018; Qi et al., 2019; Jobira et al., 2020; Sun et al., 2023). Patients with PCOS have reduced diversity and altered composition of the gut microbiota, such as a decrease in *Lactobacillus* and *Bifidobacterium*, and an impaired intestinal mucosal barrier, compared to those without any health problems. The alterations in the gut microbiota have been linked to levels of inflammation and insulin resistance, by altering the stability of the intestinal mucosa and metabolites in patients with PCOS (Sun et al., 2023).

There are four pathogenic mechanisms of PCOS: hyperandrogenism (HY), insulin resistance (IR), folliculogenesis dysfunction (FC), and neuroendocrine axis dysfunction (ND) (Hao et al., 2016; Crespo et al., 2018; Torres et al., 2018; He and Li, 2020). Hyperandrogenism is a key contributor to PCOS pathogenesis (Witchel et al., 2022). Previous studies have shown that hyperandrogenism is associated with the gut microbiota (Torres et al., 2019). Androgen affects the gut microbiota, including a decrease in the diversity of the gut microbiota and an enrichment of some bacteria. For example, dihydrotestosterone induces an increase in *Chlamydia* but a decrease in *Escherichia coli* in PCOS mice induced by dihydrotestosterone (Zhang et al., 2019); female rats exposed to high levels of androgens at birth have reduced gut microbiota diversity and an increased risk of metabolic disorders in their adult offspring (Sun et al., 2023). On the other hand, gut microbiota affect testosterone secretion (Torres et al., 2018); for example, transplantation of gut microbiota from male to female mice resulted in abnormal basal metabolism and elevated testosterone levels in the latter (Zhang et al., 2019; Chu et al., 2022). Insulin resistance is thought to be the key metabolic abnormality in patients with PCOS (Sun et al., 2023). Dysfunction of the gut microbiota may lead to a "leaky gut", leading to increased presence of LPS in the blood and resulting in chronic inflammation (Hao et al., 2016; He and Li, 2020; Zeng et al., 2020). The gut bacteria-induced inflammation, in turn, fosters insulin resistance, obesity, autoimmune diseases, and PCOS in women of reproductive age (Lindheim et al., 2017; Liu et al., 2017; Insenser et al., 2018). It's noteworthy that insulin can directly stimulate androgen production, hence, hyperinsulinemia in PCOS can concurrently provoke hyperandrogenism (Crespo et al., 2018). Folliculogenesis dysfunction is indirectly linked with gut microbiota dysbiosis. Dysfunction of gut microbiota could lead to hyperandrogenism and insulin resistance, which contribute to the development of PCOS. Excess androgen can block the pathway that converts androgen to estrogens by inhibiting the FSH-stimulated aromatase activity in granulosa cells. Low estrogen levels result in follicular atresia (Zeng et al., 2020). In addition, patients with PCOS have elevated levels of anti-Müllerian hormone (AMH), leading to increased follicular resistance to FSH action. This resistance to FSH action inhibits follicle maturation and ovulation. This leads to the formation of large numbers of immature follicles in the ovary (Crespo et al., 2018; Zeng et al., 2020). Neuroendocrine axis dysfunction (ND) mechanisms contributed to PCOS in some patients. In some PCOS patients, there is resistance to the negative feedback loop of progesterone and GnRH. As the inhibition of progesterone on GnRH production is weakened, more GnRH release from the hypothalamus will lead to excessive LH release from the pituitary and ultimately affect sex hormone production in the ovary and lead to PCOS (Crespo et al., 2018).

The microbiota of the skin and vaginal tract also play a significant role in influencing immune response and inflammation (Huttenhower et al., 2012). In women of reproductive age, a healthy vaginal microbiome typically exhibits a low pH (4.2 to 5.0) in different ethnic groups, with a limited presence of dominant facultative anaerobic bacteria, such as *Lactobacillus crispatus* and *L. inner*, which help maintain an acidic environment and produce H_2O_2 as a defense against

pathogen invasion (Greenbaum et al., 2019; Hong et al., 2020; Wright, 2021). H_2O_2-producing *Lactobacilli* colonize the vagina at rates of up to 10^6–10^7 cells/mL.

A earlier study indicated that the diversity of vaginal bacterial is more pronounced in PCOS patients than in the control group. Notably, PCOS patients demonstrated a lower prevalence of species such as *Lactobacillus crispatus, Mycoplasma, and Prevotella* (Hong et al., 2020). However, it is important to note that the authors also highlighted the necessity for further research to explore the causal relationships between PCOS and the composition of the vaginal microbiome. Dysbiosis of vaginal microbiome is associated with the change of vaginal mucosa (Chen et al., 2021; Miko and Barakonyi, 2023). The vaginal mucosa is a critical component of the female reproductive system. This mucosal lining is a complex structure that serves several essential functions, including providing a barrier against infections, facilitating reproductive processes, and contributing to the overall health of the vaginal environment. The vaginal mucosa comprises various cell types and components, including epithelial cells, macrophages, natural killer (NK) cells, dendritic cells (DC), and neutrophils, T lymphocytes, B cells and immunoglobulin IgG and IgA (Chen et al., 2021; Miko and Barakonyi, 2023). The change of vaginal microbiome will not only result a shift of the microbe components, but also linked with the immune responses of the host.

In addition to bacteria, phages are important components of the human microbiome, and they are thought to play a crucial role in shaping microbial composition, driving bacterial diversity and facilitating horizontal gene transfer. Research on the role of phages in the human microbiome and disease remains relatively limited (Zhang et al., 2023). Major studies have focused on the prevalence of phages in the gut of individuals affected by diseases such as ulcerative colitis and diabetes, with a particular focus on exploring the potential relevance of phages to disease. These investigations have shown a correlation between higher numbers of phages in the gut of individuals with ulcerative colitis and diabetes. However, to date, there is no definitive link between the widespread presence of phages in the human body and the dynamics of human health and disease. Notably, phages are distributed in various parts of the human body, and reports on viruses/phages in the vaginal microbiome are still being reported. Further studies are needed to further understand whether there are any significant changes in phage/virus in the vaginal bacteria of PCOS patients and their associated roles (Łusiak-Szelachowska et al., 2020).

The vaginal microbiome has a direct connection to a woman's reproductive system, which can both influence and be influenced by ovarian health. To date, no study of the vaginal microbiome has used shotgun sequencing to examine the more detailed microorganisms (e.g., phages) in addition to bacteria. The entire microbiome should and does play an important role in the regulation of the vaginal mucosa in patients with PCOS and obesity. Therefore, we performed deeper shotgun sequencing than the 16s test to obtain more specific microbiome information and to explore disease-related factors.

2 Materials and methods

2.1 Participants

According to the inclusion criteria, PCOS patients admitted to the First Affiliated Hospital of Xiamen University in Fujian Province from June 2018 to June 2019, and healthy volunteers were recruited. Meanwhile, healthy volunteers were recruited for the non-PCOS group. All participants signed an informed consent form before the experiment. All participants underwent a general physical examination and an ultrasonic-B scan before enrollment (Supplementary Table 1).

2.2 Diagnostic criteria

All patients with PCOS fulfilled the Rotterdam 2003 diagnostic criteria:

(1) Sparse ovulation or anovulation.
(2) Hyperandrogenic manifestations or hyperandrogenemia.
(3) Ovarian polycystic changes (12 or more follicles 2 ~ 9mm and ovarian volume ≥10mL in one or both ovaries).
(4) Patients meeting 2 of the above 3 criteria but with hyperandrogenism caused by other diseases (congenital adrenal hyperplasia, Cushing's syndrome, ovarian or adrenal tumor) were excluded.
(5) Individuals with a BMI > 28 were considered obese.

2.3 Inclusion criteria

(1) All participants with PCOS were women aged 18–40 years who fulfilled the Rotterdam diagnostic criteria.
(2) The control group consisted of 18–40 year old women with regular menstruation and no clinical signs of hyperandrogenism.

2.4 Exclusion criteria

(1) Adolescents who did not have regular menstrual periods.
(2) Individuals with inflammatory bowel disease.
(3) Individuals with thyroid dysfunction, hyperlactatemia, congenital adrenal hyperplasia, Cushing's syndrome, ovarian or adrenal tumors, and other diseases of the endocrine system.
(4) Patients with other chronic diseases, such as tumors, may require chemotherapy, radiation therapy, or long-term oral medications.

(5) Oral or intravenous use of antibiotics or probiotics within the last three months and during the study.

(6) Vegetarian or irregular eater.

(7) Smokers.

(8) Circumstances that the investigator deems unsuitable for participation in the study.

2.5 Sample collection and shotgun library preparation

According to BMI and the PCOS diagnosis results, the 41 volunteers were divided into four groups, 11 individuals in the PCOS with obesity (PO) group, eight individuals in the non-PCOS with obesity (NO) group, 8 individuals in the PCOS and non-obese (PS) group, and 14 individuals in the non-PCOS and non-obese (NS) control group. Microbiota samples were collected from the posterior fornix of the vagina were collected using a clean and sterile cotton swab during the non-menstrual period (Greenbaum et al., 2019). Swabs were placed into collection tubes containing fixative (YXA, Xiamen, China) and then delivered to the laboratory in dry ice boxes and stored in a −80 °C refrigerator until nucleic acid extraction. Total DNA from the swaps was then extracted using the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany), as recommended by the Human Microbiome Project. The DNA libraries for NGS sequencing were prepared using the Tn5 tagmentation-based protocol (Hennig et al., 2018). DNA segmentation and addition adaptors simultaneously with the Tn5 enzyme were performed according to Hennig's protocol (Hennig et al., 2018). The purified DNA libraries were then sequenced on the Illumina HiSeq X sequencing platform using a pair-end (PE150) strategy.

2.6 Analysis of sequencing data and microbiota profiling

The raw shotgun sequencing data were used for taxonomic analysis using MetaPhlAn3 (Metagenomic Phylogenetic Analysis, Version 3.1.0) (Beghini et al., 2021) and Kraken2 (Wood et al., 2019). The retrieved microbiota analyzed included bacteria, viruses, and fungi. The function of the microbiota was analyzed using Humann3 (Beghini et al., 2021), followed by megahit assembly (Dinghua et al., 2015).

2.7 Statistical methods

The one-way ANOVA test was used to test the significance of the difference in abundance of specific species between groups. The correlation between the bacteria and phage/virus species was analyzed using the R package Hmisc and plotted using the R package pheatmap. The differentially enriched microbial species in the PCOS and obesity groups were analyzed using LEfSe tool (Segata et al., 2011). The significance of the difference in relative bacterial abundance between PCOS and non-PCOS was analyzed with a t-test. The PCA analysis of the vaginal microbiome in samples from different groups was performed using the R package DESeq2 and plotted using ggplot2.

3 Results

3.1 NGS sequencing results

A total of 41 PE150 sequencing data sets were obtained, and each sample received at least 20 million reads. The data have been uploaded to Sequence Read Archive (SRA) database (Accession No. SRR24561304 to SRR24561344).

3.2 Microbiota profiling

3.2.1 Bacteria

The microbiome profile shows 4 phyla (*Actinobacteria, Firmicutes, Proteobacteria*, and *Cyanobacteria*), 4 classes, 9 orders, 11 families, 265 genera, and 333 species of bacteria. There are 220 bacterial species that were observed in only one of the 41 samples, and 33 bacterial species were identified in at least 10 samples. There are 31 genera detected in more than 10 samples, including *Staphylococcus, Lactobacillus, Clostridium, Klebsiella, Xanthomonas, Komagataeibacter, Enterobacter, Pseudomonas, Burkholderia, Streptococcus, Streptomyces, Leptospira, Neisseria, Azospira, Polynucleobacter, Citrobacter, Listeria, Bacillus, Rhodococcus, Pseudonocardia, Sinorhizobium, Corynebacterium, Cupriavidus, Ralstonia, Campylobacter, Acinetobacter, Stenotrophomonas, Enterococcus, Photorhabdus, Cohaesibacter, and Zobellia*. Among them, *Zobellia. Lactobacillus, Staphylococcus*, and *Clostridium* were detected in all the samples.

The clustering result of the samples and bacterial genera shows that all samples of the NS group clustered in the same clade, as shown in the heat map (Figure 1A). In contrast, the other 3 groups (PO, PS, and NO) clustered in different clades (Figures 1A, B). This result suggests that only the non-PCOS and non-obese group have normal vaginal microbiome, with low diversity and bacterial species of *Lactobacillus*. Thus, obesity and PCOS may lead to dysbiosis of the vaginal microbiome. Bacterial genera were grouped into four clades (Figure 1A). Lactobacillus is in the same clade as 17 other genera, such as *Klebsiella, Acinetobacter, Pseudomonas*, and *Enterococcus*. *Clostridium* is clustered with 8 other genera, including *Streptomyces* and *Gardnerella*, which are often pathogenic. *Staphylococcus* is clustered with *Streptococcus* and *Vibrio*. The fourth clade contains 17 genera, including *Listeria*, *Leptospira*, *Photorhabdus*, and *Azospira*.

There were 21 bacterial genera with significantly different abundances between groups (ANOVA-test $p < 0.05$). Ten genera, *Lactobacillus, Klebsiella, Komagataeibacter, Enterobacter, Pseudomonas, Bacillus, Ralstonia, Acinetobacter, Stenotrophomonas, and Enterococcus*, showed higher abundance in the NS group. In comparison, another 11 genera, *Staphylococcus, Clostridium, Leptospira, Azospira, Citrobacter, Listeria, Pseudonocardia, Sinorhizobium, Corynebacterium,*

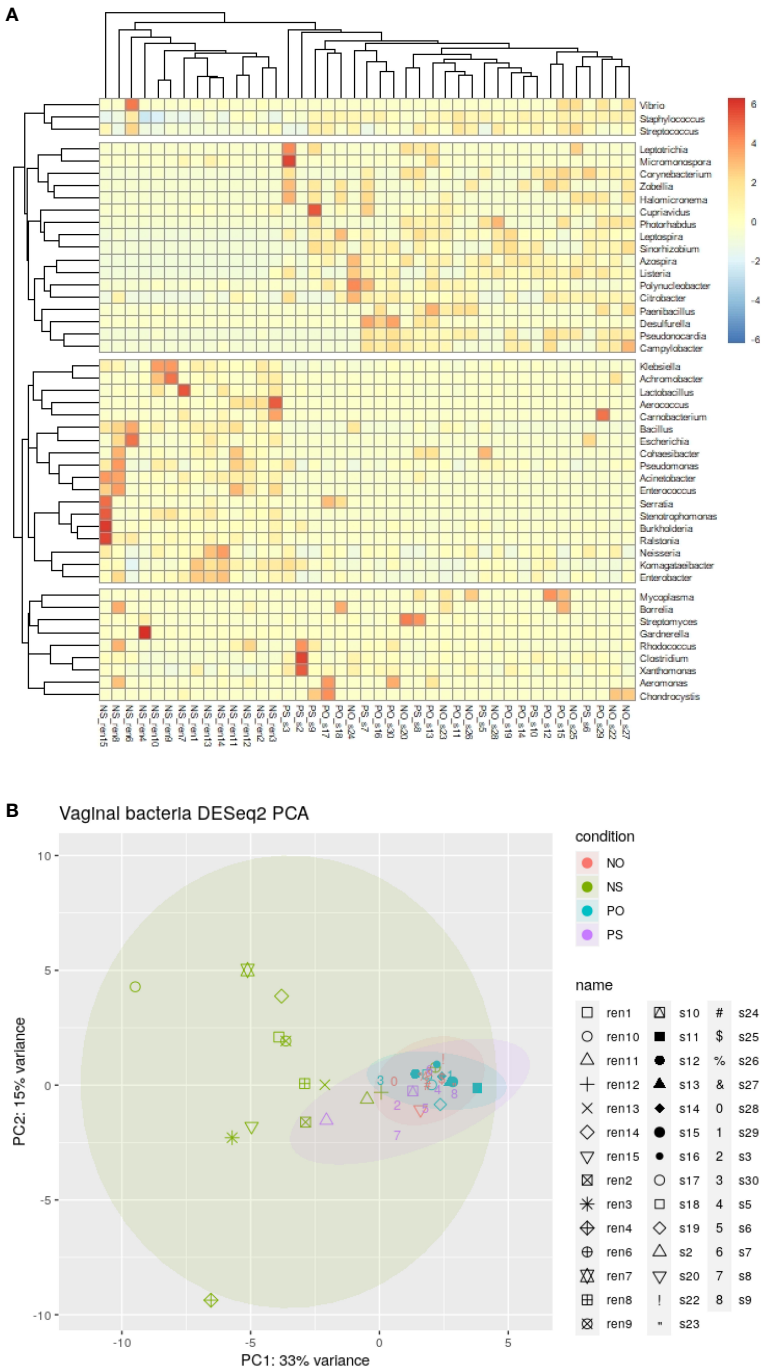


FIGURE 1
The clustering and PCA results of bacteria genera of the 41 vaginal microbiomes. **(A)** Cluster of bacteria genera; **(B)** DESeq2 PCA.

Campylobacter, and *Photorhabdus* were relatively low or not detected in the NS group (Supplementary Figure 1).

The results of the ANOVA test showed that there were 26 bacterial species and 6 phages/viruses with different abundance between the groups (Supplementary Figures 1, 2). In the NS group, 16 species show lower abundance than in other groups, including 5 pathogens and 4 opportunistic pathogens. The 5 pathogens included *Streptococcus pyogenes, Leptospira santarosai, Citrobacter amalonaticus, Listeria ivanovii*, and *Clostridium perfringens*. The 4 opportunistic pathogens included *Staphylococcus aureus, Ralstonia insidiosa, Corynebacterium diphtheriae*, and *Photorhabdus asymbiotica*.

The bacterial species detected in no less than 9 out of 41 specimens are listed in Table 1.

The clustered matrix displaying the Pearson correlation between detected phages/viruses and bacterial species is depicted. Lactobacillus bacteria and Lactobacillus phage are positively correlated. Most of the *Lactobacillus* species and *Pseudomonas* species are highly correlated. A total number of 26 *Lactobacillus* species were identified, including *Lactobacillus crispatus, L. iners, L.*

10.3389/frmbi.2023.1229723

TABLE 1 The frequently detected bacterial species and their frequency in the PCOS group and in the non-PCOS group.

Name	#Specimens detected out of 41	Frequency (%)			PCOS vs Non-PCOS	
		Mean	Min	Max	Ratio (PCOS/ Non-PCOS)	t.test p-value
Staphylococcus aureus▲	41	55.17	7.68	81.37	1.36	0.0008
Komagataeibacter rhaeticus	40	8.12	0	20.26	0.84	0.16
*Klebsiella michiganensis**	37	4.6	0	44.53	0.26	0.03
Klebsiella pneumoniae▲***	38	4.13	0	14.72	0.34	0.0002
*Lactobacillus crispatus**	27	2.29	0	43.32	0.08	0.04
*Xanthomonas citri**	39	1.87	0	9.09	1.88	0.01
Enterobacter cloacae◆*	38	1.83	0	6.24	0.55	0.01
Streptomyces sp. *ICC1*	31	1.75	0	21.51	1.05	0.47
Gardnerella vaginalis▲	9	1.72	0	67.6	0.02	0.16
Burkholderia dolosa▲	37	1.53	0	20.79	0.51	0.15
Pseudomonas fluorescens▲	38	1.44	0	4.57	1.02	0.47
Leptospira santarosai◆**	32	1.17	0	4.63	2.61	0.002
*Azospira oryzae**	26	1.05	0	4.44	1.97	0.02
Clostridium botulinum◆	38	1.04	0	13.64	2.08	0.14
Streptococcus pyogenes◆	34	0.62	0	1.84	1.45	0.08
Polynucleobacter necessaries	25	0.56	0	4.44	1.85	0.12
Lactobacillus johnsonii	26	0.46	0	4.55	0.89	0.41
*Xanthomonas euvesicatoria**	22	0.39	0	4.55	3.3	0.04
Pseudomonas aeruginosa▲	20	0.35	0	3.18	0.45	0.09
Rhodococcus opacus	16	0.31	0	4.55	3.65	0.09
*Lactobacillus delbrueckii**	28	0.3	0	1.52	2.14	0.01
Pseudomonas sp. *PONIH3*	21	0.29	0	1.98	0.87	0.38
Ralstonia insidiosa▲*	14	0.28	0	4.95	0	0.02
Lactobacillus gasseri	19	0.27	0	2.73	1.22	0.36
Listeria ivanovii▲*	20	0.27	0	1.11	2.07	0.02
Citrobacter amalonaticus◆*	21	0.24	0	1.11	1.98	0.03
Lactobacillus amylovorus	13	0.23	0	2.69	0.54	0.19
Streptomyces sp. *ICC4*	16	0.23	0	1.14	1.13	0.4
Sinorhizobium sp. *RAC02***	16	0.22	0	0.83	4.11	0.001
*Lactobacillus jensenii**	10	0.21	0	3.33	0	0.02
*Bacillus cereus****	16	0.2	0	1.52	0.1	0.0008
Corynebacterium diphtheriae◆*	14	0.2	0	0.98	3.6	0.01
Clostridium perfringens◆	15	0.18	0	1.11	1.97	0.08
Clostridium sp. *DL-VIII***	16	0.18	0	0.76	3.99	0.001
*Pseudonocardia dioxanivorans**	16	0.16	0	0.62	2.15	0.04
Cupriavidus oxalaticus	15	0.15	0	1.64	2.48	0.09

(Continued)

TABLE 1 Continued

Name	#Specimens detected out of 41	Frequency (%)			PCOS vs Non-PCOS	
		Mean	Min	Max	Ratio (PCOS/ Non-PCOS)	t.test p-value
*cyanobacterium endosymbiont of Epithemia turgida**	12	0.12	0	0.6	3.5	0.01
*Ralstonia pickettii**	11	0.12	0	1.98	0	0.02
*Zobellia galactanivorans***	10	0.12	0	0.76	5.59	0.008
Photorhabdus asymbiotica◆	11	0.11	0	0.69	1.22	0.36
Enterococcus casseliflavus	11	0.1	0	1.02	0	0.007
Halomicronema hongdechloris	9	0.1	0	0.76	3.08	0.05
Escherichia coli	9	0.07	0	0.98	0.24	0.07
Acinetobacter bereziniae▲*	10	0.06	0	0.99	0	0.02
Cohaesibacter sp. *ES.047*	10	0.06	0	0.51	0.8	0.37
*Stenotrophomonas maltophilia**	11	0.06	0	0.99	0	0.01

◆ pathogens; ▲ opportunistic pathogens; * p<0.05; ** p<0.01; *** p<0.001.

acetotolerans, L. acidipiscis, L. acidophilus, L. alimentarius, L. amylolyticus, L. amylophilus, L. amylovorus, L. apis, L. crustorum, L. delbrueckii, L. gallinarum, L. gasseri, L. heilongjiangensis, L. helveticus, L. jensenii, L. johnsonii, L. kefiranofaciens, L. kullabergensis, L. paracollinoides, L. plantarum, L. reuteri, L. salivarius, L. sanfranciscensis, and *L. sp wkB8*. Most of them are significantly positively correlated (p<0.01) (Supplementary File 1). *Lactobacillus phage.Lv.1* positively correlated with *Lactobacillus iners* (r = 0.41, p = 0.0068), *L. crustorum* (r = 0.43, p = 0.0049), *L. plantarum* (r = 0.44, p = 0.0033), *L. jensenii* (r = 0.52, p = 0.00048), and *L. amylolyticus* (r = 0.58, p = 7.0E-05). There are 68 species of bacteria and 12 species of viruses/phages that are positively correlated with the 26 *Lactobacillus* species, and only one species, *Staphylococcus aureus*, is negatively correlated with *Lactobacillus* species (Supplementary File 1).

Pseudomonas bacteria and *Pseudomonas* phages/viruses are positively correlated (Supplementary File 1). A total of 9 *Pseudomonas* species and 5 *Pseudomonas* viruses/phages were identified, including *Pseudomonas aeruginosa, P. balearica, P. fluorescens, P. pohangensis, P. putida, P. resinovorans, P. rhodesiae, P. stutzeri, and P. syringae;* and *Pseudomonas virus EL, Pseudomonas virus phiCTX, Pseudomonas virus LKA1, Pseudomonas phage PA16*, and *Pseudomonas phage F10*. *Pseudomonas phage F10* and *PA16* correlated significantly with *Pseudomonas aeruginosa, P. pohangensis, P. rhodesiae*, and *P. stutzer*i. *Pseudomonas virus EL* correlated with *P. resinovorans*; *Pseudomonas virus LKA1* correlated with *P. aeruginosa, P. rhodesiae, P. stutzeri*, and *P. syringae*. *Pseudomonas virus phiCTX* correlated with *P. aeruginosa, P. pohangensis*, and *P. rhodesiae*.

The distribution of bacteria of the genera *Lactobacillus* and *Pseudomonas* is positively correlated. Moreover, the *Lactobacillus* phages and *Pseudomonas* viruses/phages are also positively correlated. *Pseudomonas resinovorans* is highly positively correlated with *Lactobacillus acidophilus*, *L. crustorum, L. iners, L. jensenii, L. kefiranofaciens, L. kullabergensis, L. plantarum*, and *L. salivarius* (r>0.80, p=0). *Pseudomonas virus EL* is correlated with *L. crustorum* and *L. iners* (r > 0.80, p = 0).

The majority of the pathogenic bacteria had a positive correlation with each other. *Listeria ivanovii* positively correlates positively with *Clostridium perfringens*, *Clostridium sp DL VIII*, *Citrobacter amalonaticus, Azospira oryzae, Staphylococcus aureus*, and *Streptococcus pyogenes*, most of which are pathogenic. On the other hand, *L. ivanovii* negatively correlates with the benign commensals bacteria *Bacillus cereus* (Figure 2B). *Citrobacter amalonaticus, Corynebacterium diphtheriae, Escherichia virus IME08, Clostridium* sp. *DL VIII*, and *Listeria ivanovii* are all positively correlated with *Clostridium perfringens* (Supplementary File 1). The opportunistic pathogen *Staphylococcus aureus* is negatively correlated with *Lactobacillus crispatus, L. gallinarum, L. helveticus, L. kefiranofaciens, L. plantarum*, and *L. salivarius* (p<0.001).

3.2.2 Viruses/phages and their correlation with bacteria

There are 24 viruses/phages annotated at the species level (Table 2). Among these viruses, *Cyprinid herpesvirus 3* shows the highest frequency and the widest distribution. It is detected in 33 out of the 41 samples, and the average frequency is 51.00% of the viruses/phages. *Human endogenous retrovirus K* was detected in 28 samples with an average relative frequency of 21.84%.

The relative abundance of 6 virus species (*Pseudomonas virus EL*, *Pseudomonas phage PA16*, *Pseudomonas phage F10*, *Human endogenous retrovirus K*, *Cyprinid herpesvirus 3*, and *Escherichia virus IME08*) is significantly different between groups (ANOVA test, p < 0.05). The first four of them are much more abundant in the NS (healthy control) group, while *Cyprinid herpesvirus 3* is much less abundant in the NS group (Figure 2B). *Cyprinid herpesvirus 3* is negatively correlated with *Pseudomonas virus EL* (r = −0.43, p = 0.0052) and *Human endogenous retrovirus K* (r = −0.55, p = 0.00016).

10.3389/frmbi.2023.1229723

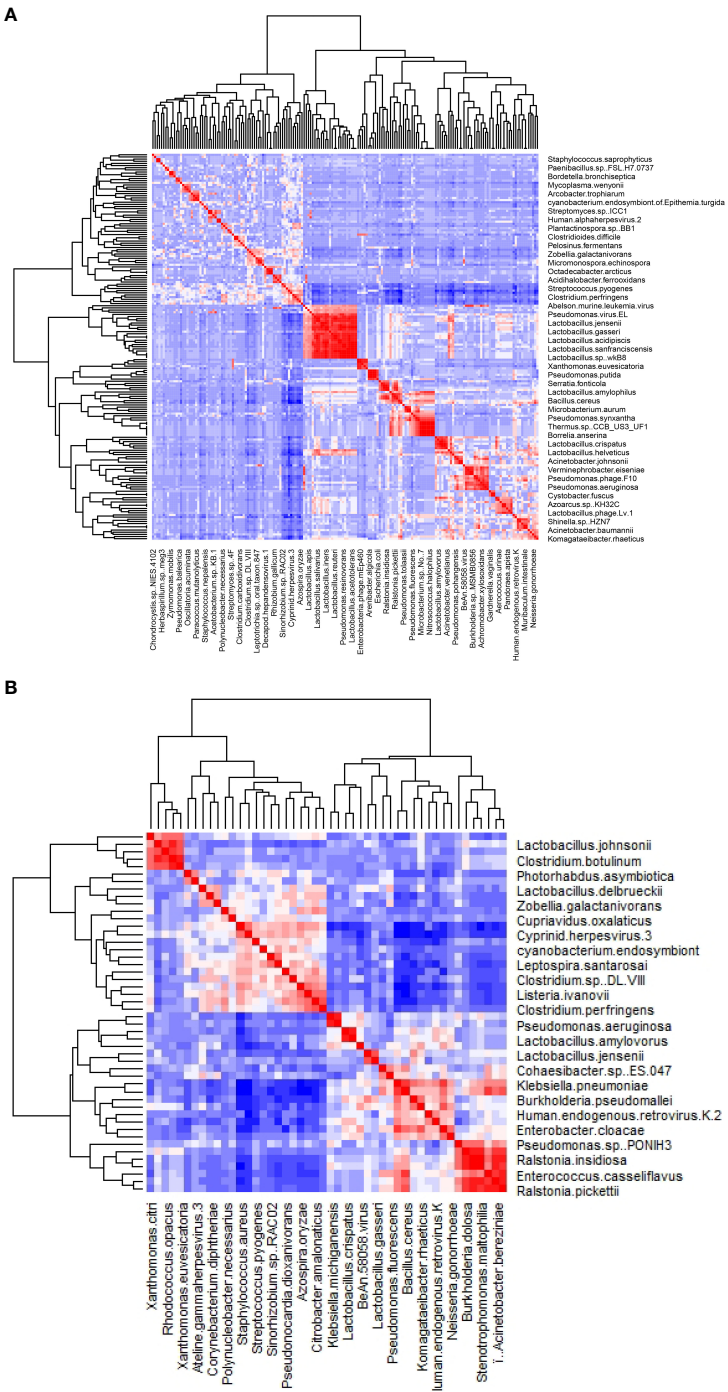


FIGURE 2
The heatmap of the correlation between detected phages/viruses and bacterial species. **(A)** All species. **(B)** Significantly correlated species with abundance above 0.1% ($p<0.01$).

3.2.3 Bacteria differentially enriched in PCOS and obesity groups

LEfSe analysis revealed that the differentially enriched bacteria in the PCOS (P) and non-PCOS (N), and obesity (O) and non-obese (S) groups were quite different (Figure 3).

As shown in Figure 3A, the bacterial species differentially enriched in the PCOS (P) groups include *Clostridium* sp. *DL-VIII*, *Xanthomonas citri*, *Xanthomonas euvesicatoria*, *Lactobacillus delbrueckii*, and *Sinorhizobium* sp. *RAC02*. However, the bacterial species differentially enriched in the obese (O) groups are different and include *Azospira oryzae*, *Citrobacter amalonaticus*, *Polynucleobacter necessaries*, *Streptococcus pyogenes*, *Staphylococcus aureus*, and *Pseudonocardia dioxanivorans* (Figure 3C). Three of the enriched O-group bacteria are known as human pathogens that may be associated with the pathogenesis of obesity through inflammation (Avire et al., 2021).

TABLE 2 The annotated virus/phage species and abundance.

Name	#Specimens detected out of 41	Frequency (%)			PCOS vs Non-PCOS	
		Mean	Min	Max	Ratio (PCOS/ Non-PCOS)	t.test p-value
Lactobacillus phage Lv 1	2	0.99	0	32.58	1.32E-01	0
*Pseudomonas virus EL***	9	3.19	0	35.49	6.38E-03	0
Pseudomonas virus phiCTX	7	0.80	0	14.37	5.15E-02	0
*Pseudomonas virus LKA1**	6	0.47	0	6.54	4.17E-02	0
Pseudomonas phage B3	4	0.43	0	16.47	1.65E-01	0
*Pseudomonas phage F10**	8	1.73	0	20.32	2.16E-02	0
*Pseudomonas phage PA16***	9	1.24	0	14.07	1.73E-02	0
Enterobacteria phage mEp460	7	0.50	0	15.70	2.22E-01	4.78
Microbacterium phage Min1	4	0.01	0	0.56	1.65E-01	0
Salmonella phage Fels 1	4	0.10	0	3.82	1.65E-01	0
*Cyprinid herpesvirus 3****	36	51.00	0	100.00	8.05E-06	3.52
*Human endogenous retrovirus K****	28	21.84	0	95.61	9.95E-04	0.16
Human alphaherpesvirus 2	12	3.62	0	97.48	2.61E-01	2.78
Human gammaherpesvirus 4	4	0.03	0	1.26	1.65E-01	0
Ateline gammaherpesvirus 3	13	3.10	0	59.64	4.23E-01	0.80
Alcelaphine gammaherpesvirus 1	8	0.24	0	5.43	3.57E-01	1.64
Abelson murine leukemia virus	7	0.19	0	3.67	5.30E-02	0
Callitrichine gammaherpesvirus 3	4	0.00	0	0.13	1.65E-01	0
Murine osteosarcoma virus	8	2.69	0	94.85	1.84E-01	13.12
Macacine alphaherpesvirus 1	4	0.01	0	0.20	1.65E-01	Infinity
Ndumu virus	4	0.02	0	0.78	1.65E-01	Infinity
Pestivirus Giraffe 1	4	0.03	0	1.02	1.65E-01	0
Ross River virus	6	0.04	0	0.77	6.43E-02	Infinity
Semliki Forest virus	4	0.07	0	2.76	1.65E-01	Infinity

The cladogram plot showed that two genera, *Xanthomonas* and *Leptospira*, and one family, *Clostridiaceae*, were enriched in the PCOS group (Figure 3B), while there are three genera, *Polynucleobacter, Azospira, Citrobacter*, and *Epsilonproteobacteria*, and two families, *Pseudonocardiaceae* and *Staphylococcaceae*, were enriched in the obesity group (Figure 3D). Bacteria of the genus *Leptospira*, which are enriched in the PCOS group, can cause a bacterial disease, leptospirosis, that affects humans and animals.

Enriched bacteria and phages in each group were listed as follows:

(1) Bacteria enriched only in PO, PS, NO, and NS groups:

Bacteria enriched in the NS group include *Komagataeibacter rhaeticus, Burkholderia pseudomallei, Enterobacter cloacae, Klebsiella pneumoniae, Bacillus cereus*, and *Lactobacillus.*

Bacteria enriched in the NO group include *Azospira oryzae, Clostridium perfringens, Streptococcus pyogenes, Listeria ivanovii, Citrobacter amalonaticus*, and *Photorhabdus asymbiotica.*

Bacteria enriched in the PS group include *Clostridium* sp.*DL_VIII, Xanthomonas citri, Corynebacterium diphtheriae, Zobellia galactanivorans*, and *Sinorhizobium* sp.*RAC02.*

Bacteria enriched in the PO group include *Leptospira santarosai, Polynucleobacter necessaries, Staphylococcus aureus*, and *Pseudonocardia dioxanivorans.*

(2) Bacteria and phages that are commonly enriched between the groups:

The bacteria and phages/viruses with significantly different relative abundances between PO and NS and between PS and NS are listed in Table 3.

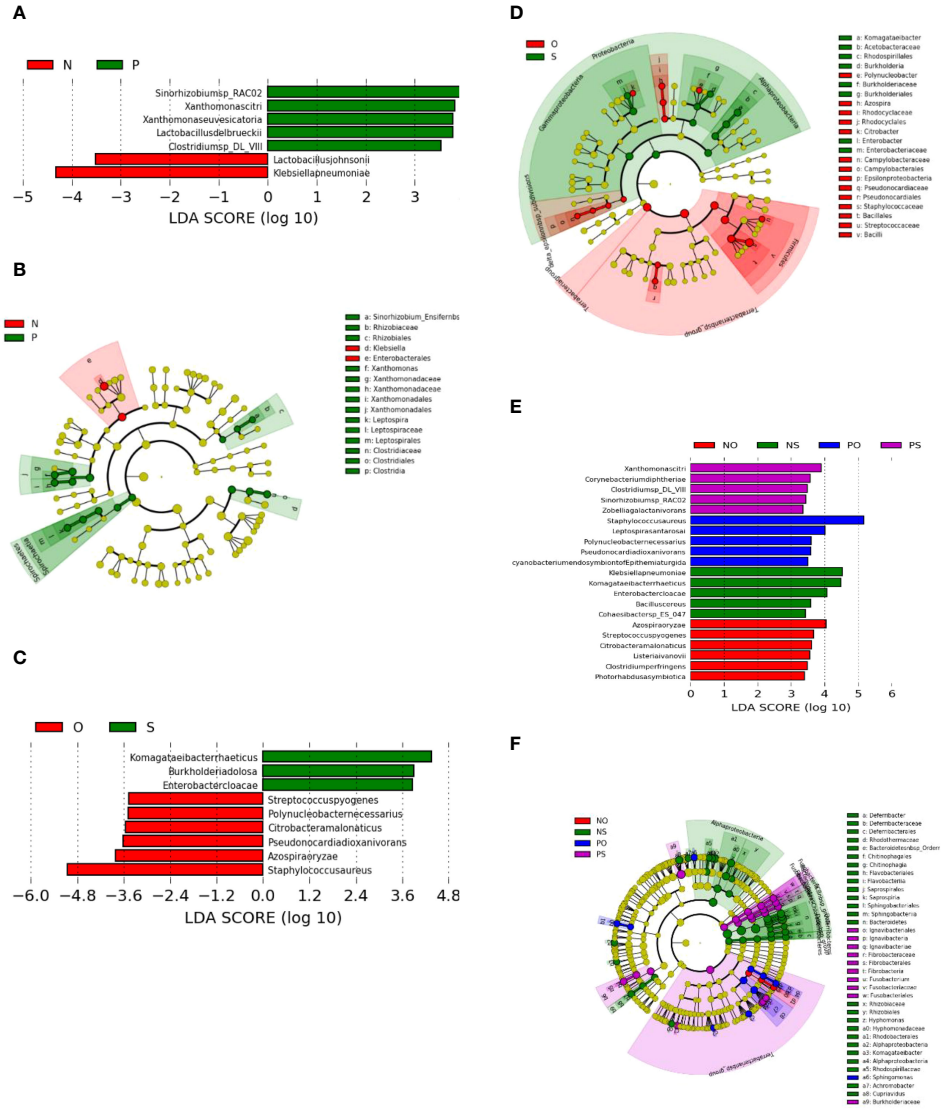


FIGURE 3
LEfSe plots of the bacteria differently enriched in the PCOS (P) group and non-PCOS (N) group, and in the Obesity (O) group and the non-obese (S) group. **(A, B)** PCOS (P) vs Non-PCOS (P); **(C, D)** Obese (O) vs Non-obese (S); **(E, F)** Non-PCOS and Obese (NO), Non-PCOS and Non-obese (NS), PCOS and Obese (PO), PCOS and Non-obese (PS).

3.3 Microbiota function annotation

The rarefaction curve of gene families retrieved from Humann and OTUs output from Kraken indicate that the sequencing depth is adequate for metagenomic analysis (Supplementary Figure 3).

The functional analysis using HUMAnN output revealed that a total of 23,438 UniRef90 genes were identified, with 1,278 classified under *Lactobacillus crispatus*, 607 under *L. iners*, and 21,553 remaining unclassified.

The pathway abundance results encompassed 27 pathways, with the majority of the samples falling within group NS and group NO. Notably, the PWY-3781 pathway, associated with aerobic respiration I (cytochrome c), was exclusively observed in the obese group. A more specific omics study should be undertaken to shed light on the functionality of the PCOS and obesity microbiome.

4 Discussion

The main findings of this paper are that the vaginal flora of non-PCOS and non-obese reproductive women are similar and have more bacterial phages, whereas both PCOS and obesity factors lead to alterations in the vaginal flora, and these alterations are diverse.

4.1 Vaginal microbiome composition and differential enrichment of bacteria

The results of this study revealed distinct differences in the composition of vaginal microbiomes between the PCOS (P) and non-PCOS (N) groups, as well as between the obesity (O) and non-obese (S) groups (Table 3). The non-PCOS group exhibited a relatively normal vaginal microbiome with low diversity. In

TABLE 3 Bacteria and phages enrichment between PO and NS, and between PS and NS.

Name	PO_vs_NS		PS_vs_NS	
	t.test p-value	Ratio	t.test p-value	Ratio
Staphylococcus aureus	1.19E-05	1.85	6.65E-04	1.62
Klebsiella michiganensis	4.39E-02	0.22	3.13E-02	0.14
Lactobacillus crispatus	ND	ND	4.35E-02	0.02
Komagataeibacter rhaeticus	7.96E-03	0.58	ND	ND
Klebsiella pneumoniae	2.06E-05	0.23	9.48E-05	0.28
Enterobacter cloacae	1.45E-03	0.34	2.57E-02	0.54
Xanthomonas citri	2.41E-03	2.12	ND	ND
Lactobacillus jensenii	2.88E-02	0	2.88E-02	0
Leptospira santarosai	1.48E-03	16.78	1.54E-03	16.85
Burkholderia pseudomallei	1.52E-03	0.12	4.31E-04	0
Bacillus cereus	1.94E-04	0.07	1.88E-04	0.06
Azospira oryzae	5.26E-05	Infinity	1.70E-03	infinity
Ralstonia insidiosa	1.40E-02	0	1.40E-02	0
Streptococcus pyogenes	4.33E-04	3.41	ND	ND
Staphylococcus aureus	1.19E-05	1.85	6.65E-04	1.62
Pseudomonas_virus_EL	4.74E-03	0	4.74E-03	0
Pseudomonas_virus_phiCTX	5.16E-02	0	5.16E-02	0
Pseudomonas_virus_LKA1	4.13E-02	0	4.13E-02	0
Enterobacteria_phage_mEp460	1.04E-01	0	2.10E-01	7.17
Lactobacillus_phage_Lv_1	1.35E-01	0	1.35E-01	0
Microbacterium_phage_Min1	1.69E-01	0	1.69E-01	0
Pseudomonas_phage_B3	1.69E-01	0	1.69E-01	0
Pseudomonas_phage_F10	2.01E-02	0	2.01E-02	0
Pseudomonas_phage_PA16	1.57E-02	0	1.57E-02	0
Salmonella_phage_Fels_1	1.69E-01	0	1.69E-01	0

ND means non-detected in both groups; infinity means non-detected in the NS group.

contrast, the PCOS and obesity groups showed signs of dysbiosis characterized by increased microbial diversity.

Among the various bacterial genera identified in the vaginal microbiomes, several genera were significantly different in abundance between the groups. Notably, *Lactobacillus*, *Pseudomonas, Bacillus*, and other genera were enriched in the non-PCOS and non-obese (NS) group, while Staphylococcus, Clostridium, Leptospira, and others were relatively more abundant in the PCOS and obesity groups. It is worth noting that several pathogenic and opportunistic pathogens, including *Streptococcus pyogenes*, *Leptospira santarosai*, *Citrobacter amalonaticus*, *Listeria ivanovii*, and *Clostridium perfringens*, were detected in higher abundance in the PCOS and obesity groups. This finding raises concerns about potential health risks associated with the alteration of the vaginal microbiome in these conditions.

The higher abundance of benign commensals bacteria could inhibit the growth of pathogens. The fact that the pathogenic bacteria were lower in the NS group could be explained by the former finding (Kang et al., 2017); the lactic acid and other chemicals produced by *Lactobacillus* could inhibit the growth of *Staphylococcus aureus* and prevent the formation of biofilm (Gulzar, 2018). Generally, the *Lactobacillus* species is the dominant bacteria in the vaginal microbiome of healthy reproductive-age women (Spurbeck and Arvidson, 2011). *Lactobacillus* species are the common bacteria that sustain health. The abundance of *Lactobacillus* species is significantly higher in the NS group. This result is concordant with the previous study (Spurbeck and Arvidson, 2011; Greenbaum et al., 2019). *Lactobacillus* species sustain the acid and anaerobic conditions for the normal vaginal microenvironment and form the front line to defend against

vaginally acquired infections (Spurbeck and Arvidson, 2011). But in the PO, PS, and NO groups, *Lactobacillus* species are not so dominant, and this suggests dysbiosis of the vaginal microbiome is linked with obesity or PCOS. The disturbance of normal vaginal bacteria has been approved to be linked with different types of diseases, such as bacterial vaginosis and virus infection (Spurbeck and Arvidson, 2011).

In this study, *Pseudomonas* species are also much higher in the NS group (Figure 2), suggesting they interact with each other and keep a balance. *Pseudomonas* species are not generally considered pathogenic, but studies have identified that they are present at a low level in the indigenous microbiota of various body sites, including the uterus, and can lead to pelvic inflammatory disease (Scales et al., 2014). The presence of Pseudomonas and Pseudomonas phages in the NS group but not in the PCOS or obesity groups suggests that, while Pseudomonas is a common bacteria in vaginal microbiomes.

Bacillus cereus is a facultative anaerobe has mechanisms for both aerobic and anaerobic respiration. Some of the major products produced from carbon sources such as sucrose or glucose during anaerobic respiration include L-lactate, acetate, formate, succinate, ethanol, and carbon dioxide (Mols et al., 2007). These metabolites produced by *B. cereus* help to sustain the acid conditions in the vaginal microenvironment and contribute to the health.

On the other hand, most of the pathogenic bacteria lower in the NS group are positively correlated with each other, such as *Listeria ivanovii*, *Clostridium perfringens*, *Citrobacter amalonaticus*, *Staphylococcus aureus*, and *Streptococcus pyogenes*; and *L. ivanovii* negatively correlated with the benign commensals bacteria *Bacillus cereus*. Dysbiosis can lead to the overgrowth of pathogens, which might be linked to an increased risk of infections, inflammation, and other health issues. Understanding these associations could have implications for diagnostic and therapeutic strategies in managing PCOS and obesity. This result suggested that in the PCOS or obesity groups, some groups of pathogens linked with the diseases. Furthermore, the pathogens linked with PCOS and obesity are quite different, suggesting the vaginal microbiome's pathogenic mechanisms are different between PCOS and obesity.

In summary, the vaginal bacteria in healthy are similar, but PCOS and obesity are associated with different bacterial shifts from the normal one which are benign commensals and produce lactic acids and help to sustain the acid conditions.

4.2 Identification of key phages and the possible function

This study annotates 14 genera (24 species) of viruses and phages. This suggests that the distribution of viruses or phages in the vaginal microbiome is very common. Bacteriophages contribute to the maintenance of ecological balance and the evolution of bacterial species (Martín et al., 2009; Wright, 2021). Regarding the function of the vaginal bacteria phage, a previous study has illustrated that the *Lactobacillus* phage can eliminate the *Lactobacillus* culture in a petri dish (Kilic et al., 2001; Miller-Ensminger et al., 2020). But in this study, *Lactobacillus phage Lv1* was only found in two specimens of the NS group, and in these two specimens, several *Lactobacillus bacteria* were detected. This implies that the *Lactobacillus* phage contributes to the ecological balance of vaginal microecology while not eliminating the *Lactobacillus*. Furthermore, the presence of the *Lactobacillus* phage is an indicator of the presence of relatively abundant *Lactobacillus* (Miller-Ensminger et al., 2020).

Lactobacillus phage Lv1 is a dsDNA virus isolated from a human vaginal *Lactobacillus jensenii* strain (Martín et al., 2010), which coexisted and positively correlated in this study. It has been illustrated that the stable coexistence of phages and bacteria is very common in the urogenital microbiota (Miller-Ensminger et al., 2020; Wright, 2021). The coexistence of lactobacilli and phages in the vaginal microbiome is widely spread, and many of the lysogenic lactobacilli could release phages at a high frequency (Kilic et al., 2001; Martín et al., 2009; Wright, 2021). In summary, lysogenic phages have a complex relationship with their host bacteria. They can both provide advantages and pose threats to the host, depending on the specific genes carried by the prophage and the environmental conditions. This interplay between lysogenic phages and bacteria is an important aspect of microbial ecology and evolution. lysogeny in vaginal lactobacilli is widely spread. Some lysogenic lactobacilli spontaneously release phages with a broad host range. Even though the phages were all temperate, they were able to cause lytic infection in various strains (Kiliç et al., 2001). The detailed relationship between *Lactobacillus phage Lv1 and Lactobacilli* in vaginal microbiome needs further study.

In this study, *Pseudomonas* phages/viruses, including *Pseudomonas virus EL, F10*, *phiCTX*, B3, PA16, and *LKA1*, have the common host *Pseudomonas aeruginosa* (Kwan et al., 2006; de Smet et al., 2017; Miller-Ensminger et al., 2020). *Pseudomonas aeruginosa* is an opportunistic pathogen that grows on the skin in moist parts, including the genital area, and it can cause severe infection when the immune system is weakened. A recent study suggested that the colonization of *Pseudomonas aeruginosa* in the genital tract can lead to infertility by inhibiting sperm mobility (Vander and Prabha, 2019). This could explain why some dysbiosis of the vaginal microbiome may lead to infertility. On the other hand, phage therapy has been proven to be an effective way to treat infection with *Pseudomonas aeruginosa* (Pires et al., 2015). The relationship between *Pseudomonas* phages/viruses and the presence of *Pseudomonas* in the vaginal microbiome, and their impacts on health needs further investigate.

The abundance of human endogenous retrovirus K (HERV-K) is significantly higher in the NS group when compared with the PS, PO, and NO groups. The most recent study suggests that HERV-K has multiple copies in the human genome and that some of its open reading frames are transcribed and translated in early embryogenesis (Garcia-Montojo et al., 2018). Though HML-2 of HERV-K has been reported to be associated with certain types of cancer and neurodegenerative diseases in adult tissues, the functional associations between HERV-K (HML-2) and diseases have not been completely clarified, and further work is still needed to confirm the specific HERV loci associated with specific diseases (Xue et al., 2020). In this study, we only sequenced DNA, not RNA, and the sequencing depth was relatively low, so it cannot reflect HERV-K transcriptional levels in the vaginal or specific HERV-K

loci in the genome. The relationship between HERV-K and PCOS needs further study.

Among these viruses, *Cyprinid herpesvirus 3* showed the highest frequency and widest distribution. It is detected in 33 out of the 38 samples, and the average frequency is 50% of the viruses/phages. *Cyprinid herpesvirus 3*, also known as CyHV-3, koi herpes virus or KHV, is a species of dsDNA virus that causes a viral disease that is very contagious to the common carp *Cyprinus carpio*. CyHV-3 has been proposed as a biological control agent for the invasive carp in Australia, and the safety to non-target species has been tested, which suggests that CyHV-3 is safe for the non-target species tested, including 14 species of fish, two species each of amphibians, and reptiles, and single species of bird, mammal, and invertebrates (McColl et al., 2016). CyHV-3 is a large double-stranded DNA herpesvirus classified in the Alloherpesviridae family in the Herpesvirales order. This order encompasses herpesviruses that infect different hosts, including mammals, avians, and reptiles. CyHV-3 is divergent from the other mammalian, avian, and reptilian herpesviruses, and bears homolog genes similar to CyHV-1, CyHV-2, AngHv-1, IcHV-1, and RaHV-1 (Ilouze et al., 2012). In this study, the annotation of *Cyprinid herpesvirus 3* from the human vaginal microbiota shotgun metagenomes may be because there is a CyHV-3-like virus in the specimens. The high frequency of this CyHV-3-like virus may contribute to diseases such as PCOS. But more information about this CyHV-3-like virus in the vaginal microbiome and its function needs further study.

4.3 Microbial correlations

The analysis of correlations between bacterial genera and phages/viruses provided insights into the complex interactions within the vaginal microbiome. Notably, there were strong positive correlations among various *Lactobacillus* species and among *Pseudomonas* species, suggesting the presence of microbial communities with shared ecological niches. The positive correlation between *Lactobacillus phages* and *Lactobacillus*, as well as *Pseudomonas phages* and *Pseudomonas* suggest lysogenic phages are also involved in the microbiome regulation and may play an role. The third correlation is the negative one between benign commensal bacteria and pathogenic bacteria, which suggests that benign commensal bacteria play a central role in the maintenance of vaginal health. It also indicates that the disruption of the benign commensal ecology will cause a dysbiosis of the vaginal flora, which lead to invasion of pathogenic bacteria and cause PCOS in women of reproductive age by inducing inflammation, causing insulin resistance, and overproduction of androgens (Azziz et al., 2009).

4.4 Clinical implications

The results of this study highlight the potential clinical significance of vaginal microbiome dysbiosis in PCOS and obesity. Dysbiosis can lead to the overgrowth of pathogens, which might be linked to an increased risk of infections, inflammation, and other health issues. Understanding these associations could have implications for diagnostic and therapeutic strategies in managing PCOS and obesity. Because the interaction between the microbiota and polycystic ovary syndrome is a bidirectional one, it is likely that restoration of the vaginal flora may contribute to the relief of PCOS symptoms by reducing inflammation and restoring normal microbiome metabolites and their signaling.

4.5 Limitations and future directions

It's important to acknowledge the limitations of this study, such as the relatively small sample size and the observational nature of the findings. Further research with larger and more diverse cohorts is needed to validate these results. Additionally, future studies may explore the mechanisms underlying the observed microbiome alterations and their potential impacts on gynecological and metabolic health.

5 Conclusion

Non-PCOS and non-obese women share a similar vaginal microbiome, but both obesity and PCOS are associated with vaginal microbiome dysbiosis with increased abundance of pathogenic bacteria. *Lactobacillus phages* and *Pseudomonas phages* are enriched in the NS group and positive correlated with their host *Lactobacilli* and *Pseudomonas*. The presence of these phages are also indicators of a healthy vaginal microbiome. The differential enrichment of bacteria in PCOS and obesity is different, suggesting that these two diseases are associated with some different microbiota mechanisms.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, SRR24561304 to SRR24561344.

Ethics statement

The studies involving humans were approved by the Human Ethics Committee of the First Affiliated Hospital of Xiamen University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SZ and HY designed the project. SZ organized the data analysis and manuscript preparation. XZ, HC, TF, XQ, and MW participated in the experiments and contributed to data curation,

investigation, review, and editing. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was partially supported by grants from Health Science and Technology Program of Xiamen, China (3502Z20174068) and the First Affiliated Hospital of Xiamen University (XYY2016002), and partially from the grants from Third Institute of Oceanography, Ministry of Natural Resources, Xiamen, China (HE170701).

Conflict of interest

Author XQ and TF were employed by Yanxuan Biotech Ltd.

The remaining 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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2023.1229723/full#supplementary-material

References

Avire, N. J., Whiley, H., and Ross, K. (2021). A review of streptococcus pyogenes: Public health risk factors, prevention and control. *Pathogens* 10, 1–18. doi: 10.3390/pathogens10020248

Azziz, R., Carmina, E., Dewailly, D., Diamanti-Kandarakis, E., Escobar-Morreale, H. F., Futterweit, W., et al. (2009). The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertility Sterility*. 91 (2), 456-488. doi: 10.1016/j.fertnstert.2008.06.035

Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3. *Elife* 10, 1–42. doi: 10.7554/eLife.65088

Chen, X., Lu, Y., Chen, T., and Li, R. (2021). The female vaginal microbiome in health and bacterial vaginosis. *Front. Cell Infect. Microbiol.* 11. doi: 10.3389/fcimb.2021.631972

Chu, W., Han, Q., Xu, J., Wang, J., Sun, Y., Li, W., et al. (2022). Metagenomic analysis identified microbiome alterations and pathological association between intestinal microbiota and polycystic ovary syndrome. *Fertil Steril* 113, 1286–1298.e4. doi: 10.1016/j.fertnstert.2020.01.027

Crespo, R. P., Bachega, T. A. S. S., Mendonça, B. B., and Gomes, L. G. (2018). An update of genetic basis of PCOS pathogenesis. *Arch. Endocrinol. Metab.* 62, 352–361. doi: 10.20945/2359-3997000000049

de Smet, J., Hendrix, H., Blasdel, B. G., Danis-Wlodarczyk, K., and Lavigne, R. (2017). Pseudomonas predators: Understanding and exploiting phage-host interactions. *Nat. Rev. Microbiol.* 15, 517–530. doi: 10.1038/nrmicro.2017.61

Dinghua, L., Chi-Man, L., Ruibang, L., Kunihiko, S., and Tak-Wah, L. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly *via* succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033

Duan, L. Y., An, X. D., Zhang, Y. H., Jin, D., Zhao, S. H., Zhou, R. R., et al. (2021). Gut microbiota as the critical correlation of polycystic ovary syndrome and type 2 diabetes mellitus. *Biomed. Pharmacother.* 142, 112094. doi: 10.1016/j.biopha.2021.112094

Garcia-Montojo, M., Doucet-O'Hare, T., Henderson, L., and Nath, A. (2018). Human endogenous retrovirus-K (HML-2): a comprehensive review. *Crit. Rev. Microbiol.* 44, 715–738. doi: 10.1080/1040841X.2018.1501345

Greenbaum, S., Greenbaum, G., Moran-Gilad, J., and Weintruab, A. Y. (2019). Ecological dynamics of the vaginal microbiome in relation to health and disease. *Am. J. Obstet Gynecol* 220, 324–335. doi: 10.1016/j.ajog.2018.11.1089

Gulzar, M. (2018). Staphylococcus aureus: A brief review. *Int. J. Veterinary Sci. Res.* 4 (1), 020–022. doi: 10.17352/ijvsr.000031

Hao, M., Yuan, F., Jin, C., Zhou, Z., Cao, Q., Xu, L., et al. (2016). Overexpression of Lnk in the ovaries is involved in insulin resistance in women with polycystic ovary syndrome. *Endocrinology* 157, 3709–3718. doi: 10.1210/en.2016-1234

He, F. F., and Li, Y. M. (2020). Role of gut microbiota in the development of insulin resistance and the mechanism underlying polycystic ovary syndrome: A review. *J. Ovarian Res.* 13, 1–13. doi: 10.1186/s13048-020-00670-3

Hennig, B. P., Velten, L., Racke, I., Tu, C. S., Thoms, M., Rybin, V., et al. (2018). Large-scale low-cost NGS library preparation using a robust Tn5 purification and tagmentation protocol. *G3: Genes Genomes Genet.* 8, 79–89. doi: 10.1534/g3.117.300257

Hong, X., Qin, P., Huang, K., Ding, X., Ma, J., Xuan, Y., et al. (2020). Association between polycystic ovary syndrome and the vaginal microbiome: A case-control study. *Clin. Endocrinol. (Oxf)* 93, 52–60. doi: 10.1111/cen.14198

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Ilouze, M., Dishon, A., and Kotler, M. (2012). Coordinated and sequential transcription of the cyprinid herpesvirus-3 annotated genes. *Virus Res.* 169, 98–106. doi: 10.1016/j.virusres.2012.07.015

Insenser, M., Murri, M., Del Campo, R., Martínez-García, M.Á., Fernández-Durán, E., and Escobar-Morreale, H. F. (2018). Gut microbiota and the polycystic ovary syndrome: Influence of sex, sex hormones, and obesity. *J. Clin. Endocrinol. Metab.* 103, 2552–2562. doi: 10.1210/jc.2017-02799

Jobira, B., Frank, D. N., Pyle, L., Silveira, L. J., Kelsey, M. M., Garcia-Reyes, Y., et al. (2020). Obese adolescents with PCOS have altered biodiversity and relative abundance in gastrointestinal microbiota. *J. Clin. Endocrinol. Metab.* 105, 1–32. doi: 10.1210/clinem/dgz263

Kang, M. S., Lim, H. S., Oh, J. S., Lim, Y. J., Wuertz-Kozak, K., Harro, J. M., et al. (2017). Antimicrobial activity of Lactobacillus salivarius and Lactobacillus fermentum against Staphylococcus aureus. *Pathog. Dis.* 75, 1–10. doi: 10.1093/femspd/ftx009

Kiliç, A. O., Pavlova, S. I., Alpay, S., Kiliç, S. S., and Tao, L. (2001). Comparative study of vaginal Lactobacillus phages isolated from women in the United States and Turkey: Prevalence, morphology, host range, and DNA homology. *Clin. Diagn. Lab. Immunol.* 8, 31–39. doi: 10.1128/CDLI.8.1.31-39.2001

Kilic, A. O., Pavlova, S. I., Alpay, S., Kilic, S. S., and Tao, L. (2001). Comparative study of vaginal lactobacillus phages isolated from women in the United States and Turkey: Prevalence, morphology, host range, and DNA homology downloaded from http://cvi.asm.org/ on february 13 , 2018 by UNIV OF CAPE TOWN LIBRARIES ELECTR JN. *Clin. Diagn. Lab. Immunol.* 8, 31–39. doi: 10.1128/CDLI.8.1.31

Kwan, T., Liu, J., DuBow, M., Gros, P., and Pelletier, J. (2006). Comparative genomic analysis of 18 Pseudomonas aeruginosa bacteriophages. *J. Bacteriol* 188, 1184–1187. doi: 10.1128/JB.188.3.1184-1187.2006

Lindheim, L., Bashir, M., Münzker, J., Trummer, C., Zachhuber, V., Leber, B., et al. (2017). Alterations in gut microbiome composition and barrier function are associated with reproductive and metabolic defects in women with polycystic ovary syndrome (PCOS): A pilot study. *PLoS One* 12, 1–20. doi: 10.1371/journal.pone.0168390

Liu, R., Zhang, C., Shi, Y., Zhang, F., Li, L., Wang, X., et al. (2017). Dysbiosis of gut microbiota associated with clinical parameters in polycystic ovary syndrome. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00324

Łusiak-Szelachowska, M., Weber-Dabrowska, B., Żaczek, M., Borysowski, J., and Górski, A. (2020). The presence of bacteriophages in the human body: Good, bad or neutral? *Microorganisms* 8, 1–15. doi: 10.3390/microorganisms8122012

Martín, R., Escobedo, S., and Suárez, J. E. (2010). Induction, structural characterization, and genome sequence of Lv1, a prophage from a human vaginal Lactobacillus jensenii strain. *Int. Microbiol.* 13, 113–121. doi: 10.2436/20.1501.01.116

Martín, R., Soberón, N., Escobedo, S., and Suárez, J. E. (2009). Bacteriophage induction versus vaginal homeostasis: Role of H2O2 in the selection of Lactobacillus defective prophages. *Int. Microbiol.* 12, 131–136. doi: 10.2436/20.1501.01.90

McColl, K. A., Sunarto, A., and Holmes, E. C. (2016). Cyprinid herpesvirus 3 and its evolutionary future as a biological control agent for carp in Australia. *Virol. J.* 13, 4–7. doi: 10.1186/s12985-016-0666-4

Miko, E., and Barakonyi, A. (2023). The role of hydrogen-peroxide (H2O2) produced by vaginal microbiota in female reproductive health. *Antioxidants* 1055, 1–17. doi: 10.3390/antiox12051055

Miller-Ensminger, T., Mormando, R., Maskeri, L., Shapiro, J. W., Wolfe, A. J., and Putonti, C. (2020). Introducing Lu-1, a novel lactobacillus jensenii phage abundant in the urogenital tract. *PLoS One* 15, 1–16. doi: 10.1371/journal.pone.0234159

Mols, M., De Been, M., Zwietering, M. H., Moezelaar, R., and Abee, T. (2007). Metabolic capacity of Bacillus cereus strains ATCC 14579 and ATCC 10987 interlinked with comparative genomics. *Environ. Microbiol.* 9, 2933–2944. doi: 10.1111/j.1462-2920.2007.01404.x

Pires, D. P., Vilas Boas, D., Sillankorva, S., and Azeredo, J. (2015). Phage therapy: A step forward in the treatment of pseudomonas aeruginosa infections. *J. Virol.* 89, 7449–7456. doi: 10.1128/jvi.00385-15

Qi, X., Yun, C., Sun, L., Xia, J., Wu, Q., Wang, Y., et al. (2019). Gut microbiota–bile acid–interleukin-22 axis orchestrates polycystic ovary syndrome. *Nat. Med.* 25, 1225–1233. doi: 10.1038/s41591-019-0509-0

Scales, B. S., Dickson, R. P., Lipuma, J. J., and Huffnagle, G. B. (2014). Microbiology, genomics, and clinical significance of the Pseudomonas fluorescens species complex, an unappreciated colonizer of humans. *Clin. Microbiol. Rev.* 27, 927–948. doi: 10.1128/CMR.00044-14

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, 1–18. doi: 10.1186/gb-2011-12-6-r60

Spurbeck, R. R., and Arvidson, C. G. (2011). Lactobacilli at the front line of defense against vaginally acquired infections. *Future Microbiol.* 6, 567–582. doi: 10.2217/fmb.11.36

Sun, Y., Gao, S., Ye, C., and Zhao, W. (2023). Gut microbiota dysbiosis in polycystic ovary syndrome: Mechanisms of progression and clinical applications. *Front. Cell Infect. Microbiol.* 13. doi: 10.3389/fcimb.2023.1142041

Torres, P. J., Siakowska, M., Banaszewska, B., Pawelczyk, L., Duleba, A. J., Kelley, S. T., et al. (2018). Gut microbial diversity in women with polycystic ovary syndrome correlates with hyperandrogenism. *J. Clin. Endocrinol. Metab.* 103, 1502–1511. doi: 10.1210/jc.2017-02153

Torres, P. J., Ho, B. S., Arroyo, P., Sau, L., Chen, A., Kelley, S. T., et al. (2019). Exposure to a healthy gut microbiome protects against reproductive and metabolic dysregulation in a PCOS mouse model. *Endocrinology* 160, 1193–1204. doi: 10.1210/en.2019-00050

Vander, H., and Prabha, V. (2019). Colonization of mouse vagina with Pseudomonas aeruginosa: A plausible explanation for infertility. *Microb. Pathog.* 134, 103602. doi: 10.1016/j.micpath.2019.103602

Witchel, S. F., Azziz, R., and Oberfield, S. E. (2022). History of polycystic ovary syndrome, premature adrenarche, and hyperandrogenism in pediatric endocrinology. *Horm. Res. Paediat.* 95, 557–567. doi: 10.1159/000526722

Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* 20, 1–13. doi: 10.1186/s13059-019-1891-0

Wright, M. L. (2021). Exploring the vaginal bacteriophage frontier. *BJOG* 128, 983. doi: 10.1111/1471-0528.16539

Xue, B., Sechi, L. A., and Kelvin, D. J. (2020). Human endogenous retrovirus K (HML-2) in health and disease. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.01690

Zeng, X., Xie, Y.-j., Liu, Y.-t., Long, S.-l., and Mo, Z.c. (2020). Polycystic ovarian syndrome: Correlation between hyperandrogenism, insulin resistance and obesity. *Clinica Chimica Acta* 502, 214–221. doi: 10.1016/j.cca.2019.11.003

Zhang, F., Ma, T., Cui, P., Tamadon, A., He, S., Huo, C., et al. (2019). Diversity of the gut microbiota in dihydrotestosterone-induced PCOS rats and the pharmacologic effects of diane-35, probiotics, and berberine. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.00175

Zhang, Y., Sharma, S., Tom, L., Liao, Y.-T., and Wu, V. C. H. (2023). Gut phageome—An insight into the role and impact of gut microbiome and their correlation with mammal health and diseases. *Microorganisms* 11, 1–21. doi: 10.3390/microorganisms11102454

frontiers | Frontiers in Microbiomes

TYPE Original Research
PUBLISHED 04 November 2024
DOI 10.3389/frmbi.2024.1399416

OPEN ACCESS

EDITED BY
Rebecca C. Mueller,
Agricultural Research Service (USDA), United States

REVIEWED BY
Neslihan Taş,
Berkeley Lab (DOE), United States
Rodrigo Gouvea Taketani,
Rothamsted Research, United Kingdom

*CORRESPONDENCE
Mia Rose Maltz
mia.maltz@uconn.edu

†PRESENT ADDRESS
Michala L. Phillips,
Pacific Island Ecosystems Research Center, United States Geological Survey, Hawaii National Park, HI, United States

RECEIVED 11 March 2024
ACCEPTED 22 August 2024
PUBLISHED 04 November 2024

CITATION
Maltz MR, Allen MF, Phillips ML, Hernandez RR, Shulman HB, Freund L, Andrews LV, Botthoff JK and Aronson EL (2024) Microbial community structure in recovering forests of Mount St. Helens. *Front. Microbiomes* 3:1399416. doi: 10.3389/frmbi.2024.1399416

Microbial community structure in recovering forests of Mount St. Helens

Mia Rose Maltz[1,2*], Michael F. Allen[2,3], Michala L. Phillips[4†], Rebecca R. Hernandez[5], Hannah B. Shulman[6], Linton Freund[3,7], Lela V. Andrews[8], Jon K. Botthoff[2] and Emma L. Aronson[2,3]

[1]Plant Science and Landscape Architecture, University of Connecticut, Storrs, CT, United States, [2]Center for Conservation Biology, University of California, Riverside, Riverside, CA, United States, [3]Microbiology and Plant Pathology, University of California, Riverside, CA, United States, [4]Department of Botany and Plant Sciences, University of California, Riverside, Riverside, CA, United States, [5]Land, Air, and Water, Resources Department, University of California, Davis, Davis, CA, United States, [6]Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, United States, [7]Genetics, Genomics, and Bioinformatics Program, University of California, Riverside, Riverside, CA, United States, [8]Tecan Genomics, Redwood City, CA, United States

Introduction: The 1980 eruption of Mount St. Helens had devastating effects above and belowground in forested montane ecosystems, including the burial and destruction of soil microbes. Soil microbial propagules and legacies in recovering ecosystems are important for determining post-disturbance successional trajectories. Soil microorganisms regulate nutrient cycling, interact with many other organisms, and therefore may support successional pathways and complementary ecosystem functions, even in harsh conditions. Historic forest management methods, such as old-growth and clearcut regimes, and locations of historic short-term gopher enclosures (*Thomomys talpoides*), to evaluate community response to forest management practices and to examine vectors for dispersing microbial consortia to the surface of the volcanic landscape. These biotic interactions may have primed ecological succession in the volcanic landscape, specifically Bear Meadow and the Pumice Plain, by creating microsite conditions conducive to primary succession and plant establishment.

Methods and results: Using molecular techniques, we examined bacterial, fungal, and AMF communities to determine how these variables affected microbial communities and soil properties. We found that bacterial/archaeal 16S, fungal ITS2, and AMF SSU community composition varied among forestry practices and across sites with long-term lupine plots and gopher enclosures. The findings also related to detected differences in C and N concentrations and ratios in soil from our study sites. Fungal communities from previously clearcut locations were less diverse than in gopher plots within the Pumice Plain. Yet, clearcut meadows harbored fewer ancestral AM fungal taxa than were found within the old-growth forest.

Discussion: By investigating both forestry practices and mammals in microbial dispersal, we evaluated how these interactions may have promoted revegetation

and ecological succession within the Pumice Plains of Mount St. Helens. In addition to providing evidence about how dispersal vectors and forest structure influence post-eruption soil microbiomes, this project also informs research and management communities about belowground processes and microbial functional traits in facilitating succession and ecosystem function.

KEYWORDS

succession, bacteria, fungi, AM fungi, Mount St. Helens, community assembly, legacy effects, diversity

1 Introduction

Natural disasters such as volcanic eruptions can potentially change change Earth's ecological systems. Volcanic ash, lava, and debris flows have transformative and ecosystem-level effects, such as burying vegetation and destroying habitat. Volcanic eruptions may also dramatically affect atmospheric conditions and precipitation, far from where a blast originated, with extensive implications for both natural landscapes and human systems, such as in agriculture, industry, art, and religion (Stothers, 1984). Patchy mosaics of new soils emerge after volcanic eruptions, yielding shifts in substrate age, soil development, and primary successional processes (Vitousek, 2004). As volcanic eruptions resurface landscapes, the effects of the blast reverberate belowground as soil microorganisms may perish, aggregate within unaltered areas (as refugia), or disassociate with their hosts (Jackson et al., 2022). Belowground ecosystem functions likely diminish following eruptions, which may inhibit microbial activity and reestablishment (Hernández Garcia et al., 2020).

Thus, eruptions indirectly hinder plant reestablishment, nutrient cycling, and both provision and regulation of ecosystem services (Banwart et al., 2019). In these novel environments, microorganisms struggle to reassemble beneath the ashfall and pyroclastic flow (Dale et al., 2005).

The U.S. State of Washington is home to Mount St. Helens, a volcano known for numerous historic eruptions. During volcanic eruptions, fragmented material known as tephra is produced, ejected, and then deposited within the blast zone (Dale et al., 2005). At Mount St. Helens, tephra consisting of smaller particles (< 4 mm ash) to larger blocks (> 32 mm angular stones) is intercalated by buried forest floor layers in profiles that vary in organic matter content and nutrient concentrations (Allen et al., 2005). Mount St. Helens's cataclysmic explosion in May 1980 destroyed more than 350 km^2 of coniferous forest and montane habitats in the Cascades Mountains of the western United States, influencing not only aboveground vegetation and habitat for wildlife, but also free-living and host-associated soil microbial communities. The vertical stratification of soil horizons provides different resources for newly developing plant roots and rhizosphere microorganisms (Phillips and Lorz, 2008). Epipedons, subsurface edaphic layers, and soil aggregates engender spatial arrangements for roots or microbes to obtain limited resources, locate refugia, or proliferate amid the regolith. Pyroclastic flows created new environments from the crater to the Pumice Plain as remnant rubble from the May 1980 eruption from the crater to the Pumice Plain, new environments were created by pyroclastic flows as remnant rubble from the May 1980 eruption. These environments contained buried organic material and layers of tephra, with much of the blastfall zone covered in thick layers of snowpack, ash, and fallout from the blast (Findley, 1981).

Volcanic soils in forested ecosystems are exposed to high—often transformative—temperatures, as soil carbon (C) and nitrogen (N) transformation rates shift, leading to less N turnover, more N immobilization, and greater increases in C storage than are found in nonvolcanic soils (Yokobe et al., 2020). Post-eruption soils also tend to be more acidic, with more gas emissions (i.e., methane, CH_4; carbon dioxide, CO_2; hydrogen sulfide, H_2S; hydrogen; H_2 gases), giving rise to novel microbial communities with varied nutritional modes, such as those that are chemolithoautotrophic (Picone et al., 2020).

Although many lifeforms perished as the eruption transformed or translocated sterile tephra, some remnant soil organisms survived (Allen et al., 1984, 2018). Besides those buried under the ashfall, immigrant or ruderal taxa may have sought out shelter and found refugia in stark environments and the Pumice Plain's sterile tephra, which initially consisted of pumice emerging from the pyroclastic flow. The pumice ranged from 10 – 20 mm in diameter and contained no measurable carbon (C) or nitrogen (N) (Fruchter et al., 1980; Baross et al., 1982). In 1981, scattered individuals of *Lupinus lepidus* (i.e., lupine) managed to establish patches (Allen and MacMahon, 1988; Allen et al., 1992, 2005) and likely associated with ruderal microorganisms, such as endosymbiotic rhizobial bacteria that are capable of atmospheric nitrogen fixation within the rhizosphere (Vuong et al., 2017; Coba de la Pena et al., 2018).

Lupine also served another important role in successional dynamics: as a prime food source for pocket gophers (*Thomomys talpoides*, Richardson), thus, inviting these ecosystem engineers (Reichman and Seabloom, 2002) to seed the Pumice Plain with biotic propagules. Gophers are known as "fossorial species," meaning "hole diggers." Findings from a fundamental study

recounted by Logan (2007) showed that a single gopher can move 227 kg of soil per month, with gopher populations translocating 38,000 kg of soil per acre per year. Given that the disturbance intensity at Mount St. Helens differentially impacted pocket gophers, these parameters affected successional processes within both patchy mosaics and disturbance intensity gradients (Crisafulli et al., 2005; Dale et al., 2005). These patches subsequently metastasized as biotic communities grew proximally and distally (Franklin, 2005; Dale et al., 2005) within the blastfall zone. While the type, abundance, and distribution of biotic and abiotic legacies corresponded with this complex post-eruption landscape, a confluence of chemical, soil, and extant—or dead—biotic legacies may obfuscate patterns arising from the eruption, endogenous inputs, or effects of distant propagule sources.

At Mount St. Helens, the time required for landscapes to recover varied across systems and by the pre-eruption conductions. Spatial heterogeneity, responses to disturbance, and widespread biotic dispersal influenced the speed and variability of terrestrial and aquatic systemic response trajectories. After the 1980 eruption, lakes and streams became enriched with nutrients and resumed pre-eruption productivity levels far sooner than terrestrial ecosystems, as forest floors were blanketed by nutrient-poor tephra in response to dominant tree communities being extinguished. Geophysical forces, wildlife, and human activities contributed to the multifaceted dynamics during succession, including facilitation, inhibition, tolerance, and relay succession, with communities changing deterministically from pioneer to climax taxa (McCook, 1994; Pulsford et al., 2016). Yet researchers at Mount St. Helens also observed contemporaneous occurrences of early-seral and late-seral species, herbivory, predation, and mutualisms. Furthermore, ruderal and established taxa altered environmental parameters, such as by changing C and N concentrations to emulate more livable habitats (Halvorson et al., 2005).

In the clearcut, the ashfall covered the intact meadow soil to a depth of 12 cm. Because a deep snowpack was still in place in May, soil organisms, such as the pocket gopher, survived the eruption (Allen et al., 1984, 2018). The gophers burrowed back to the surface in spring, bringing old soil, reducing bulk density, and mixing in different soil layers and "stored topsoil", with the ash mixture at approximately 1:1 soil to ash (Allen et al., 1984). These "islands" of inoculated soil included a diverse soil biota, mycorrhizal fungi, and nutrients (Allen et al., 2005). In 1980 (MacMahon and Warner, 1984), the organic C of the tephra was 2.2g.kg^{-1} and nitrate N 0.9mg.kg^{-1}. However, the buried soil was highly organic (C was 39.6g.kg^{-1} with NO_3-N of 10.9mg.kg^{-1}). By 2000 (Allen et al., 2005), the organic C and N in the tephra remained low (1.46g.kg^{-1} and 10mg.kg^{-1}, respectively), whereas the buried soil C averaged 17.5g.kg^{-1} and N averaged 390mg.kg^{-1}). By 2015, the forb and shrub vegetation were dense.

The Pumice Plain site was within the pyroclastic flow consisting of sterile tephra, greater than 20m deep, upslope from Spirit Lake (Figure 1). Initially, the predominant feature was the texture of the growth medium, composed of marble- to golf ball-sized stones of pumice, with no measurable C or N (Fruchter et al., 1980; Baross et al., 1982). A few widely scattered individuals of *Lupinus lepidus* Dougl. had established during 1981. In 1982, these had reproduced, forming small patches. One of these became the focus of our research (Allen and MacMahon, 1988; Allen et al., 1992, 2005, 2018). In 1982, an individual pocket gopher (*Thomomys talpoides* [Richardson]) from a nearby clearcut was placed in a 1m^2 enclosure

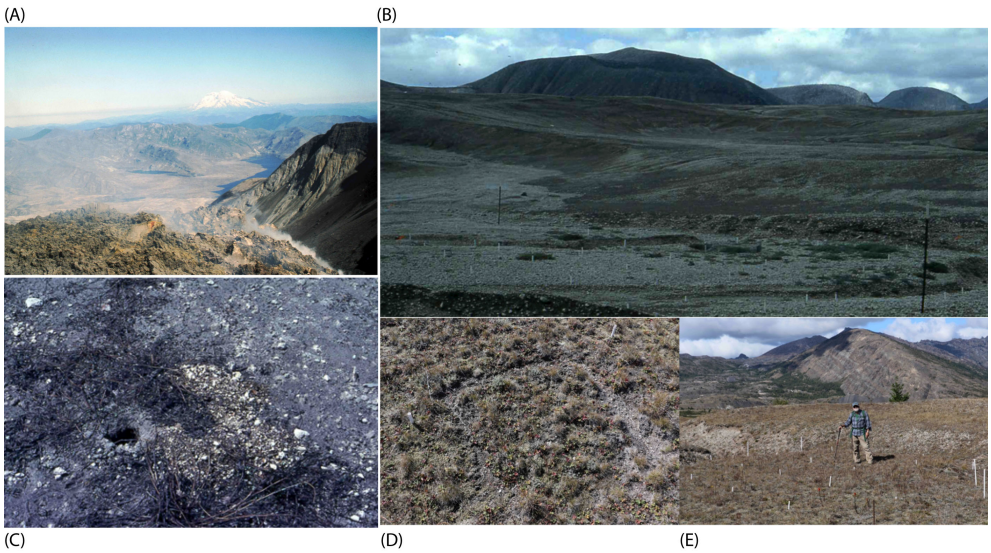


FIGURE 1
Overview of the Mount St. Helens study area. Aerial photograph above the dome in the crater, taken from a helicopter, above Mount St. Helens in 1982 **(A)**. Ashfall extends over Bear Meadow, with the Pumice Plain in the foreground and Mount Rainier in the background. (photograph taken by M.F. Allen). Pumice Plain from 1983 plots shown in **(B)**. **(C)** shows gopher activity and residual ash with old soil, transported by gophers in the Bear Meadow clearcut site. Photographic documentation shows an overview of the vegetation plots **(D)** in 1982 in the Pumice Plain, which corresponded to the year that gopher enclosures were set up following the pyroclastic flow, amidst the residual ash and tephra. Recent site visit in 2018 to historic gopher enclosure plots (M.F. Allen, shown in the foreground) within the Pumice Plain **(E)**.

around a single *L. lepidus* individual for 24 hours (Allen and MacMahon, 1988). By 1987, in the interspace total C was only 590 mg.kg^{-1} and N 17 mg.kg^{-1} (Halvorson et al., 2005). Under the plant, though, soil was already forming, and total C was 840mg.kg^{-1} and total N53mg.kg^{-1}. By the year 2000, total C was 1.28 g.kg^{-1}, and total N 114 mg.kg^{-1} in the interspace and C had increased to 4.6 g.kg^{-1} and total N to 384 mg.kg^{-1}. Other characteristics can be found in Halvorson and colleagues (Halvorson et al., 2005). By 2015, the individual *L. lepidus* was long dead and gone, but a diverse suite of plants occupied the Pumice Plain extending up the mountain. Gophers had migrated back onto and up the Pumice Plain (Allen et al., 2018).

This seminal backdrop for studying secondary succession, in part due to the magnitude of the 1980 event, conveys valuable biological information to contextualize meaningful ecological issues, commensurate with thorough, protracted, and robust detection and monitoring of both the *types* and *levels* of remnant biological legacies herein (Franklin; Dale et al., 2005). As forest practices or gopher enclosures (as per Allen and MacMahon, 1988) overlay manipulative or experimental approaches, we are more equipped to disentangle the rules governing community assembly in this system subjected to natural disasters. The recent advent of cutting-edge technologies may facilitate fine-scale investigations of how top-down (e.g., clearcuts) or bottom-up (e.g., discrete or assisted microbial migration) factors influence long-term pedological trajectories and biotic succession of vegetation and macrofauna. Next-generation tools (Zimmerman et al., 2014), such as -omics approaches (Quince et al., 2017) and targeted-amplicon sequencing of fungal or bacterial communities (Taylor et al., 2016), may now be leveraged to track the arrival—and subsequent encroachment—of free-living, mutualistic, or commensal microorganisms in substantially sterile, harsh environments (Allen et al., 1984). In the wake of this historic blast, microbial communities—or their assisted migration—may have played a role in altering long-term successional processes and community assembly (Allen and MacMahon, 1988; Allen et al., 1992). Using next-generation tools, we can evaluate how, or to what extent, these effects may have lingered or attenuated over decadal scales since eruption events.

Human activities that altered landscapes, such as conventional forestry practices or modifying of fish and game species, may have influenced rates and patterns of ecological responses to the eruption. The spread of nonnative species, changes in primary production, and vegetation type conversion may exert selective pressures on soil microbial communities and likewise affect food webs (including birds or invertebrates), biodiversity, and ecosystem multifunctionality (Wagg et al., 2014). Dove and Keeton (2015) suggest that forest management practices directly influence microhabitat characteristics, such as those critical to forest fungi's survival, which could indirectly affect decomposition and nutrient cycling. Forest structural complexity and functional diversity likewise affect these key soil processes. According to Grime's (1998) mass-ratio hypothesis, changes to dominant plant species, and rates of primary production, such as by clearcut forestry management practices, could feed back to affect soil C and biogeochemical cycling, which may have implications for habitat quality, stability, and resilience to disturbance.

This study examined the effects of pre-eruption forest management and post-eruption gopher additions on bacterial and fungal communities, including plant-root-associated AM fungal mutualists. We hypothesized that there would be greater fungal and bacterial richness in old-growth forests after the blast than in previously clear-cut sites, e.g., Bear Meadow. Taxa originating in old-growth forests may be more resilient to disturbance or natural disasters than those altered by clearcutting forest practices (Bowd et al., 2022; Delgado-Baquierzo et al., 2018). In areas laden with volcanic debris, we predicted that old-growth and clearcut forest management processes may have opposing effects on fungal and bacterial groups over decadal time scales and correspond to changes in C: N ratio in soils. These conjectural effects may be related to the relative abundance of particular microbial groups, the functions performed by microbes, or litter quality from the surrounding vegetation within these contrasting landscapes.

We hypothesized that microbial and fungal community composition would vary in long-term lupine plots with gopher additions in the Pumice Plain as compared to those without historic gopher activity, and that we would find more mycorrhizal fungal mutualists in lupine plots containing historic gopher enclosures. We further hypothesized that the presence of animal dispersal vectors, such as the pocket gophers, will influence mycorrhizal guilds, community composition, and soil biogeochemical processes because of preferential foraging and other behaviors, including bioturbation, elimination (e.g., fecal deposits), and soil translocation.

2 Methods

2.1 Study site description

Prior to the May 18, 1980, eruption, the vegetation of our study sites on Mount St. Helens (46°16'N, 121°53'W) was primarily comprised of a Douglas fir forest interspersed with clearcuts (Allen et al., 1984, 1992). The area has predominately cool, wet winters and mild, dry summers, with heavy snowpack, often greater than 3 m. For this study, we focused on two locations: the Pumice Plain long-term lupine plot and Bear Meadow, a high-ashfall area in the path of the 1980 eruption.

The eruption itself created multiple zones or layers of disturbance and mortality, ranging from the crater, which is still active, to the Pumice Plain, the pyroclastic flow running out to Spirit Lake, down to ash deposition, from about 10 km to the northward extending as far north as Mount Rainier, and as far east as Lincoln, Nebraska. An overview is provided in Figure 1, Supplementary Figure S1, and in many publications (Allen et al., 1984; Dale et al., 2005; Allen et al., 2018). Additionally, Bear Meadow was a high-deposition area (Supplementary Figure S1). The Bear Meadow site consisted of two parts: a clearcut that became a meadow by 1980, and an adjacent old-growth Douglas fir forest.

2.2 Soil collection

On July 28 and 29, 2014, we collected soils from the Pumice Plain long-term lupine plots and from Bear Meadow, the high-ashfall area, which was in the path of the 1980 eruption. Three 9 cm diameter × 12 cm deep soil cores were collected and composited for each sample, for a total of 12 individual cores from the Pumice Plain (46°14'43.1"N 122° 10'50.9"W), consisting of six composite samples from plots with historic Gopher enclosures and six composite samples from paired No Gopher control plots, also from the Pumice Plain. From Bear Meadow, 26 composite soil cores were collected across the clearcut (46° 15'52.4"N 122°04'46.9"W) and old-growth (46°18'35.6"N 122° 01'18.3"W) forests, with 11 samples from the clearcut and 15 from the old-growth. Surface litter and debris were removed before coring. Samples were immediately placed on dry ice for transportation to the laboratory at UC Riverside (Riverside, CA, USA). Soils were used for extraction and targeted amplicon sequencing, as well as to measure C and N concentrations.

2.3 Molecular analyses

We extracted microbial DNA from frozen soil (Pall, NY, USA). After collection in the field, samples were kept frozen in a -20°C freezer at the University of California, Riverside, until subsequent DNA extraction using a MoBio PowerSoil Power Lyzer DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA; Qiagen Inc. Valencia, CA, USA), followed by cell lysis directly on a MoBio PowerLyzer tissue disrupter (Qiagen Inc., Valencia, CA, USA), as per the manufacturer's instructions. Prior to tissue lysis, we weighed 0.25g of soil from each sample, and subsequently homogenized these aliquots, following a modified protocol from the PowerSoil PowerLyzer DNA Extraction Kit protocol (Protocols.io link; MO BIO Laboratories, Carlsbad, CA, USA).

Extracted DNA was further purified with a PEG and carboxylated magnetic bead solution, using a 1:1 ratio of sample-to-bead solution as described in Rohland and Reich (2012). Purified DNA extracts were quantified using a NanoDrop 2000/2000c UV-Vis spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and concentrations normalized to 10 ng/µl.

Normalized and purified DNA was used as a template to amplify each of three different taxonomically informative loci by polymerase chain reaction (PCR). We used a tailed primer strategy for sequencing library construction to first amplify each target locus with a subsequent amplification to add molecular indexes and Illumina flowcell adapter sequences (Alvarado et al., 2018). Bacteria and Archaea were characterized according to the v4 region of the 16S rRNA gene using the 515F and 806R primers (Caporaso et al., 2010a). General fungal communities were characterized according to the Internal Transcribed Spacer 2 rRNA region, with primer sets ITS4-fun and 5.8S-fun (Taylor et al., 2016); finally, we used AMF-specific primers WANDA (Dumbrell et al., 2011) and AML2 (Lee et al., 2008). Individual sequencing libraries were quantified by PicoGreen (Thermo Fisher Scientific) fluorescence and pooled for sequencing.

Sequencing was carried out on a MiSeq Desktop Sequencer (Illumina Inc, San Diego CA) running in paired end 2×300 mode at the University of California, Riverside Institute for Integrative Genome Biology. Sequences were submitted to the National Center for Biotechnology Information Sequence Read Archive under BioProject PRJNA1124544. Total sequencing depths overall for 16S were 4097169 reads; for ITS2, read depth was 1728368; for SSU, depth was 1601559 reads.

2.4 Bioinformatics

The fungal sequences were demultiplexed in QIIME 1.9.1 (Caporaso et al., 2010b; Caporaso et al. 2012), and quality control measures were employed (Edgar and Flyvbjerg, 2015) in USEARCH. We characterized the percentage of reads passing through filters across different read lengths; we then determined the appropriate lengths to trim and still retain high-quality reads before per-base quality plummeted. After trimming, we ran trimmed sequences through FastQC to compare reports and to ensure that adapters were absent and per-base and per-sequence qualities had improved. Since there was at least a 12 base pair overlap, we used paired ends of both forward and reverse reads for both the V4 region of the 16S rRNA gene for bacteria and archaea and from the fungal ITS2 region to infer amplicon sequence variants (ASV) assignments from our sequences using the Divisive Amplicon Denoising Algorithm (DADA2; dada). ASVs were assigned based on 99% sequence identity (Prodan et al., 2020; Callahan et al., 2017).

2.4.1 Arbuscular mycorrhizal fungal SSU

We filtered demultiplexed files from the SSU locus using multiple_split_libraries_fastq.py command in QIIME 1.9.1 with Q (Phred) scores of 20 (q = 19) as a quality control parameter; we did not allow any low-quality base calls (r = 0), and we only retained reads possessing 95% of initial sequence length following quality truncation (p = 0.95). We used VSEARCH (Rognes et al., 2016) in uchime_denovo mode to screen for chimeras. We used Swarm (Mahe et al., 2014) with a d4 resolution for picking operational taxonomic units (i.e., OTUs) for our SSU data. This d4 resolution for local clustering threshold collapsed sequences with fewer than or equal to 4 differences into a single representative OTU, as long as it passed our stringent quality filtering threshold of q20. We used BLAST to assign taxonomy, with an e-value of 0.001, via the MaarjAM (Opik et al., 2010) database. Prior to analysis, we truncated the reference databases to include only the region of interest, as well as to limit any spurious results. We used the metagenomeSeq package of Bioconductor (Ihaka and Gentleman, 1996) in R (R: Development core team, 2004) to normalize our OTUs using cumulative sum scaling (CSS-normalization) prior to further analyses (CSS normalization attempts to avoid biases associated with marker gene surveys, which could result from uneven sequencing depth). Read counts are rescaled against a quantile that was determined by assessing the count deviations for each sample, as compared to the distribution of counts across all of our other samples.

2.4.2 Bacterial 16S and fungal ITS2

We separated our sequence files for the 16S and the ITS2 regions before the Illumina MiSeq data was processed interactively on the High Performance Computing Cluster at UC Riverside. We conducted further quality control (Kassambara, 2018) and trimming in DADA2 and used the filterAndTrim command to filter reads based on read quality, read length, number of unknown bases, expected errors, and matches to the PhiX genome, as per Freund (2023; 10.5281/zenodo.5794553). The region lengths vary for ITS2; therefore, truncating based solely on specific lengths may be unreliable. We used paired-end ITS2 reads when assigning taxonomic features to the ASVs via DADA2, as detailed in Freund (Lee, 2019) (2023; v1.0.1 10.5281/zenodo.8264886). The DADA2 algorithm as per Callahan et al. (2016) uses a parametric error model with both inference and estimation as it determines error rates per sample. We constructed our feature tables for both loci (i.e., ASV tables), removed chimeric sequences, and assigned taxonomy to our assigned ASVs using the Silva database and the Ribosomal Database Project for 16S, the UNITE database for fungal ITS2, and the MaarjAM data base for SSU.

To examine the AMF community's (18S) responses within a fungal functional group framework, we assigned families of Glomeromycotina to AMF functional groups (rhizophilic, edaphophilic, and ancestral) using AMF resource allocation patterns defined in previous studies (Phillips et al., 2019; Weber et al., 2019). We did not include sequences assigned to the taxon *Geosiphon pyriformis*, as reads reportedly identified as *G. pyriformis* are not considered as rhizophilic, ancestral, or edaphophilic AMF. This functional group framework is based on studies demonstrating that AMF are host-specific and exhibit diverse resource allocation patterns. Arbuscular mycorrhizal fungal families with high allocation to extraradical hyphae, including Gigasporaceae and Diversisporaceae, are members of the guild "edaphophilic"; those with high allocation to root colonization (i.e., Glomeraceae, Claroideoglomeraceae, and Paraglomeraceae) are characterized as belonging to the "rhizophilic" guild; families that allocate lower AM fungal hyphae to either root colonization or soil foraging than the aforementioned guilds are known as "ancestral" AMF families, and include Ambisporaceae, Archaeosporaceae, and Acaulosporaceae. Previous studies show that fungal taxa from the rhizophilic AMF functional group are important for protecting their host plant from pathogen colonization, while edaphophilic AM fungi putatively improve plant nutrient uptake.

To examine fungal and bacterial microbiomes, as well as the relationship between these communities from Bear Meadow and the long-term lupine plot, sequences were processed with DADA2, with components modified from Quantitative Insights Into Microbial Ecology 2 (QIIME2; Bolyen et al., 2019). This approach was used to determine the relationship between bacterial microbiome communities and host diet, rearing, and sterility variables. We removed low-quality and chimeric sequences and computed core microbiomes in QIIME; we amplified sterile PCR-grade water, as a negative control, which was processed alongside the DNA samples. After samples were extracted, amplified, and sequenced, any OTUs or ASVs that were present in the negative controls were removed from downstream analyses.

To account for uneven sequencing depth samples' bacterial and fungal alpha diversity was determined using rarefied ASV counts. For 16S, we used a rarefied ASV table to 17644 reads per sample; for ITS2 used a rarefied ASV table with 20069 reads per sample. Distance matrices were constructed using Bray-Curtis metrics through the "diversity()" function in the vegan package in R, using the Shannon-Wiener index. Species richness was calculated using raw ASV counts via the "specnumber()" function in the vegan package for R. Alpha diversity and species richness across treatment conditions and treatment sources were compared, respectively, using Wilcoxon tests, performed with the "compare_means" function in the "ggpubr" package for R, as well as t.test commands through the base R package. We used the "total" option in the "decostand()" function in the vegan package in R, upstream of compositional analyses, which relativizes varying ASV counts by dividing row total from each cell in the row, standardizing all of the counts to each other within each respective sample (Borcard et al., 2011; Legendre and Legendre, 2012).

To determine if fungal or bacterial composition differed among samples, Bray-Curtis distance matrices were used to compare community samples and to visualize and explain differences among microbial communities. We used nonmetric multidimensional scaling (NMDS) plots of the Bray Curtis distances and PERMANOVA analyses of microbial community data using the "adonis()" function in the vegan package of R. NMDS stress values (MDS; "metaMDS()" command, package vegan) were reported for fungi and bacteria. PERMANOVA was used to compare fungal bacterial community structures across all forest management, assisted migration, or control groups based on the ASV composition. To summarize broad-level taxonomic patterns, we parsed our count matrix by taxonomy to view the relative abundance of major taxa or microbial richness by categories; we used a 1.0% similarity threshold to explore the relationships among the relative abundances of abundant taxa. Percent N, percent C, and C:N ratios were used, along with Shannon diversity and species richness (Supplementary Figure S2) to further compare across treatment sources and conditions.

3 Results

3.1 Land use histories, soil chemistry, and microbial communities

We found that microbial communities varied among forestry practices (n=26) and between long-term lupine sites with and without gophers (n=12). These findings related to detected differences in carbon (C) or nitrogen (N) concentrations in soil or tephra samples across our study sites (Supplementary Table 1). Contrary to our first hypothesis, our data suggests that old-growth forests did not harbor greater microbial diversity or taxa richness than clearcut forests, which led us to reject that hypothesis. Our

second hypothesis, that microbial and fungal community composition would vary across different forestry regimes, was supported and likewise correlated to higher C:N in old-growth, as compared to clearcut meadows in Bear Meadow forests. Although our methods did not allow us to characterize the successional trajectory, we were able to examine differences in multiple components of the soil and microbial community within a post-disturbance context as related to microbial rescue (Mueller et al., 2020; Shade, 2023) and land use histories.

3.2 Soil chemistry

Carbon (C) and nitrogen (N) levels varied across land use histories and gopher activity in the Pumice Plain. Samples recovered from lupine plots in the Pumice Plain had lower C ($p = 0.02$) and N ($p = 0.04$) concentrations than samples from the surrounding forests (Supplementary Table 1). Lupine plots without gophers were more depleted in C ($p = 0.03$), with lower C:N ($p = 0.04$), than was detected in lupine plots with historic gopher activity. Nitrogen concentrations in soils from lupine plots with gophers were higher than the percentage of N from soils from lupine plots without historic gopher activity ($p = 0.046$; at alpha value threshold, corresponding to the threshold between significant and *marginally* significant findings).

Soils from old-growth forests have higher concentrations of C ($p = 0.01$) and N ($p = 0.01$) than were found in historically clearcut forests; with C:N significantly lower in clearcut soils than in those from old-growth forests ($p < 0.001$).

3.3 Bacterial and archaeal taxa communities

Counterintuitively, bacterial and archaeal taxa richness (i.e., microbial richness) and alpha diversity (i.e., microbial diversity) in Bear meadow ($n = 26$) were higher in clearcut soils than in soils from old-growth forests (richness $p < 0.001$; diversity $p < 0.001$). Indeed, these microbial communities from clearcuts were more diverse and characterized by higher bacterial and archaeal taxa richness than those from remnant old-growth forests. Overall, microbial diversity was higher in the long-term lupine plots in the Pumice Plain ($n = 12$) than was found in the surrounding forests ($p = 0.03$), yet microbial taxa richness was nearly equivalent (i.e., marginally higher; $p = 0.045$) in lupine plots to the nearby forested locations.

Although microbial richness and diversity were equivalent in Pumice Plain plots with or without historical gopher activity (richness, $p = 0.52$; diversity, $p = 0.22$), Pumice Plain lupine plots that contained historical gophers harbored significantly greater microbial taxa richness ($p = 0.03$) and diversity ($p < 0.001$) than were found in the surrounding old-growth forests from Bear Meadow. Microbial diversity in Pumice Plain plots without historic gopher activity was equivalent to levels in clearcut forests ($p = 0.014$). Yet we found higher microbial diversity in these lupine plots than in old-growth forests ($p=0.002$). However, microbial richness in lupine plots with and without historical gopher activity was equivalent to richness detected in either old-growth ($p=0.10$) or clearcut forests ($p=0.10$). Additionally, we did not detect any clinal differences with respect to richness or diversity, as deeper layers below the eruption had similar microbial taxa richness ($p = 0.32$) and diversity ($p = 0.86$) as the top layers in the profile with these old-growth forests.

The composition of microbial communities in the lupine plots differed significantly from that of the surrounding forests ($p < 0.001$), with old-growth microbial communities differing from the clearcut forests ($p = 0.002$; Figure 2). PERMANOVA was used to compare bacterial community structures across all forest management (MDS stress value 0.132), assisted migration (MDS stress value 0.164), or control groups based on the ASV composition), as visualized via NMDS. Taxa from Acidomicrobiia and Planctomycetes were more common in old-growth than in clearcut forests, with taxa from Gammaproteobacteria more commonly found in clearcut forests than were detected in old-growth forests within this system.

PERMANOVA analyses revealed that microbial and fungal community composition correlated significantly with soil C:N ($p < 0.001$), as well as the percent C ($p < 0.001$) and N ($p < 0.001$) of soil samples. Additionally, microbial community composition varied by layer in the profile or regolith, as microbial community structure from the top layer of soil in the Bear Meadow old-growth forests differed significantly from those in the layers below ($p = 0.02$) the previous eruption. Additionally, the composition of microbial communities from the lupine plots with gophers differed significantly from the composition of those without gophers ($p = 0.01$).

3.4 Fungal communities

Fungal taxa richness (i.e., fungal richness) and alpha diversity (i.e., fungal diversity) in clearcut soils were equivalent to soils from old-growth forests (richness $p = 0.06$; diversity $p = 0.05$; Supplementary Figure S2). Overall, fungal diversity was higher in the lupine plots in the Pumice Plain than was found in the surrounding forests ($p = 0.01$); yet no differences were detected in fungal taxa richness in lupine plots, as compared to samples from the surrounding forested locations.

Fungal diversity was significantly greater in lupine plots with gophers than in clearcut forests ($p < 0.02$); however, fungal richness was similar across these same conditions ($p = 0.10$). Fungal richness was equivalent across lupine plots with gophers and the surrounding old-growth forest (richness, $p = 0.38$). Although lupine plots without gophers also harbored more diverse fungal communities than were found in old-growth forests ($p < 0.001$), fungal taxa richness was equivalent in old-growth forests as compared to those in the lupine plots without historic gopher activity.

Fungal composition varied significantly by forest management strategies, with clearcut forests characterized by different fungal communities than those found in the surrounding old-growth forests ($p < 0.001$; Figure 1; MDS stress value 0.127). Likewise, we

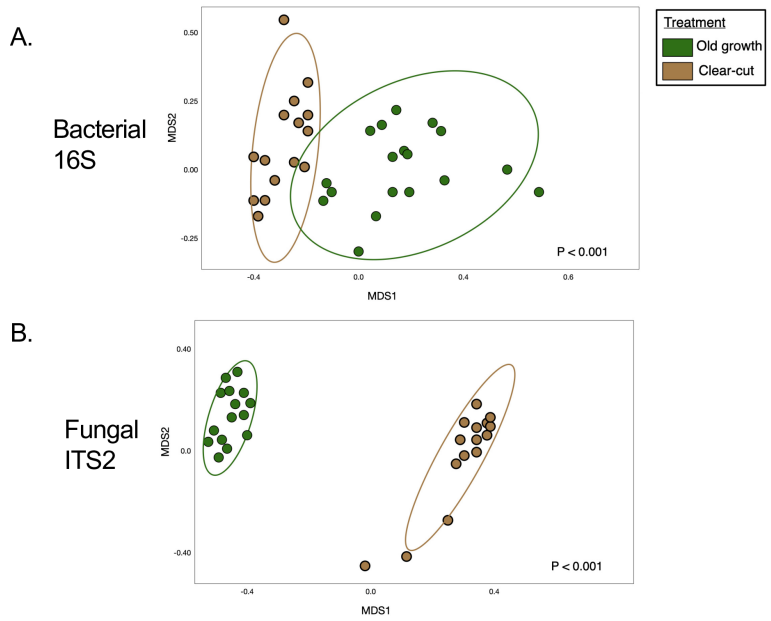


FIGURE 2
Microbial community composition shifts by land use histories in the forests surrounding Mount St. Helens. Non-metric multidimensional scaling (NMDS) based on bray-curtis distance matrices of bacterial **(A)**, and fungal **(B)**, communities revealed from targeted amplicon sequencing of 16S marker genes and fungal ITS2 region, illustrated by structural shifts by land use category in clearcut vs. old-growth forests **(A, B)** and lupine plots in the pumice plain. NMDS stress values (MDS; 'metaMDS()' command, package vegan) were 0.1272 for bacteria and 0.1031 for fungi. PERMANOVA analyses illustrated by structural shifts in the ordination by land use category in clearcut vs. old-growth forests **(A, B)**

detected structural differences in fungal communities in gopher plots as compared to "no-gopher" lupine plots ($p = 0.01$; assisted migration MDS stress value 0.150).

Fungal community composition likewise varied significantly by the percentage of N ($p < 0.001$), C ($p < 0.001$), and the C:N ($p < 0.001$) of soil or tephra samples. Moreover, fungal community composition in the old-growth forests of Bear Meadows differed with depth, with significantly different fungal communities found in the top layer of the old growth, as compared to the layers below ($p = 0.03$).

A majority of the fungi detected in this system belonged to the Ascomycota, with a greater proportion of the old-growth forest represented by Basidiomycota, and specifically Agaricomycetes, than were present in the clearcut forests. Mushroom-forming fungal genera, such as *Inocybe*, *Cortinarius*, and *Craterellus* sp., and ectomycorrhizal fungi, including *Piloderma* and *Wilcoxina* sp., were more commonly found in old-growth forests. Ascomycete fungi, such as the plant disease-forming *Paraphoma* and *Paraphaeosphaeria*, as well as plant-supporting *Mortierella* (Mucoromycotina) and basidiomycetous yeasts and jelly fungi, including *Solicoccozyma*, more often in the soil from clearcut forests. More taxa from the Glomeromycotina were found in lupine plots with gophers than were found in locations without historic gopher activity. These arbuscular mycorrhizal fungi from Glomeromycotina varied according to functional guilds, with greater abundance of "ancestral" (as per Powell et al., 2009; Weber et al., 2019) mycorrhizal fungi in old-growth forests than clearcut forests. Ancestral taxa include those from Archaeosporacacae, Ambisporacacae, Pacisporacacae (as per Varela-Cervero et al., 2015), and Acaulosporacacae fungal families. Additionally, in long-term lupine plots with gophers from the Pumice Plain, we detected a greater richness of "rhizophilic" mycorrhizal fungi (i.e., Families: Glomeraceae, Claroideoglomeraceae, and Paraglomeraceae; Hart and Reader, 2002; Varela-Cervero et al., 2016a, 2016; Figure 3). With greater richness of rhizophilic AM taxa found in the lupine plots with gophers, we had partial support for our hypotheses related to more mycorrhizal fungi—in this case, specifically taxa from the rhizophilic guild—being detected in these historic gopher-containing plots within the Pumice Plain.

4 Discussion

We found evidence supporting long-term differences in microbial communities from the legacy effects of land use and assisted migration (i.e., gopher dispersal vectors) following the 1980 eruption of Mount St. Helens. The transformative effects of volcanic eruptions on soil-forming factors and biotic communities in surrounding ecosystems beget the trajectory dynamics of post-disturbance recovery, especially as important plant, animal, and microbial taxa return to establish on previously sterile pumice and tephra. The structure and legacy of forests, including pre-eruption aboveground biomass and volcanic eruptions, likely influenced soil microbial communities (Chen et al., 2023). Our study showed that microbial and fungal communities from the high-ashfall area in the path of the 1980 eruption remained vertically stratified with regolith

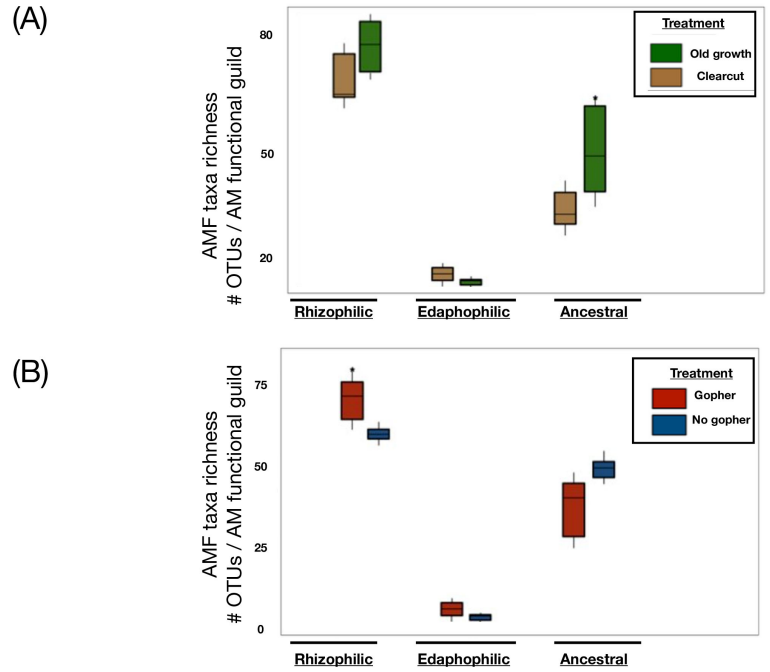


FIGURE 3
Arbuscular mycorrhizal fungal guilds shifted by dispersal vectors and land use histories in the forests surrounding Mount St. Helens. Arbuscular mycorrhizal composition shifts by historical land use and dispersal vectors in the forests surrounding Mount St. Helens. Richness of AMF guilds refers to OTU's (operational taxonomic unit, equivalent to molecular species concept) in fungal families represented by functional guilds (as per Weber et al., 2019) in these different forest conditions **(A)** or plots differing by gopher presence or absence **(B)**. Taxa and guild assignments revealed from targeted amplicon sequencing of 18S region and universal eukaryote Wanda primer set to assign families of Glomeromycotina to AMF functional guilds. Far left groupings in each panel have been assigned to rhizophilic AM families, middle groups were assigned to edaphophilic AMF, and the far right group in each panel encompass the taxonomic richness calculated from within ancestral AMF families. Taxa richness within these guilds were compared within AM functional group in plots of old growth vs. clearcut forests **(A)**, while the **(B)** depicts plots comparing AM taxa richness within guilds in either 'Gopher' or 'No gopher' plots.

in the profile, even decades after the 1980 blast; selective environmental pressure originating from multiple eruption events over time may have impacted the resultant soil microbiomes.

Biotic factors influenced long-term lupine plots, as gophers translocated both plant and microbial propagules around the Pumice Plain. As gophers naturally bioturbated, defecated, and mixed soil and tephra layers, these actions amended depauperate locations following the pyroclastic flow. In our study, we detected the presence of more types of root-associated mycorrhizal taxa (i.e., rhizophilic) in gopher plots than we found in long-term lupine plots without historic gopher activity. These mutualistic, rhizophilic AMF taxa may have promoted plant establishment within hostile terrain while also protecting plant roots from threats such as competition from root-associated pathogens (Borowicz, 2001; Powell et al., 2009; Weber et al., 2019). Plots with historic gopher activity harbored more diverse bacterial and fungal communities than the surrounding old-growth forests. We also found more diverse fungal communities in these long-term lupine gopher plots than in forests that were historically clearcut, prior to the 1980 eruption, nearby at Bear Meadow. Over 40 years since the blast, legacy or priority effects may have led to differences in fungal and microbial community composition across contrasting forestry regimes, or in pairwise treatments from within the long-term lupine plot at the Pumice Plain, upslope from Spirit Lake (Supplementary Figure S3). Indeed, the most dramatic compositional differences were detected in fungal communities from the previously clearcut forest, as compared to the old-growth forest surrounding Bear Meadow.

Ecologists are increasingly finding evidence that land use changes and forest management strategies could affect the long-term structure, functioning, and resilience of managed ecosystems. When lava cools, successional (Ni et al., 2023) forces initiate novel ecosystems and assemblages; invading organisms and immigrating propagules may be dispersed via aeolian processes. As subsets of biotic propagules survive within a patchy mosaic throughout the landscape (Allen et al., 1992), bacteria are some of the first organisms to succeed in primary succession, followed by other microorganisms and ruderal taxa that can either modify or tolerate these harsh conditions within the ashfall and blast zone (Abdulla, 2009). Some extremotolerant taxa could be some of the first pioneers in these novel environments (Ghosh et al., 2010). Some microorganisms exude proteins, such as those that break down harmful oxygen species or molecules that promote osmoregulation or extreme temperature tolerance, or find refugia in forming biofilms that protect bacterial cells physically, mechanically, or chemically; these microbial products ostensibly support their survival in toxic soils disturbed by volcanic eruptions (Memoli et al., 2018; Chen et al., 2023). As extreme conditions alter biodiversity and competitive interactions, or enhance abiotic or environmental filtering, this may both limit invasions by exotic species and preclude extant taxa from persisting or proliferating in

harsh environments. These factors likely exert selective pressure on the local community, which may hinder the establishment of immigrating taxa (Williamson and Harrison, 2002) and drive community assembly processes.

We found that bacterial community composition varied, based on pre-eruption forest management strategies or through altering propagule pressure via biotic introductions. While some members of biotic communities may have survived the blast, deep in profile, others—albeit smaller taxa—likely arrived via wet deposition as exogenous inputs from entrained ash or air currents (Maltz et al., 2022). Although bacteria vary in their tolerance of stressful environmental conditions, those that proliferated following the pyroclastic flow likely were equipped with a particular combination of traits that allowed them to persist in low-nutrient environments, such as traits related to carbohydrate synthesis and storage or the ability to form biofilms or spores (Brewer et al., 2019).

As the cataclysmic As the cataclysmic Mount St. Helens eruption destroyed thousands of hectares of forest and montane habitat, it blanketed the surrounding adjacent old-growth *Pseudotsuga menziesii* (i.e., Douglas fir) forest and clearcut in Bear Meadow with nearly 20 cm of ash. Remarkably, deeply packed snow protected remnant soil organisms that survived the eruption. Other taxa survived the blast by a combination of protection via rotten logs, soil, erosional forces, and topographic features (Andersen, 1982; Andersen and MacMahon, 1985; Allen et al., 2005). Animals, such as gophers, ants, mountain-dwelling pika (*Ochotona daurica*), and other taxa, served as vectors of dispersal. These ecosystem engineers amid the volcanic rubble were wont to move deep organic soil from below the ashfall to the surface, effectively mixing decayed organic matter with the ash and depositing nutrients via physical bioturbation.

Gophers translocated and deposited diverse soil microbiomes, seeds, and mycorrhizal spores from their fecal matter (Allen et al., 2005) into recipient ecosystems. Microbial communities were likely seeded by gophers and other remnant, immigrant, or ruderal taxa serving as dispersal vectors. This could have accelerated physical and biochemical weathering of the Andisol volcanic parent material with unconsolidated Andic Haplocryods soils and material (Web Soil Survey, Staff). Organic acids and residues from biotic introductions may also have enhanced the production of stable soil aggregates. Although measuring soil aggregation was beyond the scope of this study, we demonstrate that such interactions were also important in directly affecting microbial community composition and habitat, as well as the C:N of soil environments.

Pumice Plain plots with historic gopher activity had more C and a higher C:N than those locations without gopher interventions or activity. Additionally, N and C concentration and C:N were higher in old-growth forests than in the Bear Meadow clearcut, which was related to fungal and microbial community composition. In forest soils, soil microbes may play key roles in biotic immobilization of NH_4^+ to soil organic matter (Brookes et al., 1985), as microbially-derived organic N. Halvorson et al. (as per Dale et al., 2005) showed that prairie lupine at the pyroclastic flow site maintained higher nitrogenase activity, which corresponded with diurnal patterns of plant photosynthesis and heightened N-fixation along the successional trajectory in the Pumice Plain. Likewise, nitrate leaching may be related to microbial community activity, as C:N may influence the release of essential minerals, such as zinc and phosphorus, as well as N mineralization, immobilization, and nitrification rates (van Veen et al., 1984; Bengtsson et al., 2003). Our study shows that both fungal and microbial community composition were related to C:N and percent N within these recovering forest soils. This suggests that these environmental microbiomes not only responded to environmental cues, but their role in biogeochemical cycling also may have affected their surrounding environments, such as through enhancing plant acquisition, bioavailability, or the lability of limited nutrients and minerals within this system.

Although the previously clearcut forests in Bear Meadow harbored more bacterial taxa than nearby old-growth forests, fungal taxa richness and diversity were unchanged across Bear Meadow regardless of forest management methods. Yet we detected more diverse bacterial, archaeal, and fungal communities in the lupine plots within the Pumice Plain than we did at Bear Meadow. Microbial biodiversity is often linked to enhanced ecosystem function and to both ecological and evolutionary drivers of terrestrial ecosystem response to environmental changes (Bardgett and van der Putten, 2014). Rather than diversity or taxa richness being the overarching indicator of forest recovery, our work detected varying degrees of microbial and fungal sensitivity to forest management, introductions of historical gophers, and soil C and N concentrations. Although we did not have comparable molecular evidence characterizing composition pre-eruption, our findings suggest that the compositional differences across extant conditions may be long-lasting, or may not attenuate, even at decadal time scales.

This study showed that old-growth forests harbored more types of root-associated mutualistic fungi from the ancestral AMF guild than were found in the clearcut forest from Bear Meadow. Arbuscular mycorrhizal fungi are affected by soil moisture (Augé, 2001) and nutrient availability (Egerton-Warburton and Allen, 2000). Although we still have more to learn about the functional role of this cryptic AM group, which neither allocates excessive fungal hyphae within plant roots (i.e., intraradical) nor extends substantially into the interstitial places within the soil (i.e., extrametrical or extraradical) hyphae (Hart and Reader, 2002; Powell et al., 2009; Varela-Cervero et al., 2016a; Weber et al., 2019), we may aggregate AM communities based on trait-based approaches that assign AMF to functional groups based on physiological traits. These "ancestral AMF" may tolerate low soil moisture or nutrient levels while promoting plant performance in stressful or harsh environments, such as the Pumice Plain. AM life-history strategies and morphological features such as extraradical hyphae production may be advantageous for nutrient cycling, while genomic attributes, including genes for metabolic maintenance and survival, such as trehalose synthesis, may confer benefits in resource-limited soils (Brewer et al., 2019). Given their unclear allocation preferences, additional research into the functional variation within ancestral AM families may elucidate mechanisms for ecological specialty and inform our predictive framework for understanding which plant species are sensitive to benefits conferred by ancestral AM. Moreover, environmental parameters, such as water or N deficiencies, may also prime ancestral AM support of their host plants or fungal biogeochemical cycling within

novel ecosystems. As the frequency and severity of natural disasters increase, and as water or nutrient availability becomes limited, ancestral AM associations may be increasingly important in understanding successional dynamics, such as after major volcanic eruptions.

Multi-trophic interactions, such as AMF spore dispersal by small mammals, may ameliorate the impacts of severe disturbance and promote succession on the Pumice Plain of Mount St. Helens. Microbial propagules dispersed throughout the Pumice Plain likely pave the way for ecological succession by promoting soil aggregation, thereby preventing erosion and topsoil sloughing. As pumice material eroded downslope, it exposed new soils, and erosional rills and gullies served as "hotspots" for resprouting residual plants (Allen et al., 2005, in Dale et al., 2005). In this recovering system, spores also may have been deposited and established in the soil profile, along with plants and microbial residues (e.g., organic acids, microbial exudates, and sticky glycoproteins), which accumulate and aggregate soils, thereby influencing soil structure in these sensitive ecosystems.

Land use patterns may alter topsoil retention and soil health, which could play an important role in ecosystem resilience. Microorganisms illustrate environmental preferences, which could contribute toward our capacity to predict their responses to environmental change (Delgado-Baquierizo et al., 2018). Moreover, as ecosystems are altered by eruption, losses of dominant species could open niches to transient taxa. However, Grime's (1998) mass-ratio hypothesis suggests that disproportionate impacts on key soil processes, soil C, and biogeochemical cycling could be driven by primary production rates and the functional diversity of dominant plants. Likewise, belowground successional pathways and long-term retention or recovery of ecological communities could be affected differentially by forest management methods, such as whether conventional tree removal or sustainable forestry methods were employed in the decades preceding the 1980 eruption. Different trends in community assembly may be explained by land use or forest management methods, as well as whether stochastic or deterministic processes dominate successional dynamics.

Findings from our study could help us evaluate whether the rules that govern macro-community assembly in natural systems also apply to microbial communities following volcanic eruptions (Pérez-Hernández and Gavilán, 2021). Our work examined fungal and bacterial communities in the historic blast zone of Mount St. Helens. It highlighted structural differences in microbial groups exhibiting particular life history strategies or traits that may underscore their own survival or their host retention. Despite their small size, microbes may be key to understanding how communities respond to severe disturbances in novel ecosystems. Our investigation of these natural processes is germane for addressing contemporary climate crises and ecosystems recovering from anthropogenic activities. Likewise, we can learn much from examining the strategies and structure of macro- and microorganisms that establish and proliferate in resource-limited, harsh, or inhospitable environments, aided by dispersal vectors. As successional processes, biotic legacies, and priority effects interact to create novel biotic communities in unlikely places following eruptions, microbial processes influence their proximal environments in creating microsite conditions, which may be conducive to biotic survival, owing to spatial constraints. Findings from our study show that, after disturbance, microbial reestablishment may have been partitioned by historic forest management practices and augmented by assisted migrations (as per Shade et al., 2023; microbiome rescue), which could have long-term impacts on successional dynamics, microbial communities, and host-symbiont interactions.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI SRA, accession PRJNA1124544, http://www.ncbi.nlm.nih.gov/bioproject/1124544.

Author contributions

MM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. MA: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing, Visualization, Writing – original draft. MP: Data curation, Investigation, Methodology, Validation, Writing – review & editing, Formal analysis, Software. RH: Conceptualization, Data curation, Investigation, Methodology, Supervision, Validation, Visualization, Writing – review & editing. HS: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – review & editing. LF: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – review & editing, Visualization. LA: Data curation, Investigation, Methodology, Software, Validation, Writing – review & editing, Resources, Supervision. JB: Investigation, Methodology, Resources, Supervision, Writing – review & editing, Conceptualization, Project administration. EA: Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing, Data curation, Funding acquisition, Validation.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. MM was supported by a UC Presidential Postdoctoral Fellowship, Metabolic Studio Research Grant, and Wildlife Conservation Society's Climate Adaptation Fund; other funds supporting authors' time were from grants awarded to ELA: NSF EAR 2012878, NSF RISE 1541047, and NIH 1U54MD013368-01A1.

Acknowledgments

We wish to dedicate this manuscript to the memory of our mentor, colleague, and friend James MacMahon, who initiated our

research program at Utah State University. The authors appreciate colleagues who assisted with the initial surveying, investigation, and subsequent field work, including JA MacMahon, CM Crisafulli, EB Allen, S Morris, and others, as well as those who helped with sample preparation, such as the undergraduate research assistants, volunteers, and staff from the Aronson laboratory at UC Riverside, including M Zaza, R Bond, and A Wilson. We appreciate all members of the Allen labs for laboratory assistance and feedback, SE Weber for his contribution to our understanding of AM fungal guilds, UCR's Center for Conservation Biology, and D Lyons at the UCR Environmental Science Research Laboratory for the use of analytical instruments for characterizing soil chemistry from our Mount St. Helens soil and tephra samples. MM was supported by a UC Presidential Postdoctoral Fellowship, Metabolic Studio Research Grant, and the Wildlife Conservation Society's Climate Adaptation Fund. Because of the magnitude of the Mount St/ Helens eruption, there were many opportunities for interdisciplinary teams of scientists to study the recovery process and better understand how to conserve sensitive and threatened ecosystems. Currently, our efforts to safeguard our study sites from development are hindered and these locations are at risk of being destroyed by planned construction and forest roads into the heart of the Pumice Plain within the Mount St. Helens National Volcanic Park. This study is dedicated to the hundreds of scientists, students, and researchers who have provided technical support and contributed to the historic and ongoing research at Mount St. Helens. We would like to acknowledge the teams and individuals who continue to explore opportunities and face challenges associated with conducting research within this system of great ecological importance.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frmbi.2024.1399416/full#supplementary-material

References

Abdulla, H. (2009). Bioweathering and biotransformation of granitic rock minerals by actinomycetes. *Microbial Ecol.* 58, 753–761. doi: 10.1007/s00248-009-9549-1

Allen, M. F., Crisafulli, C., Friese, C. F., and Jeakins, S. L. (1992). Re-formation of mycorrhizal symbioses on Mount St Helens 1980–1990: interactions of rodents and mycorrhizal fungi. *Mycological Res.* 96, 447–453. doi: 10.1016/S0953-7562(09)81089-7

Allen, M. F., Crisafulli, C. M., Morris, S. J., Egerton-Warburton, L. M., MacMahon, J. A., and Trappe, J. M. (2005). *Mycorrhizae and Mount St. Helens: story of a symbiosis. Ecological responses to the 1980 eruption of Mount St. Helens*. New York, NY: Springer. 221–231.

Allen, M. F., and MacMahon, J. A. (1988). Direct VA mycorrhizal inoculation of colonizing plants by pocket gophers (*Thomomys talpoides*) on Mount St. Helens. *Mycologia* 80, 754–756. doi: 10.1080/00275514.1988.12025615

Allen, M. F., MacMahon, J. A., and Andersen, D. C. (1984). Reestablishment of Endogonaceae on Mount St. Helens: survival of residuals. *Mycologia* 76, 1031–1038. doi: 10.1080/00275514.1984.12023947

Allen, M. F., O'Neill, M. R., Crisafulli, C. M., and MacMahon, J. A. (2018). "Succession and mycorrhizae on mount st. Helens," in C. Crisafulli and V. Dale (eds) *Ecological responses at Mount St. Helens: revisited 35 years after the 1980 eruption*. New York, NY: Springer. 199–215.

Alvarado, P., Teixeira, M. D. M., Andrews, L., Fernandez, A., Santander, G., Doyle, A., et al. (2018). Detection of Coccidioides posadasii from xerophytic environments in Venezuela reveals risk of naturally acquired coccidioidomycosis infections. *Emerging Microbes infections* 7, 1–13. doi: 10.1038/s41426-018-0049-6

Andersen, D. C. (1982). Observations on Thomomys talpoides in the region affected by the eruption of Mount St. Helens. *J. Mammalogy* 63, 652–655. doi: 10.2307/1380271

Andersen, D. C., and MacMahon, J. A. (1985). The effects of catastrophic ecosystem disturbance: the residual mammals at Mount St. Helens. *J. Mammalogy* 66, 581–589. doi: 10.2307/1380942

Augé, R. M. (2001). Water relations, drought and vesicular-arbuscular mycorrhizal symbiosis. *Mycorrhiza* 11, 3–42. doi: 10.1007/s005720100097

Banwart, S. A., Nikolaidis, N. P., Zhu, Y. G., Peacock, C. L., and Sparks, D. L. (2019). Soil functions: connecting earth's critical zone. *Annu. Rev. Earth Planetary Sci.* 47, 333–359. doi: 10.1146/annurev-earth-063016-020544

Bardgett, R. D., and van der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning. *Nature* 515, 505–511. doi: 10.1038/nature13855

Baross, J. A., Dahm, C. N., Ward, A. K., Lilley, M. D., and Sedell, J. R. (1982). Initial microbiological response in lakes to the Mt St Helens eruption. *Nature* 296, 49–52. doi: 10.1038/296049a0

Bengtsson, G., Bengtson, P., and Månsson, K. F. (2003). Gross nitrogen mineralization-, immobilization-, and nitrification rates as a function of soil C/N ratio and microbial activity. *Soil Biol. Biochem.* 35, 143–154. doi: 10.1016/S0038-0717(02)00248-1

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

Borcard, D., Gillet, F., and Legendre, P. (2011). *Numerical ecology with R* Vol. 2 (New York: springer), 688).

Borowicz, V. A. (2001). Do arbuscular mycorrhizal fungi alter plant–pathogen relations? *Ecology* 82, 3057–3068. doi: 10.1890/0012-9658(2001)082[3057:DAMFAP]2.0.CO;2

Bowd, E. J., Banks, S. C., Bissett, A., May, T. W., and Lindenmayer, D. B. (2022). Disturbance alters the forest soil microbiome. *Mol. Ecol.* 31, 419–447. doi: 10.1111/mec.16242

Brewer, T. E., Aronson, E. L., Arogyaswamy, K., Billings, S. A., Botthoff, J. K., Campbell, A. N., et al. (2019). Ecological and genomic attributes of novel bacterial taxa that thrive in subsurface soil horizons. *MBio* 10, 10–1128. doi: 10.1128/mBio.01318-19

Brookes, P. C., Landman, A., Pruden, G., and Jenkinson, D. S. (1985). Chloroform fumigation and the release of soil nitrogen: a rapid direct extraction method to measure microbial biomass nitrogen in soil. *Soil Biol. Biochem.* 17, 837–842. doi: 10.1016/0038-0717(85)90144-0

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi: 10.1038/ismej.2017.119

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010a). "QIIME allows analysis of high-throughput community sequencing data," in *Nature methods*, vol. 7 (5), 335–336. doi: 10.1038/nmeth.f.303

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624. doi: 10.1038/ismej.2012.8

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., et al. (2010b). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Nat. Acade. Sci.* 108, 4516–4522. doi: 10.1073/pnas.1000080107

Chen, J., Xiao, Q., Xu, D., Li, Z., Chao, L., Li, X., et al. (2023). Soil microbial community composition and co-occurrence network responses to mild and severe disturbances in volcanic areas. *Sci. Total Environ.* 901, 165889. doi: 10.1016/j.scitotenv.2023.165889

Coba de la Pena, T., Fedorova, E., Pueyo, J. J., and Lucas, M. M. (2018). The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organelle? *Front. Plant Sci.* 8, p.2229.

Dale, V. H., Swanson, F. J., and Crisafulli, C. M. (2005). *Ecological responses to the 1980 eruptions of mount st* (New York: Helens. Springer-Verlag).

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., et al. (2018). A global atlas of the dominant bacteria found in soil. *Science* 359, 320–325. doi: 10.1126/science.aap9516

Dove, N. C., and Keeton, W. S. (2015). Structural complexity enhancement increases fungal species richness in northern hardwood forests. *Fungal Ecol.* 13, 181–192. doi: 10.1016/j.funeco.2014.09.009

Dumbrell, A. J., Ashton, P. D., Aziz, N., Feng, G., Nelson, M., Dytham, C., et al. (2011). Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. *New Phytol.* 190, 794–804. doi: 10.1111/j.1469-8137.2010.03636.x

Edgar, R. C., and Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31, 3476–3482. doi: 10.1093/bioinformatics/btv401

Egerton-Warburton, L. M., and Allen, E. B. (2000). Shifts in arbuscular mycorrhizal communities along an anthropogenic nitrogen deposition gradient. *Ecol. Appl.* 10, 484–496. doi: 10.1890/1051-0761(2000)010[0484:SIAMCA]2.0.CO;2

Findley, R. (1981). Mount St. Helens, Mountain with a death wish. *Natl. geographic* 159, 3–33.

Franklin, J. F. (2005). *Reconfiguring disturbance, succession, and forest management: The Science of Mount St. Helens*. Eds. V. H. Dale, F. J. Swanson and C. M. Crisafulli (Springer-Verlag, New York: Ecological Responses to the 1980 Eruptions of Mount St. Helens), V–VII.

Freund, H. L. (2023). doi: 10.5281/zenodo.8264886

Fruchter, J. S., Robertson, D. E., Evans, J. C., Olsen, K. B., Lepel, E. A., Laul, J. C., et al. (1980). Mount St. Helens ash from the 18 May 1980 eruption: chemical, physical, mineralogical, and biological properties. *Science* 209, 1116–1125. doi: 10.1126/science.209.4461.1116

Ghosh, S., Osman, S., Vaishampayan, P., and Venkateswaran, K. (2010). Recurrent isolation of extremotolerant bacteria from the clean room where Phoenix spacecraft components were assembled. *Astrobiology* 10, 325–335. doi: 10.1089/ast.2009.0396

Grime, J. P. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *J. Ecol.* 86, 902–910. doi: 10.1046/j.1365-2745.1998.00306.x

Halvorson, J. J., Smith, J. L., and Kennedy, A. C. (2005). "Lupine effects on soil development and function during early primary succession at Mount St. Helens," in *Ecological responses to the 1980 eruption of Mount St. Helens* (Springer New York, New York, NY), 243–254).

Hart, M. M., and Reader, R. J. (2002). Taxonomic basis for variation in the colonization strategy of arbuscular mycorrhizal fungi. *New Phytol.* 153, 335–344. doi: 10.1046/j.0028-646X.2001.00312.x

Hernandez Garcia, M., Calabi-Floody, M., Conrad, R., and Dumont, M. (2020). Analysis of the microbial communities in soils of different ages following volcanic eruptions: Microbial communities in volcanic soils. *Pedosphere* 30, 126–134. doi: 10.1016/S1002-0160(19)60823-4

Ihaka, R., and Gentleman, R. (1996). R: A Language for Data Analysis and Graphics}. *Journal of Computational and Graphical Statistics*. (5)3:299-314.

Jackson, A. C., Jorna, J., Chaston, J. M., and Adams, B. J. (2022). Glacial legacies: microbial communities of Antarctic refugia. *Biology* 11, p.1440. doi: 10.3390/biology11101440

Kassambara, A. (2018). *ggpubr: 'ggplot2'-based publication ready plots. R package version*, Vol. 2.

Lee, J., Lee, S., and Young, J. P. W. (2008). Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiol. Ecol.* 65, 339–349. doi: 10.1111/fem.2008.65.issue-2

Lee, M. D. (2019). Happy Belly Bioinformatics: an open-source resource dedicated to helping biologists utilize bioinformatics. *J. Open Source Educ.* 4, p.53. doi: 10.21105/jose.00053

Legendre, P., and Legendre, L. (2012). *Numerical ecology* (Vol. 24). Elsevier. New York, NY.

Logan, W. B. (2007). *Dirt: The ecstatic skin of the earth* (New York, NY: WW Norton & Company).

MacMahon, J. A., and Warner, N. (1984). "Dispersal of mycorrhizal fungi: processes and agents," in *VA mycorrhizae and reclamation of arid and semiarid lands* (University of Wyoming Press, Laramie), 28–41.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2014). Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* 2, e593. doi: 10.7717/peerj.593

Maltz, M. R., Carey, C. J., Freund, H. L., Botthoff, J. K., Hart, S. C., Stajich, J. E., et al. (2022). Landscape topography and regional drought alters dust microbiomes in the sierra nevada of california. *Front. Microbiol.* 13, 856454. doi: 10.3389/fmicb.2022.856454

McCook, L. J. (1994). Understanding ecological community succession: Causal models and theories, a review. *Vegetatio* 110, 115–147. doi: 10.1007/BF00033394

Memoli, V., Eymar, E., García-Delgado, C., Esposito, F., Panico, S. C., De Marco, A., et al. (2018). Soil element fractions affect phytotoxicity, microbial biomass and activity in volcanic areas. *Sci. total Environ.* 636, 1099–1108. doi: 10.1016/j.scitotenv.2018.04.327

Mueller, E. A., Wisnoski, N. I., Peralta, A. L., and Lennon, J. T. (2020). Microbial rescue effects: how microbiomes can save hosts from extinction. *Funct. Ecol.* 34, 2055–2064. doi: 10.1111/1365-2435.13493

Ni, G., Lappan, R., Hernández, M., Santini, T., Tomkins, A. G., and Greening, C. (2023). Functional basis of primary succession: Traits of the pioneer microbes. *Environ. Microbiol.* 25, p.171. doi: 10.1111/1462-2920.16266

Öpik, M., Vanatoa, A., Vanatoa, E., Moora, M., Davison, J., Kalwij, J. M., et al. (2010). The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytol.* 188, 223–241.

Perez-Hernandez, J., and Gavilan, R. G. (2021). Impacts of land-use changes on vegetation and ecosystem functioning: old-field secondary succession. *Plants* 10, 990. doi: 10.3390/plants10050990

Phillips, J. D., and Lorz, C. (2008). Origins and implications of soil layering. *Earth-Science Rev.* 89, 144–155. doi: 10.1016/j.earscirev.2008.04.003

Phillips, M. L., Weber, S. E., Andrews, L. V., Aronson, E. L., Allen, M. F., and Allen, E. B. (2019). Fungal community assembly in soils and roots under plant invasion and nitrogen deposition. *Fungal Ecol.* 40, 107–117. doi: 10.1016/j.funeco.2019.01.002

Picone, N., Hogendoorn, C., Cremers, G., Poghosyan, L., Pol, A., van Alen, T. A., et al. (2020). Geothermal gases shape the microbial community of the volcanic soil of Pantelleria, Italy. *MSystems* 5, 10–1128. doi: 10.1128/mSystems.00517-20

Powell, J. R., Parrent, J. L., Hart, M. M., Klironomos, J. N., Rillig, M. C., and Maherali, H. (2009). Phylogenetic trait conservatism and the evolution of functional trade-offs in arbuscular mycorrhizal fungi. Proceedings of the Royal Society B. *Biol. Sci.* 276, 4237–4245.

Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A. H., Nieuwdorp, M., and Levin, E. (2020). Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PloS One* 15, e0227434. doi: 10.1371/journal.pone.0227434

Pulsford, S. A., Lindenmayer, D. B., and Driscoll, D. A. (2016). A succession of theories: purging redundancy from disturbance theory. *Biol. Rev.* 91, 148–167. doi: 10.1111/brv.12163

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* 35, 833–844. doi: 10.1038/nbt.3935

R: Development core team. (2004). R: A language and enviroment for statistical computing. *R foundation for statistical computing*, Vienna, Austria. http://www.r-project.org

Reichman, O. J., and Seabloom, E. W. (2002). The role of pocket gophers as subterranean ecosystem engineers. *Trends Ecol. Evolution.* 17, 44–49. doi: 10.1016/S0169-5347(01)02329-1

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e258. doi: 10.7717/peerj.2584

Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22, 939–946. doi: 10.1101/gr.128124.111

Shade, A. (2023). Microbiome rescue: directing resilience of environmental microbial communities. *Curr. Opin. Microbiol.* 72, p.102263. doi: 10.1016/j.mib.2022.102263

Stothers, R. B. (1984). The great Tambora eruption in 1815 and its aftermath. *Science* 224, 1191–1198. doi: 10.1126/science.224.4654.1191

Taylor, D. L., Walters, W. A., Lennon, N. J., Bochicchio, J., Krohn, A., Caporaso, J. G., et al. (2016). Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing. *Appl. Environ. Microbiol.* 82, 7217–7226. doi: 10.1128/AEM.02576-16

van Veen, J. A., Ladd, J. N., and Frissel, M. J. (1984). Modeling C and N turnover through the microbial biomass in soil. *Plant Soil* 76, 257–274. doi: 10.1007/BF02205585

Varela-Cervero, S., López-García, Á., Barea, J. M., and Azcón-Aguilar, C. (2016a). Differences in the composition of arbuscular mycorrhizal fungal communities promoted by different propagule forms from a Mediterranean shrubland. *Mycorrhiza* 26, 489–496. doi: 10.1007/s00572-016-0687-2

Varela-Cervero, S., López-García, Á., Barea, J. M., and Azcón-Aguilar, C. (2016b). Spring to autumn changes in the arbuscular mycorrhizal fungal community composition in the different propagule types associated to a Mediterranean shrubland. *Plant Soil* 408, 107–120. doi: 10.1007/s11104-016-2912-3

Varela-Cervero, S., Vasar, M., Davison, J., Barea, J. M., Öpik, M., and Azcón-Aguilar, C. (2015). The composition of arbuscular mycorrhizal fungal communities differs among the roots, spores and extraradical mycelia associated with five Mediterranean plant species. *Environ. Microbiol.* 17, 2882–2895. doi: 10.1111/1462-2920.12810

Vitousek, P. M. (2004). *Nutrient cycling and limitation: Hawai'i as a model system* (Princeton New Jersey: Princeton University Press).

Vuong, H. B., Thrall, P. H., and Barrett, L. G. (2017). Host species and environmental variation can influence rhizobial community composition. *J. Ecol.* 105, 540–548. doi: 10.1111/1365-2745.12687

Wagg, C., Bender, S. F., Widmer, F., and van der Heijden, M. G. (2014). Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proc. Natl. Acad. Sci.* 111, 5266–5270. doi: 10.1073/pnas.1320054111

Weber, S. E., Diez, J. M., Andrews, L. V., Goulden, M. L., Aronson, E. L., and Allen, M. F. (2019). Responses of arbuscular mycorrhizal fungi to multiple coinciding global change drivers. *Fungal Ecol.* 40, 62–71. doi: 10.1016/j.funeco.2018.11.008

Williamson, J., and Harrison, S. (2002). Biotic and abiotic limits to the spread of exotic revegetation species. *Ecol. Appl.* 12, 40–51. doi: 10.1890/1051-0761(2002)012[0040:BAALTT]2.0.CO;2

Yokobe, T., Hyodo, F., and Tokuchi, N. (2020). Volcanic deposits affect soil nitrogen dynamics and fungal–bacterial dominance in temperate forests. *Soil Biol. Biochem.* 150, 108011. doi: 10.1016/j.soilbio.2020.108011

Zimmerman, N., Izard, J., Klatt, C., Zhou, J., and Aronson, E. (2014). The unseen world: environmental microbial sequencing and identification methods for ecologists. *Front. Ecol. Environ.* 12, 224–231. doi: 10.1890/130055

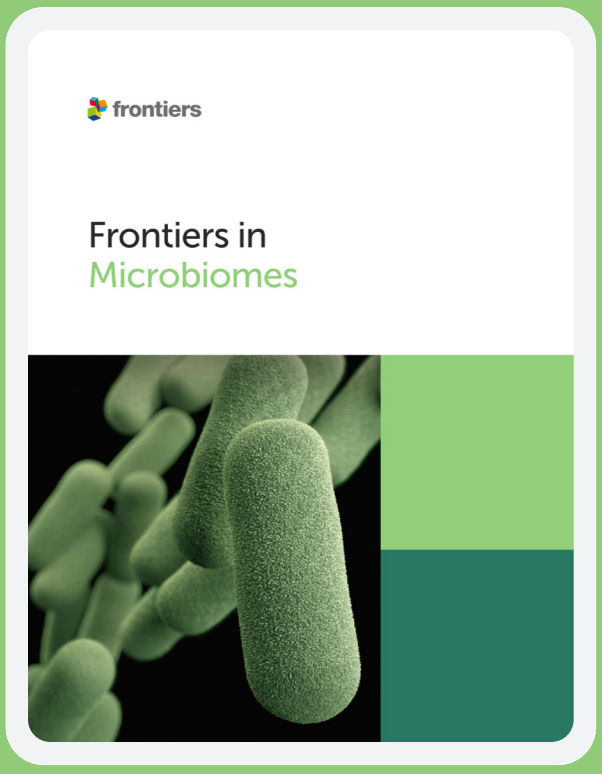
Frontiers in Microbiomes

Explores impactful connections between microbiomes and the world around us

Advances our understanding of how microbiomes generate positive or negative outcomes for their hosts and environments.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact