UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CHARACTERIZING THE EFFECT OF HABITAT STRUCTURE ON MICROBIAL COMMUNITY ASSEMBLY AND SECONDARY METABOLISM FOR NATURAL PRODUCT DISCOVERY

A DISSTERTAION

SUBMITTED TO THE GRADUATE FACILITY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

BY

EMILY NEWBY JUNKINS

Norman, Oklahoma

2021

CHARACTERIZING THE EFFECT OF HABITAT STRUCTURE ON MICROBIAL COMMUNITY ASSEMBLY AND SECONDARY METABOLISM FOR NATURAL PRODUCT DISCOVERY

A DISSERTATION APPROVED FOR THE

DEPARTMENT OF MICROBIOLOGY AND PLANT BIOLOGY

BY THE COMMIITTEE CONSISTING OF

Dr. Bradley S. Stevenson, Chair

Dr. Robert W. Cichewicz

Dr. Anne K. Dunn

Dr. Laura-Isobel McCall

Dr. Tassie K. Hirschfeld

This is for my mother, for her contagious curiosity.

Acknowledgments

First and foremost, I would like to thank Dr. Bradley Stevenson, for his constant support and enthusiasm. His eagerness to hear new ideas and talk about what observations mean was a crucial factor to my success as a student and my development as a microbial ecologist. He taught me how to ask a question and to craft a story from the answer; for that I am eternally grateful.

I extend my utmost appreciation to my committee members: Dr. Anne Dunn, Dr. Tassie Hirschfeld, Dr. Robert Cichewicz, and Dr. Laura-Isobel McCall. Their examples shaped me into the scientist I am. I focused on chemistry more so than I anticipated during this dissertation. So, I especially thank Dr. Robert Cichewicz for telling me that 'You can always learn new facts' and Dr. Laura-Isobel McCall for helping me do it. I also thank my other advisors over the years: to Dr. Krithi Sankaranarayanan, who laid my foundation in sequencing, computing, and microbiome sciences; to Dr. Catherine Mintler who helped me voice my reasons for pursuing graduate studies and cultivating my 'why', and to Dr. Liz Karr, who always advised me on my future with my best interests in mind. To my peers, Dr. Jane Lucas and Dr. Evan Gora, thank you for always talking ecological shop; and to my lab mates Dr. Heather Nunn, Dr. Brian Bill, and James Floyd for all of the guidance and all of the coffee walks.

I would not be the scientist I am without the example set for me by Dr. David Carter, I thank him for showing me my potential and for setting an example of curiosity and passion-driven research. He showed my what a life of research is, and without his support, guidance, and, most of all, friendship; I would not have pursued this doctorate.

I especially thank The Pat Tillman Foundation and my fellow Tillman Scholars for being my constant inspiration and embodying the values of service, scholarship, humble leadership,

and impact which kept me focused. No single experience has guided the direction of my life more than being a part of this community.

Most of all I thank my husband and partner, John Mark, for his unwavering encouragement.

Table of Contents

List of Tables

Chapter 2

List of Figures

Chapter 2

Chapter 3

Chapter 4

Abstract

Microorganisms exist in dynamic networks that catalyze reactions and respond to stimuli. They organize into populations that make up the microbial communities we know to drive large-scale biogeochemical cycles that can determine the fate of all organisms in an ecosystem. For something so minute, their impact is extensive. What culminates in global-scale processes starts with the small interactions microorganisms have with their environment, and importantly, each other.

Microbial communities are dynamic systems; one of the most intricate traits of a microbial community is the interactions taking place among its members. Consider a microbial community a small neighborhood, where interactions define the network and structure of a community. Many of these interactions are mediated by secondary metabolites as these are the molecules microorganisms use to communicate and compete. Because of these bioactive roles, microbial secondary metabolites represent promising antimicrobial drug targets, and their structures make up a large proportion of our current antibiotics on the market. The diversity of secondary metabolites in the microbial world is promising, as the increase in antimicrobial resistance among pathogens has intensified the need for new drugs and strategies.

With advances in high throughput sequencing and metabolomic methods, research efforts continue to characterize the untapped biosynthetic potential of microbial communities as they interact with each other and the environment. Considering the network in which communities assemble and interact, this dissertation sought to leverage the ecology behind the production of secondary metabolites in an effort to cultivate antimicrobial producers. Since microbes exist in structured environments, we aimed to characterize the effect of habitat structure on the production of secondary metabolites. We hypothesized that manipulating the habitat during

cultivation may select for more competitive interactions as organisms compete for space and nutrients, potentially producing secondary metabolites as a means to be successful.

This dissertation showed that manipulating the interaction landscape of a microbial community can alter the microbial community assembly patterns and the chemodiverisity in structured habitats. These studies add to continuing efforts to (1) characterize the drivers of microbial community assembly, (2) associated changes in microbial diversity with chemical diversity, and (3) determine cultivation methods that select for diverse microbial assemblages and that elicit a secondary metabolic response for both natural product discovery and chemical ecology.

Chapter 1: Habitat Structure, Secondary Metabolisms, and Natural Product Discovery in Cultivated Microbiomes

Overview

The ecological dynamics that govern microbial communities are still being uncovered. Much of what is known to be true of how communities assemble and how its members interact is similar to macroecology. However, dynamics can change with scale, and microbial communities offer a unique challenge in adapting macroecological theory at the microscopic scale. Within and between populations, habitat structure and physical space can dictate potential interactions, and often, those interactions are mediated by chemistry (e.g., secondary metabolites). Many of these interactions are competitive, for resources and/or space, where microorganisms are producing antimicrobial compounds to warn or even gain an advantage (Foster & Bell, 2012; Hibbing et al., 2010). These dynamics are all taking place on what I am calling the interaction landscape. The interaction landscape is the physical space in which microorganisms are communicating and competing, where the chemical gradients in the environment are affecting the microorganisms, and where the production of antimicrobial metabolites have the greatest effect. Manipulating the interaction landscape during cultivation may select for more competitive interactions and potentially the production of secondary metabolites as a population goes through community assembly.

Drivers of Community Assembly in Structured Habitats

Microbial communities are dynamic systems comprised of constantly interacting populations. Microbial ecologists have long sought to characterize and the define the rules that govern the

spatiotemporal landscape in which populations are distributed. The interplay among ecological forces like selection, diversification, drift, and dispersal acting on a given community are difficult to disentangle, thus making their effects difficult to predict. However, selection plays a crucial role in microbial communities as it pertains to interactions among its members. Selection encompasses powerful abiotic forces, like the environment, and biotic forces, like competition. One highly selective force is the anatomy of a habitat. We refer to habitat as the physical space a population inhabits, the gradients that exist or are constructed, and the interactions among the community.

Complex interplay between abiotic and biotic factors influences habitat structure. Structure can be influenced abiotically from the formation of microscale gradients in nutrients or chemicals like oxygen or light, as well, as, physical barriers, like surfaces. Structure will also have a biotic component, through the metabolic activity of microorganisms themselves via nutrient degradation and the production of metabolites involved in signaling or competition that can establish these microscale gradients (McNally & Brown, 2015). Overall, gradients can result in barriers. These barriers imposed by the habitat are dynamic too, in that they will both influence the structure of a microbial community and its spatiotemporal dynamics. But barriers are also subject to alteration by microorganisms themselves. For example, some organisms can alter the pH to drive out competitors, thus altering the chemical structure of the environment and changing selective pressures on other community members (Ratzke & Gore, 2018). Another example would be organisms that create physical structure through biofilm formation, but over time become susceptible to cheaters or non-biofilm producing populations, resulting in the collapse of the biofilm and changes in the distribution and abundance of community members

(Rainey & Rainey, 2003). Structured habitats can foster diverse community members by providing ecological opportunity. Over time, a structured habitat can lead to co-existence and cooperation (Geyrhofer & Brenner, 2020; Wakano et al., 2009) as populations occupy niches and may no longer compete for space or nutrients. But disturbances or changes in the structure of a habitat can select for competition, and potential competitive exclusion (Tilman, 1999), as populations seek to reassemble. Overall, structured environments can support diverse, complex microbial communities by providing biological and chemical heterogeneity.

Microbial Interactions

Perhaps one of the most complex traits of a community are the interactions taking place among its members. These interactions generate a network among community members; and are often mediated by secondary metabolites. These interactions can be positive or negative, and they can oscillate between cooperation and competition (Coyte et al., 2015; Oliveira et al., 2014; Jousset et al., 2013; Foster & Bell, 2012). Interactions within a microbial community will depend on deterministic factors like the environment where drivers of microbial interactions are resource-based, and members of populations and communities compete for nutrients and space (Traxler & Kolter, 2015; Slattery et al., 2001). Interactions within a microbial community will also depend on the richness and abundance distribution of the community members (Otto-Hanson & Kinkel, 2020; Chao & Levin, 1981). The results of microbial interactions can be chemical products (e.g., antimicrobial compounds or signaling molecules) and they can elicit changes (e.g., changes in diversity), though these will be heavily linked. For example, the production of a certain compound can be detrimental to one population resulting in a decline, altering the abundance distribution of the community, and triggering other populations to

produce certain compounds in reaction, such is the case for organisms that would rely on quorum-sensing regulated antimicrobial production, like *Pseudomonas* spp. (Brint & Ohman, 1995; Pierson et al., 1994), *Streptomyces* spp. (Corre et al., 2008), or *Bacillus* sp. (Stein et al., 2002). The products of microbial interactions will alter the chemical ecology of the interaction landscape, which would then alter the community structure, and this may be cyclic and continue as a community seeks to reach dynamic equilibrium and stability. In other words, dynamics of an interaction landscape may resemble cyclic, feedback mechanisms more so than end-point transactions. Lastly, the chemical products of microbial interaction have utility not just for microbial community dynamics in nature, but the function of these compounds can be applied to therapeutics (Newman & Cragg, 2016; Dias et al., 2012). For example, actinobacteria produce antimicrobial compounds as signaling molecules for their developmental life cycle and to ward off competitors (van der Meij et al., 2017); these same molecules are used as antimicrobial drugs, like erythromycin (Chng et al., 2008; McGuire et al., 1952), and scaffolds for synthetic analogs, like clarithromycin (Alvarez-Elcoro & Enzler, 1999).

Using Ecology to Find New Natural Products

Growing antibiotic resistance in pathogenic bacteria is surpassing the rate of drug discovery, intensifying the need for new drug therapies (Blaskovich, 2018) . Most natural products used as drugs are a result of microbial secondary metabolisms, often providing a competitive advantage for the producer (Traxler & Kolter, 2015). However, most microorganisms remain to be cultivated, representing an untapped resource of biosynthetic potential and drug discovery (Lloyd et al., 2018). Recent studies have begun to molecularly describe the genetic capability for uncharacterized drug pathways in many systems like the

mammalian gut (Donia et al., 2014) and soil (Crits-Christoph et al., 2018). However, these biosynthetic pathways are not expressed and/or the producing organism has not been isolated in culture. There are multiple strategies to approach this problem. Some efforts use the "one strain, many compounds" (OSMAC) approach that alters cultivation medium or conditions (e.g., pH or temperature), and uses co-cultivation with other strains to activate different silent gene clusters (Pan et al., 2019). Once these clusters are identified, some efforts focus on heterologous expression of biosynthetic pathways (Kang & Kim, 2021; Hao et al., 2019; Ongley et al., 2013; Wenzel et al., 2005). However, molecular efforts often rely on isolated strains, which rely on efforts to cultivate as many organisms as possible and screen each (Motley et al., 2017). In these cases, cultivation strategies that maximize the cultivable diversity from various systems will increase the number of organisms isolated and screened for the production of antimicrobial compounds. However, each approach has issues, whether that be the challenge of cloning large biosynthetic gene clusters and activating them, or only being able to cultivate the same organisms repeatedly. Focusing on cultivation strategies that both uncover previously uncultivated microorganisms and that elicit a native secondary metabolic response resulting in natural product production represent a promising avenue to explore for new drugs.

Since many compounds of natural product interest mediate microbial interactions, understanding the purpose behind the production of a compound may inform strategies to elicit that response. Specifically, using ecological theory to characterize the pressures that drive the distribution and expression of these secondary metabolites may provide a basis for isolating new drug targets in the future. Since these drug targets are often produced in response to microbial interactions (Traxler et al., 2013; Yim et al., 2007; Long & Azam, 2001), one area of focus for

discovery is bacterial interactions in diverse communities, specifically competitive interactions in which the production of antimicrobial compounds would offer an advantage to the producer. Here, we sought to use cultivation strategies that would force organisms to interact and select for the production of secondary metabolites. By allowing communities to assemble during cultivation, forcing them to interact, and allowing for chemical gradients to form in a structured habitat, we aimed to select for compounds that gave organisms an advantage by providing a cultivation condition in which expression of secondary metabolite pathways would be useful, like structured habitats with populations in close proximity and a diversity of ecological niches. During cultivation, both the initial environment (i.e., medium and cultivation conditions) and the colonizing community (i.e., inoculum) will affect what populations can grow, how those populations interact within and between populations, and how they then alter the interaction landscape. These conditions provide an opportunity to characterize the community-level, secondary metabolic response to changes in habitat structure, like disturbance.

Using New Tools to Answer Old Questions

The work presented in this dissertation sought to leverage competitive ecological interactions to enrich for antimicrobial compound producing organisms. Competitive interactions can be driven by the spatial structure of the habitat and the organisms present (García-Bayona & Comstock, 2018; Kraemer et al., 2017; Hibbing et al., 2010; Yim et al., 2007; Czárán & Hoekstra, 2003). We hypothesized that *manipulating the interaction landscape will select for competitive interactions and potentially the production of secondary metabolites.* Focusing on community diversity, habitat structure, and the metabolome provided the framework to ask the following research questions:

1. Can we maximize cultivable diversity through targeted cultivation strategies?
2. Does habitat structure influence cultivable diversity and metabolite production?
3. How does the microbial community structure change over time in response to disturbance in a structured habitat?

The first question is addressed in chapter two and describes the cultivable diversity of several inocula using different media types and treatments and quantifies a proportion of the cultivable diversity that produces antimicrobial compounds. This study sought to address the problems with natural product re-discovery by cultivating from a little-explored, but bioactive reservoir—roadkill. High throughput sequencing of cultivated communities revealed that much of what was cultivable was not detected in sequencing, highlighting the importance of cultivation in describing total diversity and confirming that using multiple cultivation strategies on a single inoculum will increase the number of organisms to be screened for natural product discovery.

The third and fourth chapters focused on habitat structure and disturbance by revisiting two, well-known experiments, but instead of using clonal populations, the studies presented here focused on diverse, mixed communities from a wastewater treatment facility. Using a complex inoculum, rather than a clonal population, allowed us to test the effect of habitat structure and disturbance on a whole community and to characterize the chemical response. Specifically, the second question is addressed in the third chapter by revisiting an experiment that showed static cultivation selected for phenotypic diversification in *Pseudomonas fluorescens* compared to a constantly mixed cultivation condition (Rainey & Travisano, 1998). This experiment used a

clonal population to demonstrate that a structured habitat selected for different colony morphologies (diversity) as populations partitioned along a gradient, particularly an oxygen gradient (Rainey & Travisano, 1998). In our study, we applied the same selective pressure to a mixed wastewater community. We hypothesized that structed habitats would select for higher diversity of the microbial community and the metabolome as members are forced to interact as they partition along a gradient.

The third question is addressed in the fourth chapter by revisiting an experiment that showed habitat structure and population size determined if a colicin-producing *Escherichia coli* could outcompete the sensitive *E. coli* strain (Chao & Levin, 1981). Here, we repeated the structured portion of that experiment by immobilizing mixed wastewater communities in soft agar overlays during cultivation and propagating the community over ten transfers to determine which organisms were most successful at reassembly after disturbance and used metabolomics to uncover a potential mechanism.

This dissertation sought to address the knowledge gap associated with effects of habitat structure on the selection and cultivability of natural product-producing organisms and the interaction-mediated production of antimicrobial compounds. Using cultivated mesocosms and altering the structure of the habitat, we were able to manipulate the interaction landscape of a microbial community and change the microbial and chemical diversity of microbial communities. We were able to use ecological theory to determine cultivation methods that elicit a secondary metabolic response with the goal of identifying ways to select for natural-product producers. Further, this dissertation adds to larger efforts to (1) characterize the drivers of

microbial community assembly, (2) correlate changes in microbial and chemical diversity, and (3) determine cultivation methods that produce diverse enrichments and that stimulate a secondary metabolic response for chemical and microbial ecology.

References

Alvarez-Elcoro, S., & Enzler, M. J. (1999). The macrolides: erythromycin, clarithromycin, and azithromycin. *Mayo Clinic Proceedings*, *74*(6). https://doi.org/doi: 10.4065/74.6.613

Blaskovich, M. A. T. (2018). The fight against antimicrobial resistance is confounded by a global increase in antibiotic usage [article-commentary]. *ACS Infectious Diseases*, *4*(6), 868–870. https://doi.org/10.1021/acsinfecdis.8b00109

Brint, J. M., & Ohman, D. E. (1995). Synthesis of multiple exoproducts in *Pseudomonas aeruginosa* is under the control of RhlR-RhlI, another set of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family. *Journal of Bacteriology*, *177*(24), 7155–7163. https://doi.org/10.1128/jb.177.24.7155-7163.1995

Chao, L., & Levin, B. R. (1981). Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(10 I), 6324–6328. https://doi.org/10.1073/pnas.78.10.6324

Chng, C., Lum, A. M., Vroom, J. A., & Kao, C. M. (2008). A key developmental regulator controls the synthesis of the antibiotic erythromycin in *Saccharopolyspora erythraea*. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11346–11351. https://doi.org/10.1073/pnas.0803622105

Corre, C., Song, L., O'Rourke, S., Chater, K. F., & Challis, G. L. (2008). 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17510–17515. https://doi.org/10.1073/pnas.0805530105

Coyte, K. Z., Schluter, J., & Foster, K. R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science*, *350*(6261), 663–666. https://doi.org/10.1126/science.aad2602

Crits-Christoph, A., Diamond, S., Butterfield, C. N., Thomas, B. C., & Banfield, J. F. (2018). Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature*, *558*(7710), 440–444. https://doi.org/10.1038/s41586-018-0207-y

Czárán, T. L., & Hoekstra, R. F. (2003). Killer-sensitive coexistence in metapopulations of micro-organisms. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1522), 1373–1378. https://doi.org/10.1098/rspb.2003.2338

Dias, D. A., Urban, S., & Roessner, U. (2012). A historical overview of natural products in drug discovery. *Metabolites*, *2*(2), 303–336. https://doi.org/10.3390/metabo2020303

Donia, M. S., Cimermancic, P., Schulze, C. J., Wieland Brown, L. C., Martin, J., Mitreva, M., Clardy, J., Linington, R. G., & Fischbach, M. A. (2014). A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell*, *158*(6), 1402–1414. https://doi.org/10.1016/j.cell.2014.08.032

Foster, K. R., & Bell, T. (2012). Competition, not cooperation, dominates interactions among culturable microbial species. *Current Biology*, *22*(19), 1845–1850. https://doi.org/10.1016/j.cub.2012.08.005

García-Bayona, L., & Comstock, L. E. (2018). Bacterial antagonism in host-associated microbial communities. *Science*, *361*(6408). https://doi.org/10.1126/science.aat2456

Geyrhofer, L., & Brenner, N. (2020). Coexistence and cooperation in structured habitats. *BMC Ecology*, *20*(1), 1–15. https://doi.org/10.1186/s12898-020-00281-y

Hao, T., Xie, Z., Wang, M., Liu, L., Zhang, Y., Wang, W., Zhang, Z., Zhao, X., Li, P., Guo, Z., Gao, S., Lou, C., Zhang, G., Merritt, J., Horsman, G. P., & Chen, Y. (2019). An anaerobic bacterium host system for heterologous expression of natural product biosynthetic gene clusters. *Nature Communications*, *10*(1), 1–13. https://doi.org/10.1038/s41467-019-11673-0

Hibbing, M. E., Fuqua, C., Parsek, M. R., & Peterson, S. B. (2010). Bacterial competition: Surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, *8*(1), 15–25. https://doi.org/10.1038/nrmicro2259

Jousset, A., Eisenhauer, N., Materne, E., & Scheu, S. (2013). Evolutionary history predicts the stability of cooperation in microbial communities. *Nature Communications*, *4*, 1–7. https://doi.org/10.1038/ncomms3573

Kang, H. S., & Kim, E. S. (2021). Recent advances in heterologous expression of natural product biosynthetic gene clusters in *Streptomyces* hosts. *Current Opinion in Biotechnology*, *69*, 118–127. https://doi.org/10.1016/j.copbio.2020.12.016

Kraemer, S. A., Soucy, J. P. R., & Kassen, R. (2017). Antagonistic interactions of soil pseudomonads are structured in time. *FEMS Microbiology Ecology*, *93*(5), 1–9. https://doi.org/10.1093/femsec/fix046

Lloyd, K. G., Ladau, J., Steen, A. D., Yin, J., & Crosby, L. (2018). Phylogenetically novel uncultured microbial cells dominate Earth microbiomes. *MSystems*, *3:e00055-1*. https://doi.org/10.1128/mSystems.00055-18

Long, R. A., & Azam, F. (2001). Antagonistic interactions among marine pelagic bacteria. *Applied and Environmental Microbiology*, *67*(3–12), 4975–4983. https://doi.org/10.1128/aem.67.11.4975-4983.2001

McGuire, J. M., Bunch, R. L., Anderson, R. C., Boaz, H. E., Flynn, E. H., Powell, H. M., & Smith, J. W. (1952). Ilotycin, a new antibiotic. *Antibiotics & Chemotherapy (Northfield)*, *2*(6), 281–283.

McNally, L., & Brown, S. P. (2015). Building the microbiome in health and disease: Niche construction and social conflict in bacteria. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1675). https://doi.org/10.1098/rstb.2014.0298

Motley, J. L., Stamps, B. W., Mitchell, C. A., Thompson, A. T., Cross, J., You, J., Powell, D. R., Stevenson, B. S., & Cichewicz, R. H. (2017). Opportunistic sampling of roadkill as an entry point to accessing natural products assembled by bacteria associated with non-anthropoidal mammalian microbiomes. *Journal of Natural Products*, *80*(3), 598–608. https://doi.org/10.1021/acs.jnatprod.6b00772

Newman, D. J., & Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *Journal of Natural Products*, *79*(3), 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055

Oliveira, N. M., Niehus, R., & Foster, K. R. (2014). Evolutionary limits to cooperation in microbial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(50), 17941–17946. https://doi.org/10.1073/pnas.1412673111

Ongley, S. E., Bian, X., Neilan, B. A., & Müller, R. (2013). Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Natural Product Reports*, *30*(8), 121–1138.

Otto-Hanson, L. K., & Kinkel, L. L. (2020). Densities and inhibitory phenotypes among indigenous *Streptomyces* spp. vary across native and agricultural habitats. *Microbial Ecology*, *79*(3), 694–705. https://doi.org/10.1007/s00248-019-01443-2

Pan, R., Bai, X., Chen, J., Zhang, H., & Wang, H. (2019). Exploring structural diversity of microbe secondary metabolites using OSMAC strategy: A literature review. *Frontiers in Microbiology*, *10*(FEB), 1–20. https://doi.org/10.3389/fmicb.2019.00294

Pierson, L. S., Keppenne, V. D., & Wood, D. W. (1994). Phenazine antibiotic biosynthesis in *Pseudomonas aureofaciens* 30-84 is regulated by PhzR in response to cell density. *Journal of Bacteriology*, *176*(13), 3966–3974. https://doi.org/10.1128/jb.176.13.3966-3974.1994

Rainey, P. B., & Rainey, K. (2003). Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, *425*(6953), 72–74. https://doi.org/10.1038/nature01906

Rainey, P. B., & Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*, *394*(6688), 69–72. https://doi.org/10.1038/27900

Ratzke, C., & Gore, J. (2018). Modifying and reacting to the environmental pH drives bacterial interactions. *PLoS Biology*, *16*(3). https://doi.org/10.1371/journal.pbio.2004248

Slattery, M., Rajbhandari, I., & Wesson, K. (2001). Competition-mediated antibiotic induction in the marine bacterium *Streptomyces tenjiimariensis*. *Microbial Ecology*, *41*(2), 90–96. https://doi.org/10.1007/s002480000084

Stein, T., Borchert, S., Kiesau, P., Heinzmann, S., Klöss, S., Klein, C., Helfrich, M., & Entian, K. D. (2002). Dual control of subtilin biosynthesis and immunity in *Bacillus subtilis*. *Molecular Microbiology*, *44*(2), 403–416. https://doi.org/10.1046/j.1365-2958.2002.02869.x

Tilman, D. (1999). The ecological consequences of changes in biodiversity: A search for general principles. *Ecology*, *80*(5), 1455–1474. https://doi.org/10.1890/0012-9658(1999)080[1455:tecoci]2.0.co;2

Traxler, M. F., & Kolter, R. (2015). Natural products in soil microbe interactions and evolution. *Natural Product Reports*, *32*(7), 956–970. https://doi.org/10.1039/c5np00013k

Traxler, M. F., Watrous, J. D., Alexandrov, T., Dorrestein, P. C., & Kolter, R. (2013). Interspecies interactions stimulate diversification of the *Streptomyces coelicolor* secreted metabolome. *MBio*, *4*(4), 1–12. https://doi.org/10.1128/mBio.00459-13

van der Meij, A., Worsley, S. F., Hutchings, M. I., & van Wezel, G. P. (2017). Chemical ecology of antibiotic production by actinomycetes. *FEMS Microbiology Reviews*, *41*(3), 392–416. https://doi.org/10.1093/femsre/fux005

Wakano, J. Y., Nowak, M. A., & Hauert, C. (2009). Spatial dynamics of ecological public goods. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(19), 7910–7914. https://doi.org/10.1073/pnas.0812644106

Wenzel, S. C., Gross, F., Zhang, Y., Fu, J., Stewart, A. F., & Müller, R. (2005). Heterologous expression of a myxobacterial natural products assembly line in pseudomonads via Red/ET recombineering. *Chemistry and Biology*, *12*(3), 349–356. https://doi.org/10.1016/j.chembiol.2004.12.012

Yim, G., Wang, H. H., & Davies, J. (2007). Antibiotics as signalling molecules. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1483), 1195–1200. https://doi.org/10.1098/rstb.2007.2044

Chapter 2: Using Plate-Wash PCR and High-Throughput Sequencing to Measure Cultivated Diversity for Natural Product Discovery Efforts

Foreword

The research in Chapter Two was conducted by Emily N. Junkins with the help of Dr. Heather Nunn and the team of Mammalian Microbiome undergraduate researchers (Audrey Byrd, Jenifer Lesiter, Brittany Hansen, Chris Hauger, Jaclyn Jacobs) who aided in media preparation and cultivation. Isolates used for the pathogen panel were kindly provided by Dr. Robert Cichewicz, Institute for Natural Products and Research Technologies (INPART), University of Oklahoma. Dr. Graham Wiley and staff at University of Oklahoma Consolidated Core Laboratories provided sequencing support and instrumentation. Emily N. Junkins performed all analyses. This work culminated in a manuscript currently (as of 2021) under review at Frontiers. Citation: Junkins, E. N., & Stevenson, B. S. (2020). Assessing and Maximizing Cultivated Diversity with Plate-Wash PCR and High Throughput Sequencing. *bioRxiv*. https://doi.org/10.1101/2020.11.19.390864

Abstract

Molecular techniques continue to reveal a growing disparity between the immense diversity of microbial life and the small proportion that is in pure culture. The disparity, originally dubbed "the great plate count anomaly" by Staley and Konopka, has become even more vexing given our increased understanding of the importance of microbiomes to a host and the role of microorganisms in the vital biogeochemical functions of our biosphere. Searching for novel antimicrobial drugs often focuses on screening a broad diversity of microorganisms. If diverse microorganisms are to be screened, they need to be cultivated. Recent innovative research has used molecular techniques to assess the efficacy of cultivation efforts, providing invaluable feedback to cultivation strategies for isolating targeted and/or novel microorganisms. Here, we aimed to determine the efficiency of cultivating representative microorganisms from a non-human, mammalian microbiome, identify those microorganisms, and determine the bioactivity of isolates. Sequence-based data indicated that around 57% of the ASVs detected in the original inoculum were cultivated in our experiments, but nearly 53% of the total ASVs that were present in our cultivation experiments were *not* detected in the original inoculum. In light of our controls, our data suggests that when molecular tools were used to characterize our cultivation efforts they provided a more complete and more complex, understanding of which organisms were present compared to what was eventually detected during cultivation. Lastly, about 3% of the isolates collected from our cultivation experiments showed inhibitory bioactivity against an *already* multidrug-resistant pathogen panel, further highlighting the importance of informing and directing future cultivation efforts with molecular tools.

Introduction

Molecular techniques and sequencing technologies have helped to describe many microbial communities and provide tools that allow us to study representative microorganisms and their activities in lieu of isolation. Application of molecular approaches to the study of microbiology have led to the discovery and characterization of rare biospheres of the most extreme environments (Sogin et al., 2006) and enabled researchers to decipher the metabolic potential and activity of microorganisms from the human gut (Donia et al., 2014) to the bottom of the ocean (Tortorella et al., 2018). Now, changes in microbial community structure and function can be monitored with molecular techniques throughout the course of a disease (Halfvarson et al., 2017) or with changing environmental conditions (Arora-Williams et al., 2018), such that we may begin to predict the roles of microbial populations and outcomes over time (Singh et al., 2010; Gilbert et al., 2016).

Isolating, or even just cultivating representative microorganisms is still considered the best means towards understanding their physiology, metabolism, and potential ecological roles, but this remains to be a major hurdle. The degree to which any microbial community is represented in culture varies considerably. Since Staley and Konopka coined the term “great plate count anomaly” (Staley and Konopka, 1985), numerous studies have sought to quantify the culturable organisms from various biomes (see (Lloyd et al., 2018) for a compiled list of most probable number (MPN) based approaches). A recent study estimated that 81% of all microbial cells on Earth belong to uncultured genera or higher classifications and about 25% of microbial cells are from uncultured phyla (Lloyd et al., 2018). Specifically, 26% of marine, 31% of host-associated, 31% of eutrophic lakes, and 30% of soil metagenomic sequences have cultured

representatives at the family level (Lloyd et al., 2018). Not surprisingly, human associated microbiomes have more cultured representatives, likely affected by a focused effort (Eloe-Fadrosh et al., 2016). Sequencing technology has shed light onto the previously unknown diversity of the microbial world, outpacing our ability to cultivate these organisms by an estimated rate of 2.4- and 2.5-fold for bacteria and archaea, respectively (Schloss et al., 2016).

Efforts to isolate yet uncultured microorganisms have gathered momentum as researchers have begun to apply new technologies and innovations (Connon and Giovannoni, 2002; Stevenson et al., 2004; Nichols et al., 2010). Microbiologists routinely manipulate the composition of growth media and treatments of the inocula in order to increase the diversity of microorganisms represented in culture (e.g., (Reasoner and Geldreich, 1985; Connon and Giovannoni, 2002; Davis et al., 2005; Kopke et al., 2005; Laiger et al. 2012; Oberhart et al., 2015; Henson et al., 2016; Kawasaki and Kamagata, 2017; Kim et al., 2019; Bartelme et al., 2020)). As an example, treating an inoculum with ethanol can select for spore-forming bacteria by killing vegetative cells and leaving spores intact that may otherwise be overgrown and remain unisolated (Browne et al., 2016). Furthermore, antioxidants in media can support the cultivation of aerobic and even anaerobic organisms, by alleviating oxidative stress (La Scola et al., 2014; Dione et al., 2016). The goals of this study were to 1) determine the efficiency of cultivating representative microorganisms from a non-human, mammalian microbiome, 2) identify cultivated microorganisms through comprehensive, whole community sequencing, and 3) screen representative isolates for antimicrobial activity. With one aim of this study being to cultivate many, different microorganisms, the approaches described above were of particular interest. We anticipated that we would select for gut-associated bacteria; antioxidants would help mitigate

oxygen stress for organisms best suited for microoxic or anoxic environments, and ethanol treatment would reduce the abundant vegetative cells, allowing us to better access the many spore-forming taxa that reside in the gut.

When cultivation approaches are coupled with molecular analyses, the total diversity and identity of cultivated taxa can be determined (Justice et al., 2017; Yang et al., 2015; Medina et al., 2017), and specific, targeted taxa can be detected (Stevenson et al., 2004). More recently, high throughput sequencing has made it possible to quantify exactly what proportion of an original sample has been cultivated (Medina et al., 2017; Zehavi et al., 2018; Cummings et al., 2020). Studies using this approach showed that roughly 23% of microorganisms in bovine rumen fluid were cultivable in liquid enrichments, some of which were not detected in direct sequencing efforts (Zehavia et al., 2018), while 96% of those in human bronchoalveolar lavage samples (Cumming et al., 2020), and 60% of those detected in/on American toads (Medina et al., 2017) were cultivated. A recent study characterized the microbial diversity within the same inoculum spread onto both high- and low-nutrient media by harvesting biomass on each plate and identifying all taxa present (Medina et al., 2017). Their approach built upon a previous study that sought to isolate abundant but yet-uncultured taxa using defined, low nutrient media and a PCR-based screening approach using group specific primers, termed plate-wash PCR (PWPCR) (Stevenson et al., 2004). Medina et al. confirmed that low nutrient-media would allow for the growth of a more diverse subset of bacteria from the original sample and that molecular screening (i.e., 16S PWPCR) is a practical way to characterize the microbial community growing on an entire agar plate (Medina et al., 2017). This data is invaluable in identifying medium composition, inoculum treatments, and incubation conditions that can target certain taxa, or

maximize the cultivated diversity for large-scale isolation efforts (see Henson et al., 2016; Bartelme et al., 2020). However, these studies using solid agar plates may have overestimated diversity for some taxa by focusing on the relative abundance of taxa in their analyses.

Pure cultures have always been the "gold standard" for the in-depth study of microorganisms. The need for isolation is especially important when screening for the production of secondary metabolites for drug discovery. Historically, the process of isolating microorganisms for drug discovery has relied mainly on high throughput cultivation efforts that increase the number of isolates being screened (e.g., (Motley et al., 2017; Helfrich et al., 2018; Hoffmann et al., 2018)), enrich for bioactive production from known producers like *Streptomycetes* spp. (van der Meij et al., 2017), or focus on novel or rare taxa thought to have uncharacterized biosynthetic capabilities (Stierle et al., 2006; Ding et al., 2010; Rangseekaew and Pathom-aree, 2019). We aimed to do the same by sampling a host-associated microbial community and cultivating them under conditions designed to obtain as many different organisms as possible through the use of different media types. Screening cultivated microorganisms with the broadest diversity is expected to provide the greatest chance of uncovering natural products that are novel drug targets. Previous studies have focused on extreme environments for novel natural products (Smanski et al., 2015), but we expand on this idea to include non-human mammals. We predicted that this would increase the diversity of microorganisms and compounds screened for useful antimicrobial compounds and argue that the shared life-history between a host and members of its microbiomes may produce compounds that are less cytotoxic to the host. This idea was originally supported by a survey of microbiomes

collected from roadkill mammals, which produced two new cyclic lipodepsipeptides of pharmaceutical interest (Motley et al., 2017).

Here, we sought to use both general and selective media along with various treatments to the inocula to collectively enrich for as many different microorganisms as possible from the oral and gut microbiomes of a roadkill mammal (racoon). The goals for this study were to (1) compare the diversity of cultured bacterial communities on multiple types of media combined with various treatments of the inoculum, (2) determine which media, treatments, or combination cultivated the highest richness of bacterial taxa, and (3) determine how many isolates produced bioactive molecules against a panel of multi-drug resistant pathogenic microorganisms.

Materials and Methods

Sample Collection and Storage

A 3 mile stretch of Oklahoma State Highway 9 in Norman, OK between the University of Oklahoma, Norman campus and Lake Thunderbird was used for our opportunistic sampling area (around 35.1879, -97.3005). This well-traveled highway runs through undeveloped prairie, scrub and forest, cattle ranches, and rural residential land, providing many mammalian roadkill specimens. Sample collection was conducted under Oklahoma Scientific Collector Permit #5250 (Motley et al., 2017). The oral cavity and rectum of a racoon (*Procyon lotor*) deceased less than 8 h was sampled in triplicate (n=3) with sterile, nylon swabs. One concern was that swabbing for sequencing first would remove or alter the biomass when the cultivation swab samples were taken. To avoid this, each orifice was swabbed three times. At each of those times, two swabs were used, one destined for cultivation the other destined for sequencing. In doing so, any alteration of the biomass would be reflected in each set of triplicate swabs, for sequencing and for cultivation. This generated a total of 6 swabs from each orifice, 3 for sequencing and 3 for cultivation. Swabs to be used for cultivation were pooled by orifice and stored in a sterile conical tube containing 3.0 mL 1X PBS, stored on ice, and immediately transferred back to the laboratory (under 45 minutes). Swabs to be used for sequencing were transferred to separate, empty, sterile conical tubes to preserve the technical replicate, stored on ice, and immediately transferred back to the laboratory (under 45 minutes). At the laboratory, swabs for sequencing were transferred to BashingBead Lysis Tubes (Zymo Research Corp., Irvine, CA) containing 750 µL of BashingBead Buffer (Zymo Research Corp.). Cells were lysed through homogenization in a Mini-BeadBeater-8 (BioSpec Products Inc., Bartlesville, OK) at maximum speed for 45 seconds. Homogenized samples were then stored at -20˚C until needed for DNA extraction.

Swabs for cultivation were vortexed to suspend cells. A 900 µL aliquot of the cell suspension intended for storage was centrifuged at 10,000 x g for 1 min, the supernatant was removed. The cell pellet was resuspended in 750 µL of BashingBead Buffer (Zymo Research Corp.) and transferred to BashingBead Lysis Tubes (Zymo Research Corp., Irvine, CA) prior to lysis through homogenization in a Mini-BeadBeater-8 (BioSpec Products Inc., Bartlesville, OK) at maximum speed for 45 seconds. Archives were stored at -80˚C. The remaining 2.1 mL of the cell suspension for each sample was serially diluted with PBS. Aliquots of 50 µL from the 10^{0}-10^{-6} dilutions were spread onto 13 different media/treatment combinations (see Table S1). In all, each inoculum was spread onto the agar media, n=6, for each medium type and incubated for 6 days at 30˚C, three plates would be used for PWPCR and three plates would be used to pick colonies for isolation. As a control, 50 µL of PBS used for sample collecting and dilutions was spread onto each medium type and incubated for 6 days at 30˚C.

Culture-Dependent Bacterial Community Analysis

Cultivation was conducted with a combination of broad-spectrum media with decreased nutrients and selective media in order to "cast a wide net" and maximize the cultivable diversity of each inoculum (see quasi-factorial design in Table S1). The media used included 0.1X strength tryptic soy broth (Bacto, USA) with 1.5% agar, ROXY and derivatives with hemin (0.1 g/L) and alpha-ketoglutarate (2.0 g/L)(referred to as ROXYakgh) (Dione et al., 2016), 0.25X strength R2A medium (Reasoner and Geldreich, 1985), and blood agar (per L: 10.0 g meat extract, 10.0 g peptone, 5.0 g sodium chloride, 15 g agar, 5% sheep's blood (ThermoFisher, USA), pH 7). To date, this is the first instance using ROXY medium and plate-wash PCR techniques. In addition to different media types, one variation included the pretreatment of the

inocula for 4 hours in 70% ethanol at 22-24˚C (Browne et al., 2016). Catalase (40,000 U/L) was added as a treatment to TSA, ROXY and R2A to remove peroxides produced during autoclaving (Kawasaki and Kamagata, 2017; Kim et al., 2019). The last treatment included the addition of streptomycin at 50 mg/L as a broad-spectrum selection agent against Gram-negative and some Gram-positive species (Schatz et al., 1944). To test which combinations of media and treatments resulted in cultivated communities that represented the diversity observed in the original sample, we grew each inoculum (mouth or rectum) on each combination of media/treatment types and sequenced the resulting biomass.

Collection of Total Biomass from Agar Plates

Biomass was harvested from each agar plate in order to identify the microorganisms present and to compare this to culture-independent analyses of the inoculum. The biomass from each agar plate was harvested by adding 2.0 mL of PBS to the surface of the plate and suspending the colony biomass by agitation with a sterile spreader. The suspended colony biomass was collected and transferred to a 2.0 mL microcentrifuge tube, pelleted at 10,000 x g for 1 min, and resuspended in 750 µL of BashingBead Buffer (Zymo Research Corp.). Each sample was transferred to a BashingBead Lysis Tube (Zymo Research Corp.) and homogenized for 45 seconds at maximum speed using a Mini-BeadBeater-8 (BioSpec Products Inc., Bartlesville, OK). The lysed samples were stored at -20˚C until DNA was extracted.

DNA Extraction and Sequencing

Before DNA extraction, each sample was thawed on ice and homogenized for 30 seconds at maximum speed using a Mini-BeadBeater-8 (BioSpec Products Inc., Bartlesville, OK). DNA

was extracted according to manufacturer specifications using the Zymo *Quick*-DNA Fungal/Bacterial Miniprep kit (Cat# D6005, Zymo Research Corp., Irvine, CA). For community characterization, a conserved region of the SSU rRNA gene of most bacteria, archaea, and eukarya was amplified using primers 515F-Y and 926R (Parada et al., 2015) via the following PCR protocol: initial denaturation at 94˚C for 2 min, followed by 30 cycles of denaturation at 94˚C for 45 s, annealing at 50˚C for 45 s, and extension at 68˚C for 90 s, with a final extension at 68˚C for 5 min. These primers produced two amplicons, a ~400 bp fragment for bacteria and archaea, and a 600 bp fragment for eukaryotes. The forward primer 515F-Y (5'-GTA AAA CGA CGG CCA G CCG TGY CAG CMG CCG CGG TAA-3') contains the M13 tag (underlined) fused to the 5' end of the forward primer (Kraus et al., 2016). The reverse primer 926R (5'-CCG YCA ATT YMT TTR AGT TT-3') was unmodified from Parada et. al 2015. Each PCR contained 5 PRIME HOT master mix (1X; 5 PRIME Inc., Gaithersburg, MD), 0.2 µM of each primer, and 3.0 µL of extracted DNA at a final volume of 50 µL. The amplified fragments were purified using Sera-Mag magnetic beads (GE) with the AmPureXP (Beckman Coulter) protocol at a final concentration of 1.8x v/v. After purification, 3 µL of each PCR product, 1x 5 PRIME HOT master mix (Quantabio, Massachusetts, USA), 0.2µM of the 926R primer, and 0.2 µM of a specific 12 bp oligonucleotide was used in a separate barcoding PCR (6 cycles) in 50 µL reactions to attach a unique barcode to amplicons of each library. The same thermocycler protocol was used as above but only run for 6 cycles. The now barcoded amplicons were purified using Sera-Mag (GE) beads with the AmPureXP (Beckman Coulter) protocol to a final volume of 40 µL, quantified using the QuBit HS DS DNA assay kit (Thermo Fisher Scientific Inc., Waltham, MA), and pooled in equimolar amounts. The pooled, barcoded amplicon libraries were then concentrated to a final volume of 40 µL (209 ng/ µL) with an Amicon-Ultra filter

(Millipore, Burlington, MA, USA) following manufacturer's protocol. The combined amplicon libraries were denatured according to Miseq library preparations protocol (Illumina, San Diego, CA, USA). The sample was loaded at a concentration of 10 pM and sequenced using 2x250 paired-end strategy on the Miseq (Illumina San Diego, CA, USA) platform for 251 cycles.

ASV Inference and Bacterial Community Characterization

Barcodes from amplified regions of the 16S rRNA gene were removed and demultiplexed using QIIME v 1.9.1 (Caporaso et al., 2010). Demultiplexed reads were trimmed for adapters and quality filtered as a part of the *dada2* pipeline (Callahan et al., 2016) and amplicon sequence variants (ASVs) were inferred using the forward reads (203 bp). Specifically, forward reads were used instead of merged forward and reverse reads due to poor quality in the overlap region. In order to retain sequencing depth, but at the expense of some taxonomic resolution, reverse reads were omitted from all analyses. In short, forward reads (V4) were trimmed and quality filtered (maxEE=2) via *filterAndTrim*, dereplicated via *derepFastq*, and ASVs were inferred (*dada*). Concerning ASV filtration, the dada2 pipeline does not infer an amplicon sequence variant that is only supported by one read, therefore, the final ASV table does not contain singletons, which could affect statistical methods for richness (though no statistics were performed on richness in this study). Lastly, chimeras were removed (*removeBimeraDenovo,* method ="consensus") and taxonomy was assigned using the SILVA database v32 (Quast et al., 2013, Yilmaz et al., 2014). Sequences mapping to chloroplasts and mitochondria were filtered from the dataset. The final dataset consisted of ~1.7 million reads resulting in 463 ASVs with a median of 17,347 sequences per sample (min = 132, max= 266,678).

Data Availability: Sequence data has been deposited at NCBI's Sequence Read Archive (SRA) database under accession number PRJNA675861.

Data Analysis

Community diversity was analyzed using the *phyloseq* (McMurdie and Holmes, 2013) and *vegan* (Dixon, 2003) packages in R. Relative abundance was only used to directly assess the diversity of the original inocula. Presence/absence and richness were used as the main metrics to compare cultivated samples because of the variation in biomass (i.e. colony size) between microbial colonies on and between agar plates. Since this procedure duplicated swabs when sampling, the data from the molecular control swabs were pooled for analysis. The variability in swabbing, as these are technical replicates, would be different than the variability of triplicates resulting from cultivation, since these could represent biological replicates resulting from the stochasticity of what cells landed on which plate.

Alpha diversity was calculated as richness of ASVs, while beta diversity was measured with non-metric multi-dimensional scaling (NMDS) using Jaccard distances. A permanova with the *adonis2* function in the *vegan* package was used to compare differences in the composition of taxa cultivated across media types, treatment, and orifices (Dixon, 2003). Due to the nestedness of orifice and cultivation conditions, orifice was used as a stratum or fixed factor, and media and treatment were used as predictor variables (~media*treatment). Differences in cultivable diversity could be due to different starting communities from each orifice. So, orifice was treated as a fixed factor to control for any influence it could have on the significant differences in diversity between cultivable communities. To fully understand the significance described in the

permanova, a permdisp with the *betadisper* function in *vegan* was used to describe any within sample variance (i.e., across replicate agar plates) that could explain any significant differences detected in the permanova. Hypothesis testing via anova with permutations (n=999) was used with permdisp to determine any significant differences in variation within samples. These statistical methods allowed us to determine if our cultivation strategies cultivated diverse subsets of organisms that were different from each other between media and treatment types.

To identify certain ASV families more associated with an orifice, medium, or treatment, a species indicator analysis was performed using the *indicspecies* package (permutations = 999) (De Cáceres and Legendre, 2009; De Cáceres et al., 2011). Species indicator analyses would help efforts to determine which medium or treatment cultivated which taxa.

Bioassays and Isolate Identification

Colonies were selected for isolation based on morphology, with the intent of sampling as many different colony morphologies as possible. From average of 8.1×10^{10} CFUs/mL cultivated across all medium types and dilutions, a total 238 colonies were chosen for isolation based on differential colony morphology from the mouth and rectum on a subset of the media types (0.1X TSA, ROXY with alpha-ketoglutarate and hemin, and 0.25X R2A). Colonies were isolated by streaking for isolation of single colony morphologies at least three times before the assumption of pure cultures. To determine bioactivity, soft agar overlays in 0.1X TSA (0.8% agar) were used to test for growth inhibition of a panel of drug-resistant pathogens (*Klebsiella pneumonia* ATCC# 13883, *Enterococcus faecium* ATCC# 51559, *Pseudomonas aeruginosa* ATCC# 10145, and *Candida albicans* ATCC# 565304). To prepare for the bioassay, pathogens were incubated

overnight, shaking at 250 rpm at 30°C, in 0.1X tryptic soy broth (TSB) (Bacto, USA). A 400 µL aliquot of each pathogen was used to inoculate 3.6 mL of molten TSA soft agar (at 55 ºC), mixed, and immediately poured onto a 0.1X TSA agar plate for a final overlay volume of 4 mL. Colony material from each isolate was transferred to the bioassay plate with a sterile toothpick by etching an 'X' into the overlay. The inoculated bioassay plates were incubated at 30 ºC for 24 hours. Inhibition of the pathogen in the overlay was characterized by a zone of clearing surrounding the isolate (Figure S1). For cryopreservation, each isolate was grown in 5 mL of 0.1X TSB in a 16 x 150 mm test tube shaking at 250 rpm for 48 h. An aliquot of 800 uL of the culture was transferred to a 2 mL screw cap tube, along with 200 uL of 80% glycerol, for a final concentration of 20% glycerol. After mixing by vortex for 30 s, it was then stored at -80ºC. Partial SSU rRNA sequences were used to identify each isolate. Genomic DNA was extracted from each isolate using QuickExtract (Lucigen, Wisconsin, USA) according to manufacturer's protocol from the above mentioned 0.1X TSB cultures. The SSU rRNA gene was amplified using primers 8F (3'-ATGC-5') and 1492R (3'-ATGC-5') at a concentration of 0.2 mM, 1x 5 PRIME HOT master mix (Quantabio, Massachusetts, USA) , and 2.0 mL of DNA template per 50 mL reaction via the following PCR protocol: initial denaturation at 94˚C for 2 min, followed by 30 cycles of denaturation at 94˚C for 30 s, annealing at 55˚C for 45 s, and extension at 72˚C for 45 s, with a final extension at 72˚C for 10 min. Amplified fragments were purified using Sera-Mag magnetic beads (GE) with the AmPureXP (Beckman Coulter) protocol at a final concentration of 1.8x v/v. Some isolates were not recoverable after the process of freeze-thawing, and some extractions did not result in enough DNA for sequencing. Purified amplicons (a total of 192 samples) were sent for Sanger sequencing using the 8F primer (Genewiz, New Jersey, USA). Resulting sequences and chromatograms were assessed for quality using

Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Sequences with low quality (< Q20) were removed, low quality ends were trimmed with MEGA X (Kumar et al., 2018), and the final sequences were identified using the SILVA v. 132 classifier online server (Quast et al., 2013). A total of 117 isolates had suitable reads and met quality thresholds for taxonomic identification.

Results

Cultivated Microbial Richness Differed Between Orifice and Cultivation Condition

As expected, a library of SSU rRNA genes from the original inocula showed that they had higher richness (alpha diversity) than the cultivated organisms from both orifices (Figure 1). The microbial community sampled from the rectum had higher overall richness (153 ASVs) compared to that of the mouth (86 ASVs) (Figure 1). Overall, we were able to cultivate 57.3% of the ASVs detected in the inocula (58.1% from the mouth and 57.5% from the rectum). Interestingly, about half (52.7%) of the ASVs detected from the cultivation experiments were not represented in either library from the inocula (i.e., molecular control) (Figure 2). The number of ASVs detected on a certain medium or treatment differed between orifices. For instance, cultivation on TSA medium treated with catalase resulted in only 29 ASVs from the rectum inoculum compared to around 70 different ASVs from the mouth. The fewest number of ASVs was detected on the ROXY medium amended with alpha-ketoglutarate and hemin from both the mouth and rectum. The cultivated beta diversity differed based on orifice and cultivation condition. Specifically, community structure differed based on orifice and treatments, more so than base medium type (Figure 3). Differences in the cultivated community composition and diversity between medium and treatment confirmed that, collectively, using different cultivation techniques increased the number of organisms cultivated from each inoculum.

Selective Treatments Increased Cultivated Richness Compared to Medium Type or Orifice

An indicator species analysis was used to determine which microbial families were significantly associated with certain conditions. A total of ten families were identified to be

significantly ($p < 0.05$) associated with orifice, medium type, and/or treatment. Only one family, the Aeromonadaceae, was significantly associated with a particular orifice, the mouth. This family was also the only family significantly ($p = 0.001$) associated with a certain base medium type, ROXY (Table 1). Treatment of the medium and/or the inoculum selected for the most diverse set of indicator species, with a total of 9 families significantly ($p < 0.05$) associated with a certain treatment or combination of treatments (Table 1). Certain bacterial families like Staphylococcaceae and Paenibacillaceae were associated with the ethanol pretreatment of inocula, whereas the Wohlfarhtiimonodaceae and Metaschinikowiceae were more associated with the addition of streptomycin to media (Table 1).

Much of the cultivable microbial diversity was shared between the mouth and the rectum, indicated by the lack of significant differences in community structure described by the permanova and permdisp analyses (discussed below). Based on molecular analyses, 27 families were shared between the mouth and rectum before cultivation, while 6 families were unique to the mouth and 6 families were unique to the rectum (Figure 4). Negative cultivation controls (e.g., PBS) yielded no growth, supporting the assumption that any biomass collected from plates originated from the inoculum. Lastly, our cultivation experiments led to the detection of 18 families from the mouth and 8 families from the rectum that were not detected through direct molecular analysis of the inoculum (Figure 4). This accounted for around half (52.7%) of the ASVs detected during cultivation (Figure 2). This could be explained by differences in biomass and growth characteristics not being captured by sequence data from the inocula (discussed below).

The Raccoon Microbiome Contained Cultivable Bioactive Bacteria

A total of 238 isolates were collected from a subset of media (0.25X R2A, 0.1X TSA, and ROXY with alpha-ketoglutarate and hemin). Despite choosing colonies based on differential morphology, many of these isolates were expected to be redundant, since the diversity of the initial inoculum was low (Figure 1). This was confirmed with many isolates redundant at the genus level and the majority of the identifiable isolates belonging to genera *Serratia* (34.2%) and *Klebsiella* (17.9%) (Figure S2). Each of the 238 isolates were assayed for antimicrobial activity against a panel of *multi-drug resistant* pathogens. A total of 7 isolates showed antimicrobial activity, identified as a zone of clearing of one of the antibiotic resistant panel organisms. This equated to a ~3% "hit" rate against already resistant organisms. Six of these isolates were recovered from the mouth, while one, a *Bacillus* sp., was recovered from the rectum. Three isolates were recovered from R2A, 3 from TSA and 1, an Enterobacteriaceae sp., from ROXYakgh. Three isolates exhibited antimicrobial activity against *Klebsiella pneumonia*, three against *Enterococcus faecium*, and one, a *Pseudomonas* sp., against *Candida albicans.* Five of the isolates were identified by partial 16S rRNA gene sequence identity (Table 2).

The Composition of Cultivated Taxa was Highly Variable

Due to the nestedness of this experiment (i.e. taxa observed during cultivation theoretically would be observed in sequencing), the stratification of taxa based on sampled orifice was corrected for hypothesis testing with permanova by defining orifice as a stratum in the *adonis2* function in the *vegan* package (Dixon, 2003)(Figure 5A). The majority of the explained variation and significant differences in cultivated taxa were due to treatment of the inoculm or medium type ($F= 2.44$, $R^2 = 0.12$, $p = 0.001$), while base medium type alone

explained only ~5% of the observed variation (F= 1.54, R^2 = 0.05, p = 0.002). The interactions of base medium and treatment were insignificant and did not fit the model (F= 1.52, R^2= 0.06, p= 0.208). An analysis of variances (permdisp) was used to determine if any significance was driven by a dispersion effect rather than a location effect for orifice, base medium type, treatments, and medium+treatments. Mean distances to centroids did not differ significantly for orifice (F=0.69, p=0.41; Mouth 0.51; Rectum 0.53) or medium+treatment (F= 1.38, p=0.20; TSA=0.47, TSAcatalase=0.52, TSAstrep=0.51, TSAethanol=0.45, R2A=0.44, R2Acatalase=0.47, R2Aethanol=0.45, Blood=0.45, Bloodethanol=0.50, Bloodstrep=0.53, ROXY=0.48, ROXYcatalase=0.49, ROXYakgh=0.52), meaning little inter-sample variation was detected between replicates. In contrast, the mean distances did differ for treatment (F=3.3, p=0.02; alpha-ketoglutarate+hemin=0.52, no treatment (base)=0.50, catalase=0.53, ethanol=0.49, streptomycin=0.56) and base medium type (F=4.4, p=0.01; Blood=0.56, R2A=0.50, ROXY=0.53, TSA=0.55). The significance in dispersion within the treatment and media could correspond to the significance in permanova results, where much of the variation explained in the permanova may be due to the intrinsic variability and stochasticity among replicates of the same inoculum.

Discussion

We used mouth and rectal samples from a roadkill mammal (raccoon, *Procyon lotor*) to (1) compare the diversity of cultured bacterial communities on multiple media types and/or treatments of the inoculum, (2) determine which media/treatments cultivated the highest richness of bacterial taxa, and (3) determine if any isolated organisms produced bioactive molecules against a drug resistant pathogen panel. The occurrence of antimicrobial resistance is increasing at a faster rate than new therapies are entering the market (Newman and Cragg, 2016; Centers for Disease Control, 2019). Since nearly two-thirds of our current antimicrobial drugs are linked to microorganisms (Clardy et al., 2006; Newman and Cragg, 2016), one of our goals was to cultivate a diverse array of bacteria from a non-human mammalian microbiome and screen these isolates for the production of antimicrobial metabolites. Growing microbial populations in the laboratory that are representative of the overall microbial diversity within an inoculum continues to be a major limitation for the field of microbiology. Here, we sought to use a combination of selective medium types and treatments of the inocula that might enrich for certain groups present in the oral and gut microbiomes of a roadkill racoon. In order to capture a more comprehensive sample of microbial diversity, we used high-throughput sequencing of 16S rRNA gene amplified libraries to characterize the original inocula and the cultivated populations that were washed from the surface of agar media.

We anticipated finding a significantly lower percentage of cultivable organisms relative to the molecular-based measures of diversity from each orifice, which we described as the cultivated percentage. We also anticipated that the identity of recovered taxa would be similar between media types and even orifices. The differences between organisms cultivated on

different media, from different orifices was low. This suggested that the original inocula were not as differentially diverse as expected, despite differing measures of richness in the molecular controls. However, when observing the significance described in the permanova and permdisp, the percentage of explained variation was low. This suggested that the stoichiometry of diverse and variable inocula could play a large part in the outcome of this and any similar cultivation experiment, particularly when plating across multiple medium types. Based on the low percentage of explained variation, we also approached this data set from the perspective of the unobserved or residual variation, such that some of this discussion will aim to pose questions about what makes cultivation so difficult. Overall, we observed that cultivation yielded not just a fraction of what was detected in sequence data, but rather expanded the total diversity that potentially described the original sample.

Cultivation Can Be Optimized at Large Scales

Convergence of microbial communities occurred between the mouth and rectum for both the molecular samples and the cultivated communities. This was expected since as an animal undergoes decomposition, the microbial diversity decreases, especially in rectal communities (Guo et al., 2016), and becomes dominated by certain taxa, namely Proteobacteria (Metcalf et al., 2013). It is also important to point out that the preservation of a "typical" oral or rectal community was unlikely to be observed, not only due to changes as decomposition mentioned above, but also that structures may not be intact after collision with a vehicle. In our samples, despite being a single time point roughly 8 hours after death, the mouth was dominated by Proteobacteria while the rectum samples contained mostly Firmicutes, Proteobacteria, and Bacteroides (Figure S3). Overall, about 87% of the microbial families detected in the molecular

controls were shared between the mouth and rectum. This increase in shared taxa was also expected, as the proportion of cultivable organisms was expected to be lower (Jensen, 1968) and the differences in diversity would not be detected through the selection of cultivability. Because of the small proportion of molecularly characterized microorganisms that are routinely recovered in the laboratory, we employed multiple variations in cultivation, like medium type and inoculum treatments to "cast a wider net."

The treatments that were used selected for organisms that could withstand ethanol stress (e.g., spore-formers), or that were streptomycin resistant. These treatments had a greater effect on the cultivable diversity compared to changes in base medium alone, likely because the base media shared many of the same oxidizable substrates. It is logical to hypothesize that the inoculum exposed to selective treatments would have decreased cultivated richness compared to the same inoculum on an untreated medium of the same type, since the selective agent would inhibit the growth of a portion of the population. Instead, the cultivable richness and diversity varied for the same treatments on different base medium types and between orifices (Figure 1 and Figure 4). This could be due to the treatments selecting for specific groups of taxa (i.e. antibiotics) or creating an environment where certain taxa were able to out-compete other taxa on non-selective medium (i.e. low nutrients, increased incubation and slow growers) (Davis et al., 2005). Specifically, medium amended with catalase or streptomycin and inocula treated with ethanol resulted in cultivated organisms not identified on other plates, or even detected in the molecular controls in some cases (Figure 4). One possible explanation would be that this was due to the inherent stochasticity associated with cultivation (discussed below). Despite this variability, we were able to cultivate roughly 57% of the species richness detected with

molecular approaches. Previous studies measuring cultivation success with high-throughput sequencing have reported similar levels of cultivable richness from the skin of toads (~60%) (Medina et al., 2017) and rumen fluid (23%) (Zehavi et al., 2018). Medina et al. reported greater variation in cultivability between inocula (different animals of the same species from the same area). Our data showed that more variability in cultivated taxa was observed between selective additions to the medium or treatments of the inocula (Figure 4), and not the inocula (different orifices) (Figure 3). Similar to Zehavi et al, different medium types and treatments increased the cultivable proportion of the original sample.

Bioactivity and Phylogenetic Diversity of Isolates

The research reported here was based on previous success of using roadkill mammals as a source of antimicrobial-producing microorganisms (Motley et al., 2017). Our intent was to quantify and broaden the diversity of cultivated bacteria by the addition of various media and treatments. This effort was measured by using plate-wash PCR and SSU rRNA gene sequencing to thoroughly assess the taxa that were cultivated. Simultaneously, colonies were also picked for isolation from replicate plates (n = 3), screened for antimicrobial production and identified through SSU rRNA gene sequencing. Colonies were chosen based on differential morphologies. Overall, 7 of the 238 recovered isolates showed bioactivity against the already antibiotic-resistant pathogen panel, resulting in a ~3% hit rate. Five bioactive isolates were sequenced. Three belonged to the phylum Proteobacteria (class Gammaproteobacteria) and two belonged to the phylum Firmicutes. Two isolates were identified as a *Bacillus* sp. and a *Pseudomonas* sp., both genera known to contain species that produce bioactive compounds (Fuller et al., 1971; Raaijmakers et al., 1997; Hamdache et al., 2011; Sumi et al., 2015; Motley et al., 2017).

Interestingly, the isolates identified as an *Enterococcus* sp. and *Klebsiella* sp. showed inhibitory activity to a member of its own genus, *Enterococcus faecium* and *Klebsiella pneumonia,* respectively. This finding was consistent with the hypothesis that closely-related organisms often have a negative effect on each other due to resource overlap (Hibbing et al., 2009), which has been demonstrated for these genera (de Vos et al., 2017). In addition, some populations could be making antimicrobial metabolites, like bacteriocins, that are effective against close relatives (Chao and Levin, 1981). Future directions include sequencing the genome of each isolate, investigating biosynthetic gene clusters, and using liquid chromatography and mass spectrometry to investigate putative compounds produced in the bioassays.

Cultivation Yields Complex Data

The cultivation and isolation of many organisms at once can be incredibly laborious. New technologies and approaches can help increase throughput, which is necessary to overcome the re-isolation of abundant, common, and easy to cultivate organisms. However, it is still important to be able to assess the efficacy of any cultivation effort. Unfortunately, we uncovered an intrinsic property of cultivation that is not readily realized unless assessed at a large scale like it was here, variability. Here, we did not detect the same cells (e.g., colonies measured in PWPCR) consistently growing on replicate plates (Figure S4).

Neither differences in cultivation strategy nor orifice sampled could explain the majority of the observed variation in community structure (Figure 5A-B). We can attribute unexplained variance (i.e. residuals) observed between communities (i.e. beta diversity) to stochastic effects of cultivation, but that is not fundamentally measurable in this case. We can speculate that

randomness would play a role in cultivation as samples are diluted (Justice et al., 2017; Zehavi et al., 2018) or the spatial heterogeneity of the inoculum is altered, as demonstrated on a much larger scale using leaf litter (Albright et al., 2019). In these instances, we could expect different inocula resulting in different assemblages of cultivated organisms. We also know that microorganisms immobilized on an agar surface are still able to interact through motility or diffusion of metabolites, making certain competitive adaptations more advantageous (Chao and Levin, 1981). In this experiment, the agar media allowed physical separation of individual cells in the inocula and constrained microbial populations to form colonies, potentially forcing interactions with neighboring colonies that could result in competition and inhibition of growth through secondary metabolite production (Chao and Levin, 1981). The randomness of which populations are in close proximity to one another could play a role in which members will thrive from the same inoculum source, affecting the outcome of which taxa are cultivated (Traxler et al., 2013).

Cultivation on agar plates warrants special consideration when discussing diversity. The relative distribution of populations on an agar surface affects their prevalence and the overall composition of the cultivated community. Further, the physiology and colony morphology of each population dictates their prevalence and that of the surrounding populations on an agar surface. Colony size, especially surface area, is impacted by traits such as growth rate, surface motility (i.e., gliding motility, swarming), presence of inhibitory metabolites, and the availability of resources. These differences in biomass distribution were important when deciding how to measure diversity, a critical metric for assessing cultivation on an agar surface. Unlike liquid cultures, varying biomass (i.e. colony size) from different taxa that might produce an equal

number of colonies can over-estimate the relative abundance of that taxon, studies relying on relative abundance to characterize diversity should consider this (Medina et al., 2017, Cummings et al., 2020). If a population (A) has more proficient growth and forms a larger colony than population (B), DNA extracted from each population would indicate that taxon A was more abundant in the original inoculum, incorrectly representing the evenness of the original inoculum. This is especially true when relative abundance is calculated based on equimolar DNA concentration during library preparation. To mitigate any biased evenness, species richness and Jaccard distances were used to measure alpha and beta diversity, respectively.

In this data set, about half of the ASVs detected after cultivation were not detected in the molecular controls. This same phenomenon was observed by Zehavi et al. when cultivated rumen OTUs outnumbered the OTUs detected in the original rumen sample and its dilutions (1,012 out of 1,698) (Zehavi et al., 2018). This discrepancy may speak to the power of cultivation relative to the power of direct molecular analyses in describing the diversity of a community. However, both of these approaches come with their own caveats. Cultivation makes assumptions about an organism's ability to grow in the laboratory, while sequencing is dependent on methodology, sequencing chemistry and sampling depth. In this study, rarefaction curves indicated that sequencing efforts were sufficient (Figure S5). One potential explanation for the large proportion of ASVs not detected in the controls could be that in combination with the different media and treatments, we were able to give some of the less abundant microorganisms a growth advantage. For example, members of the bacterial family Wohlfarhtimonodaceae and members of the yeast family Metaschinikowiceae, were more associated with the addition of streptomycin to media (Table 1), and not detected in the

molecular controls (Figure 5). Streptomycin may have selected against the more abundant or more competitive organisms, allowing taxa from these families to grow. Additionally, observing Staphylococcaceae to be significantly associated with the ethanol treatment was unintended, as we anticipated selecting for spore formers. However, this was not surprising as ethanol has been shown to increase biofilm formation in staphylococcal species (Luther et al., 2015). Lastly and in regard to high proportions of ASVs in cultivated samples, systemic contamination and the presence of spurious ASVs from the PBS should be addressed. Similar to Zehavi et al., we believe this is unlikely, as we observed no growth on the plates incubated with PBS, nor were there any identifiable patterns among more rare taxa that would skew the data (Zehavi et al., 2018). Further, because presence absence was used to describe ASVs, we would anticipate this having a large effect if spurious ASVs were present across all samples, making more rare ASVs have equal weight to abundant ASVs.

Cultivation is critical to answer broad challenges of microbial ecology like deciphering microbial metabolisms or specific challenges like obtaining new isolates for antibiotic discovery. The study of pure cultures remains the best approach to comprehensively describe an organism's physiological and metabolic properties, and efforts to optimize cultivation strategy have proven successful (Connon and Giovannoni, 2002; Stevenson et al., 2004; Marmann et al., 2014; Laiger et al., 2015; Oberhardt et al., 2015; Yang et al., 2015; Browne et al., 2016; Henson et al., 2016; Berdy et al., 2017; Bartelme et al., 2020). Our data adds to previous studies (Medina et al., 2017; Zehavi et al., 2018) that also sought to assess how medium type and treatment could increase cultivated diversity, but adds the goal of "casting a wider net" to increase the diversity of bioactive isolates. We were able to add to our library of bioactive isolates and show that

microbiomes from roadkill mammals are a useful source of bioactivity, as observed in Motley et al. Lastly, we showed that through a thoughtful cultivation approach driven by bulk molecular analysis of cultivated taxa, we could cultivate a larger proportion of the diversity within an inoculum and even taxa not detected in the molecular controls. This work adds to the growing assessment of cultivation strategies using newer tools, like high-throughput sequencing, and shows how these methods can be applied to drug discovery efforts. Cultivation is influenced by many factors, and this work highlights some of those intricacies, like variability and stochasticity. Cultivation is complex and challenging but cultivating in combination with molecular tools can expand the observed diversity of an environment and its community.

Tables and Figures

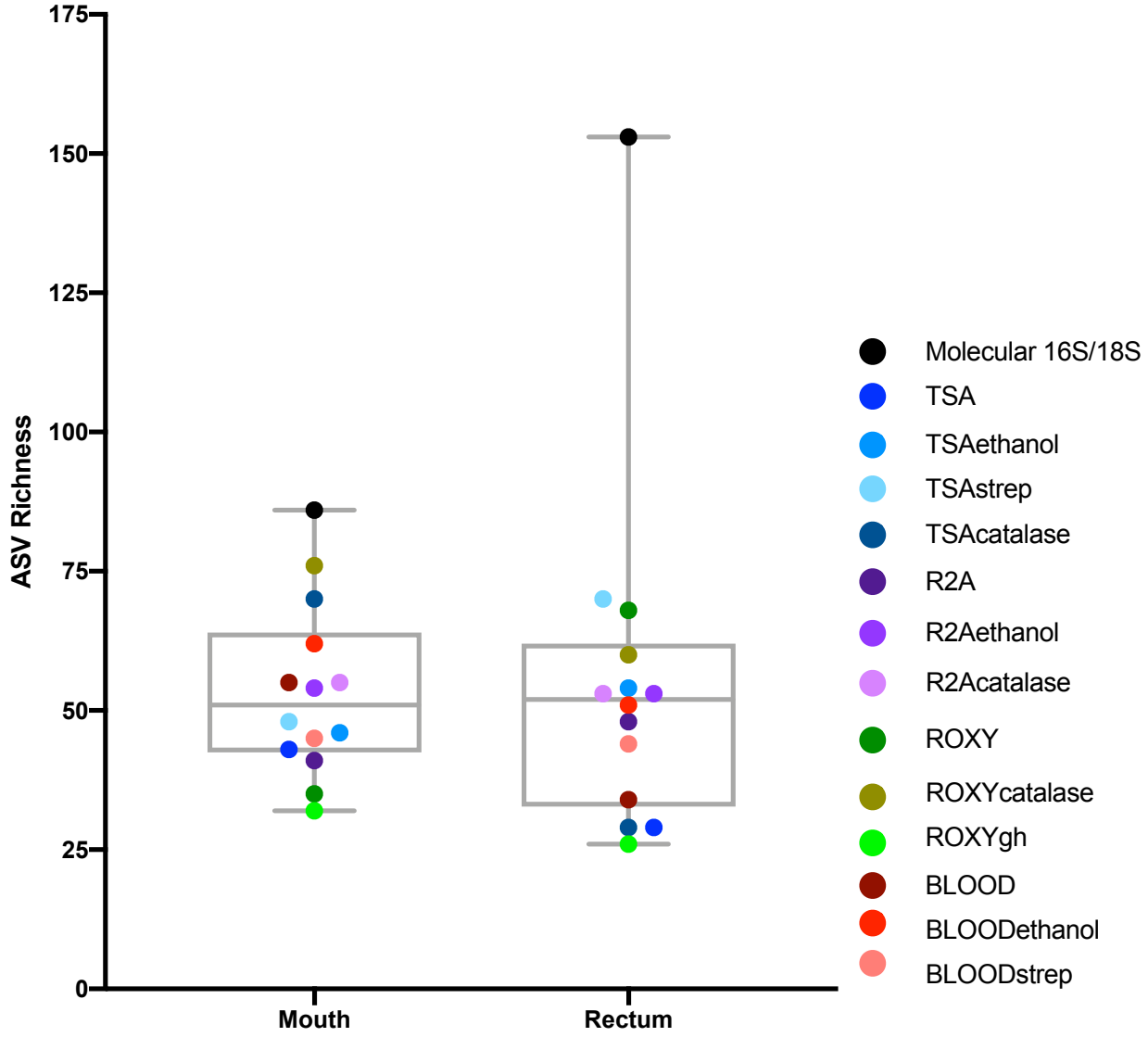


Figure 1: Alpha diversity of the mouth and rectum. For the controls, richness, measured using ASVs, was higher in the rectum than the mouth. For cultivated communities, richness varied based on medium type and treatment. Points represent each medium or treatment type, indicated by color (n=3). Molecular 16S/18S refers to the directly sequenced inoculum sample (n=3).

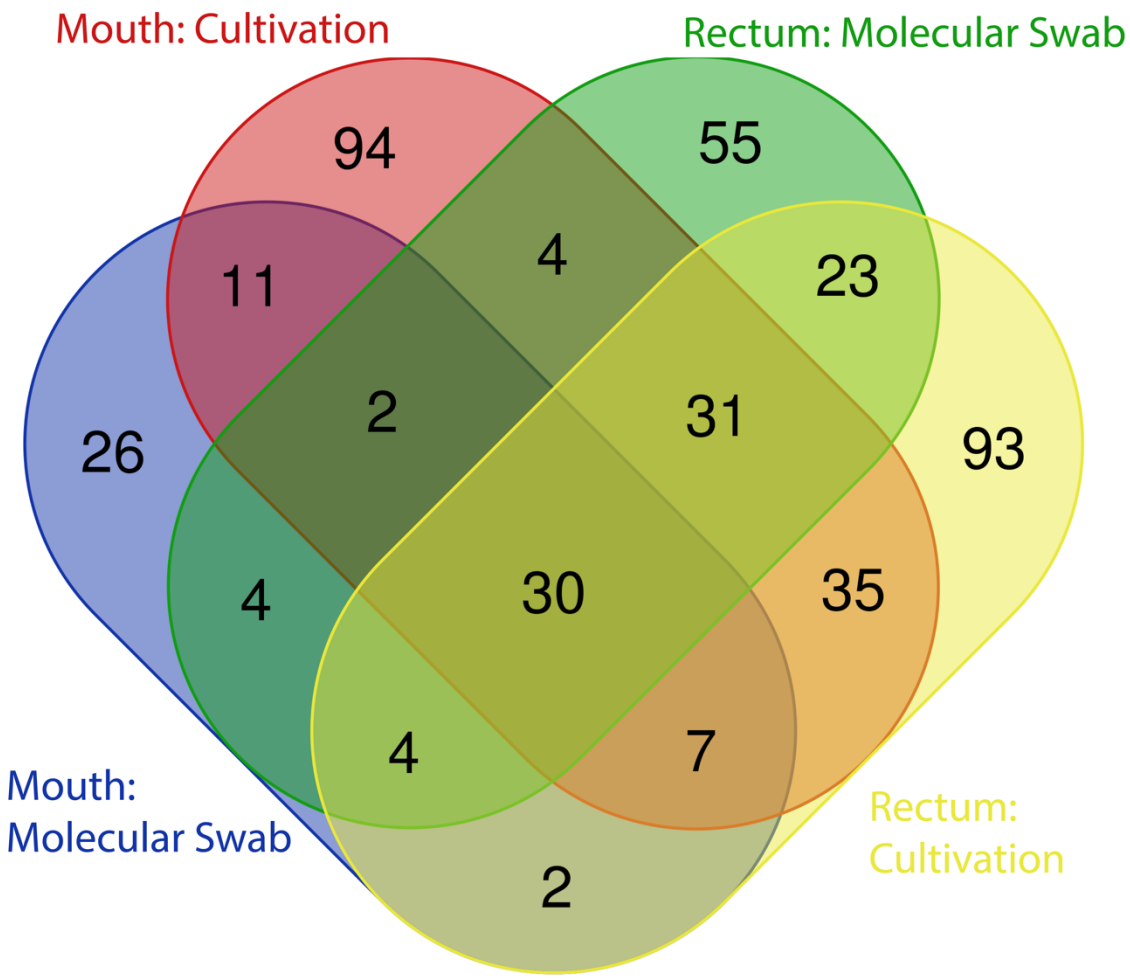


Figure 2: Distribution of ASVs between cultivation and molecular controls. The majority of ASVs detected in the original inocula were shared between the mouth and rectum, while approximately half of all ASVs detected through cultivation were not detected in the directly sequenced samples. Created with http://bioinformatics.psb.ugent.be/webtools/Venn/.

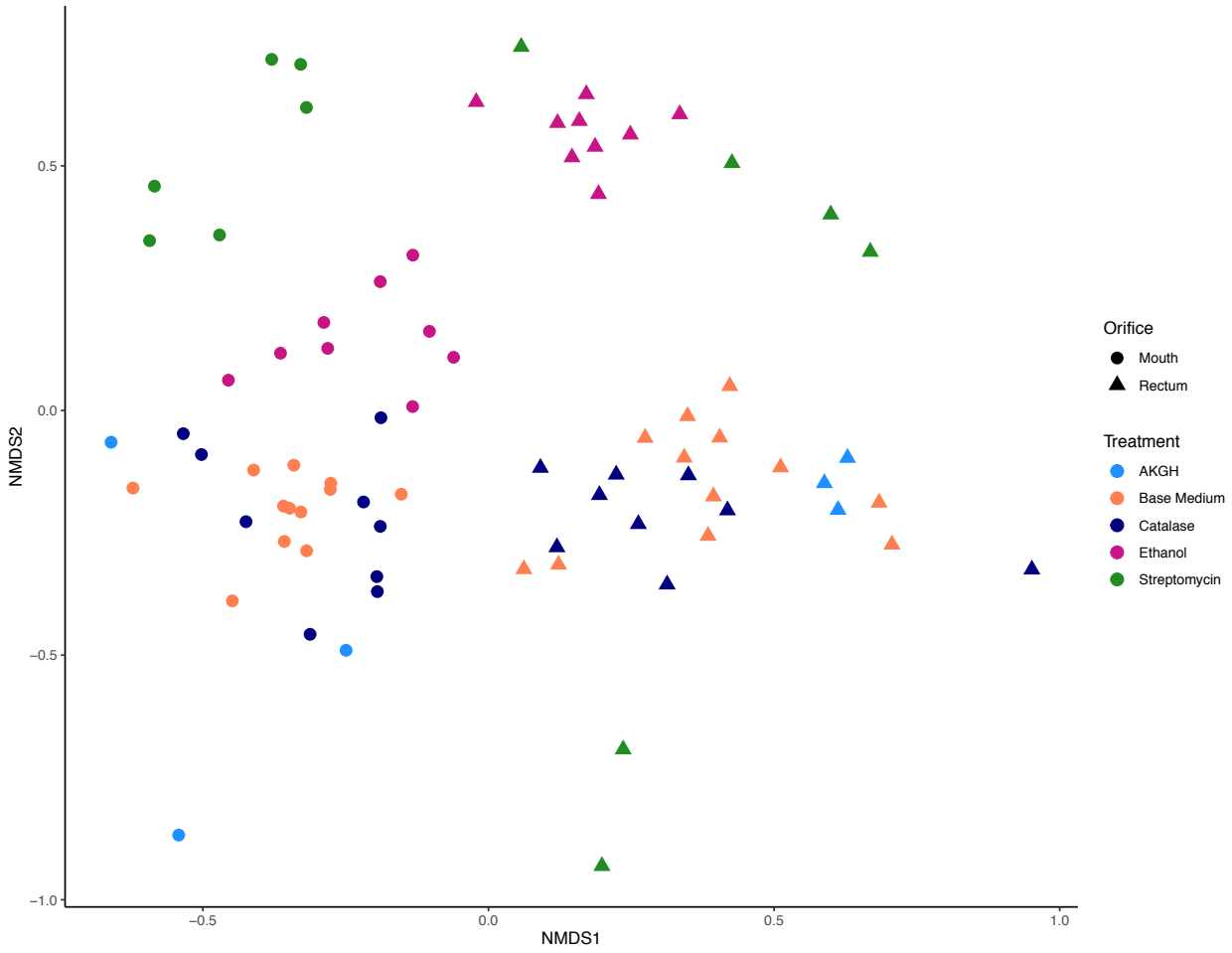
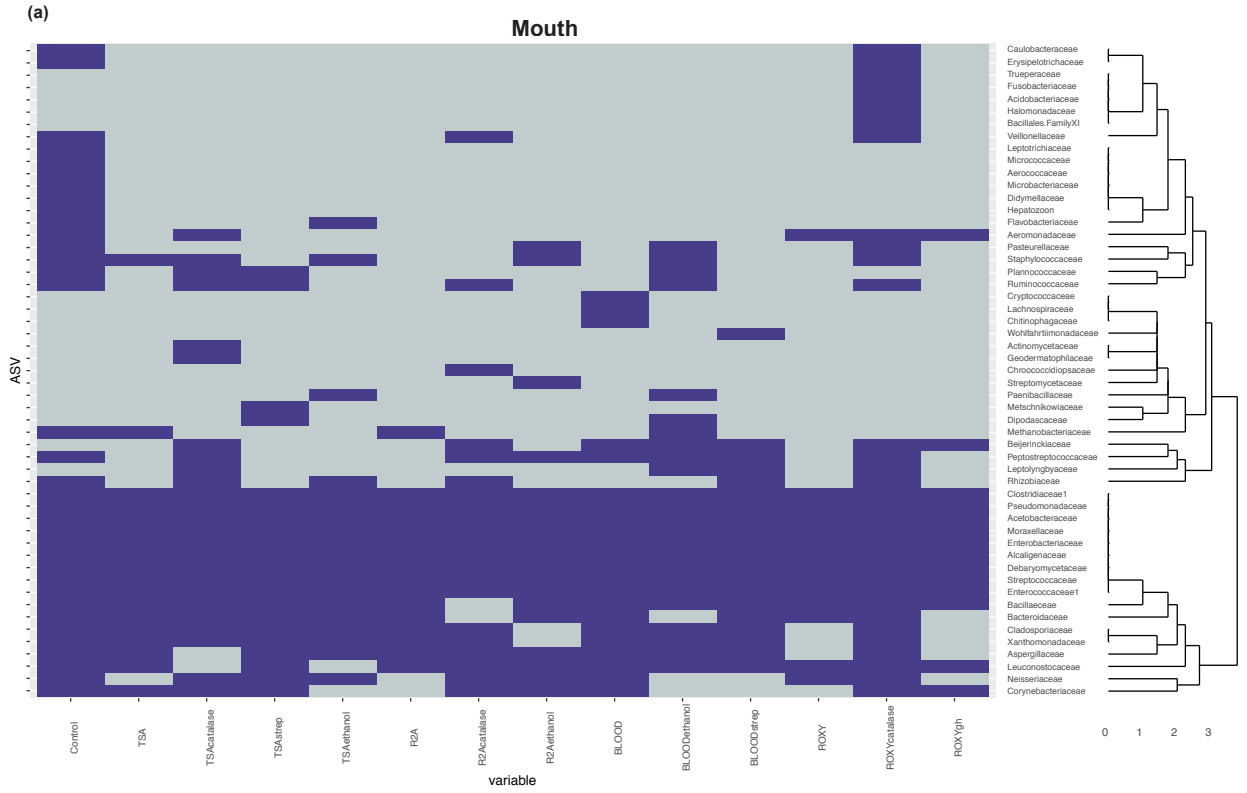


Figure 3: Beta diversity measured in Jaccard distance and NMDS ordination. Differences in microbial community structure were driven by orifice and treatment, specifically ethanol and streptomycin. Shape corresponds to orifice sampled, while color differentiates between treatments of the medium or inoculum.


(a)
Mouth
ASV
variable
Control
TSA
TSAcatalase
TSAstrep
TSAethanol
R2A
R2Acatalase
R2Aethanol
BLOOD
BLOODethanol
BLODDstrep
ROXY
ROXYcatalase
ROXYgh
Caulobacteraceae
Erysipelotrichaceae
Trueperaceae
Fusobacteriaceae
Acidobacteriaceae
Halomonadaceae
Bacillales.FamilyXI
Veillonellaceae
Leptotrichiaceae
Micrococcaceae
Aerococcaceae
Microbacteriaceae
Didymellaceae
Hepatozoon
Flavobacteriaceae
Aeromonadaceae
Pasteurellaceae
Staphylococcaceae
Plannococcaceae
Ruminococcaceae
Cryptococcaceae
Lachnospiraceae
Chitinophagaceae
Wohlfahrtiimonadaceae
Actinomycetaceae
Geodermatophilaceae
Chroococcidiopsaceae
Streptomycetaceae
Paenibacillaceae
Metschnikowiaceae
Dipodascaceae
Methanobacteriaceae
Beijerinckiaceae
Peptostreptococcaceae
Leptolyngbyaceae
Rhizobiaceae
Clostridiaceae1
Pseudomonadaceae
Acetobacteraceae
Moraxellaceae
Enterobacteriaceae
Alcaligenaceae
Debaryomycetaceae
Streptococcaceae
Enterococcaceae1
Bacillaeceae
Bacteroidaceae
Cladosporiaceae
Xanthomonadaceae
Aspergillaceae
Leuconostocaceae
Neisseriaceae
Corynebacteriaceae
0
1
2
3

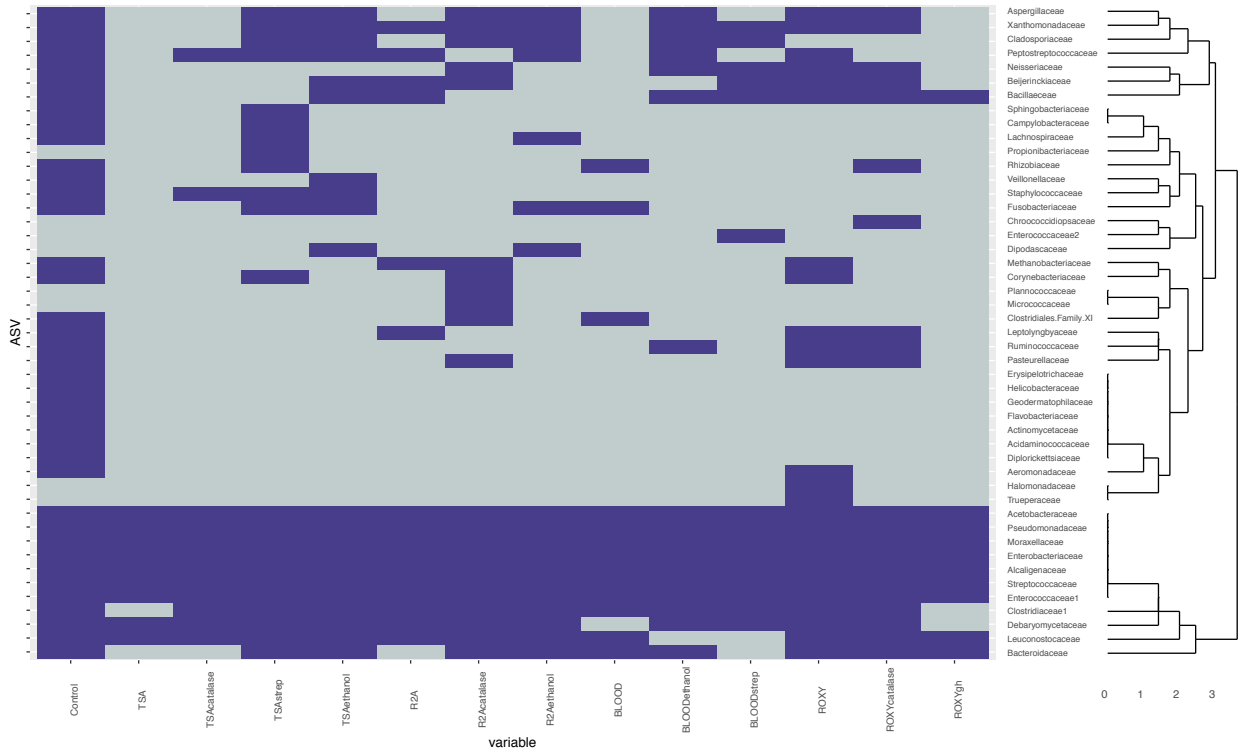

(b)
Rectum
ASV
variable
Control
TSA
TSAcatalase
TSAstrep
TSAethanol
R2A
R2Acatalase
R2Aethanol
BLOOD
BLOODethanol
BLODDstrep
ROXY
ROXYcatalase
ROXYgh
Aspergillaceae
Xanthomonadaceae
Cladosporiaceae
Peptostreptococcaceae
Neisseriaceae
Beijerinckiaceae
Bacillaeceae
Sphingobacteriaceae
Campylobacteraceae
Lachnospiraceae
Propionibacteriaceae
Rhizobiaceae
Veillonellaceae
Staphylococcaceae
Fusobacteriaceae
Chroococcidiopsaceae
Enterococcaceae2
Dipodascaceae
Methanobacteriaceae
Corynebacteriaceae
Plannococcaceae
Micrococcaceae
Clostridiales.Family.XI
Leptolyngbyaceae
Ruminococcaceae
Pasteurellaceae
Erysipelotrichaceae
Helicobacteraceae
Geodermatophilaceae
Flavobacteriaceae
Actinomycetaceae
Acidaminococcaceae
Diplorickettsiaceae
Aeromonadaceae
Halomonadaceae
Trueperaceae
Acetobacteraceae
Pseudomonadaceae
Moraxellaceae
Enterobacteriaceae
Alcaligenaceae
Streptococcaceae
Enterococcaceae1
Clostridiaceae1
Debaryomycetaceae
Leuconostocaceae
Bacteroidaceae
0
1
2
3

Figure 4: Heatmaps describing presence absence of taxonomic families observed in the (A) mouth and (B) rectum. Purple indicates that a family was observed while grey indicates that a family was absent. The dendrograms highlight clusters of families observed during cultivation. Control refers to the families identified on the molecular control.

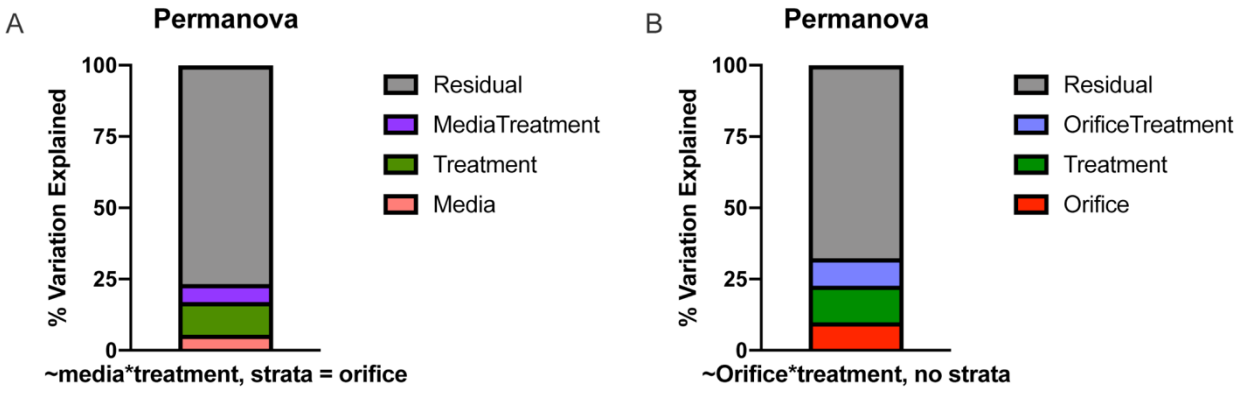


Figure 5: Percent of variation described by permanova models based on (A) media and treatment with orifice as a stratum, and (B) orifice and treatment. Largely, the differences in community structure during cultivation could not be explained by media, orifice, or treatment alone. In all graphs, residual variation (in grey), is the proportion of variance not explained by the model.

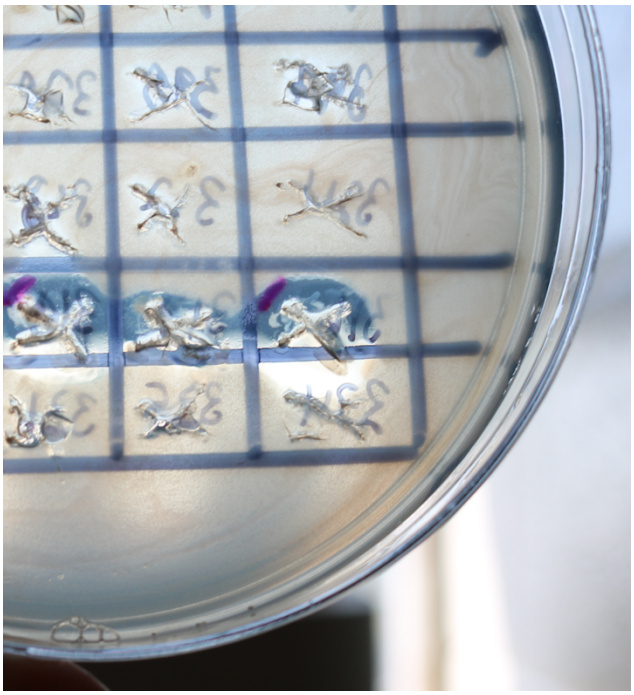
Table 1. Taxonomic associations with media type using Indicator Species Analysis

Grouping Factor	Condition	Indicator ASV Family	Index Value	p value
Orifice	Mouth	Aeromonadaceae	0.483	0.007
Media Type	ROXY	Aeromonadaceae	0.589	0.001
	Blood + TSA + R2A	Cladosporiaceae	0.604	0.026
Treatment	Ethanol	Staphylococcaceae	0.562	0.014
		Paenibacillaceae	0.408	0.041
	Streptomycin	Wohlfahrtiimonadaceae	0.408	0.034
		Metschnikowiaceae	0.408	0.030
	Ethanol + Streptomycin	Peptostreptococcaceae	0.669	0.003
	Catalase + Ethanol + Streptomycin	Aspergillaceae	0.692	0.004
		Cladosporiaceae	0.633	0.006
	Catalase + Ethanol + Streptomycin + No Treatment	Debaryomycetaceae	0.899	0.002
		Clostridiaceae 1	0.860	0.006

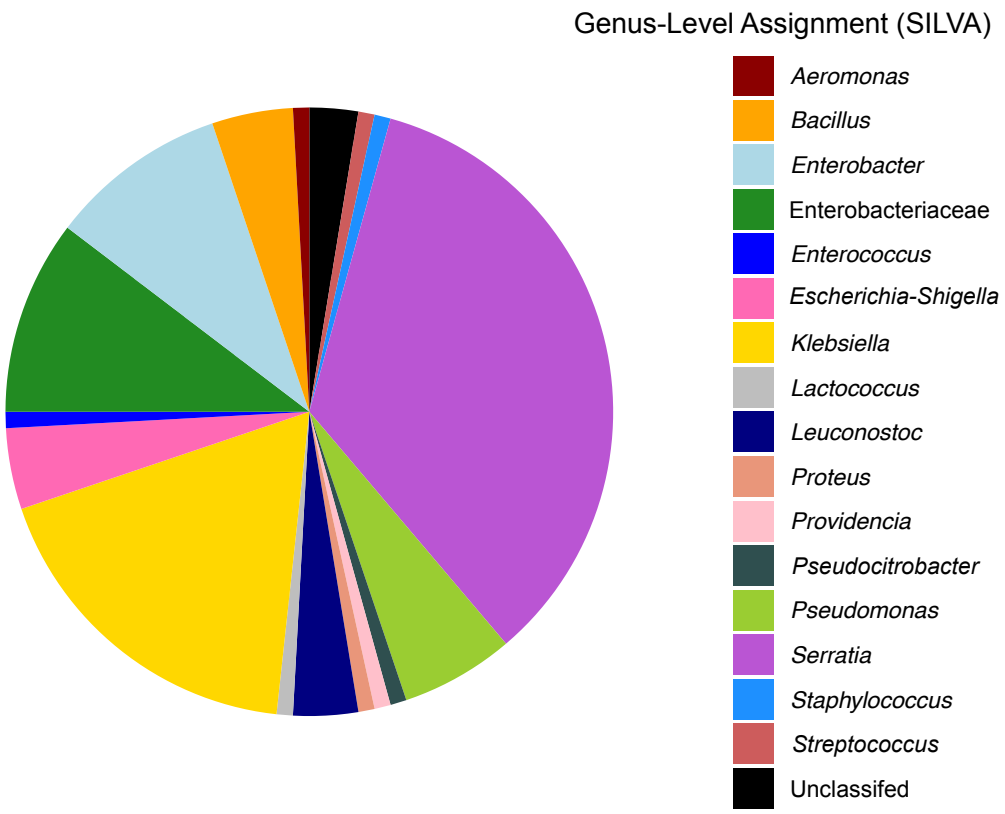
Table 2. Cultivated bioactive isolates

Isolate	SILVA Taxonomy	Orifice	Medium	Activity Against
RC2RCR2A30-212	*Bacillus* sp.	Rectum	R2A	*Klebsiella pneumoniae*
RC2MOR2A30-238	*Pseudomonas* sp.	Mouth	R2A	*Candida albicans*
RC2MORGH30-018	Enterobacteriaceae	Mouth	ROXYgh	*Enterococcus faecium*
RC2MOTSA30-050	*Enterococcus* sp.	Mouth	TSA	*Enterococcus faecium*
RC2MOTSA30-142	*Klebsiella* sp.	Mouth	TSA	*Klebsiella pneumoniae*

Supplementary Material



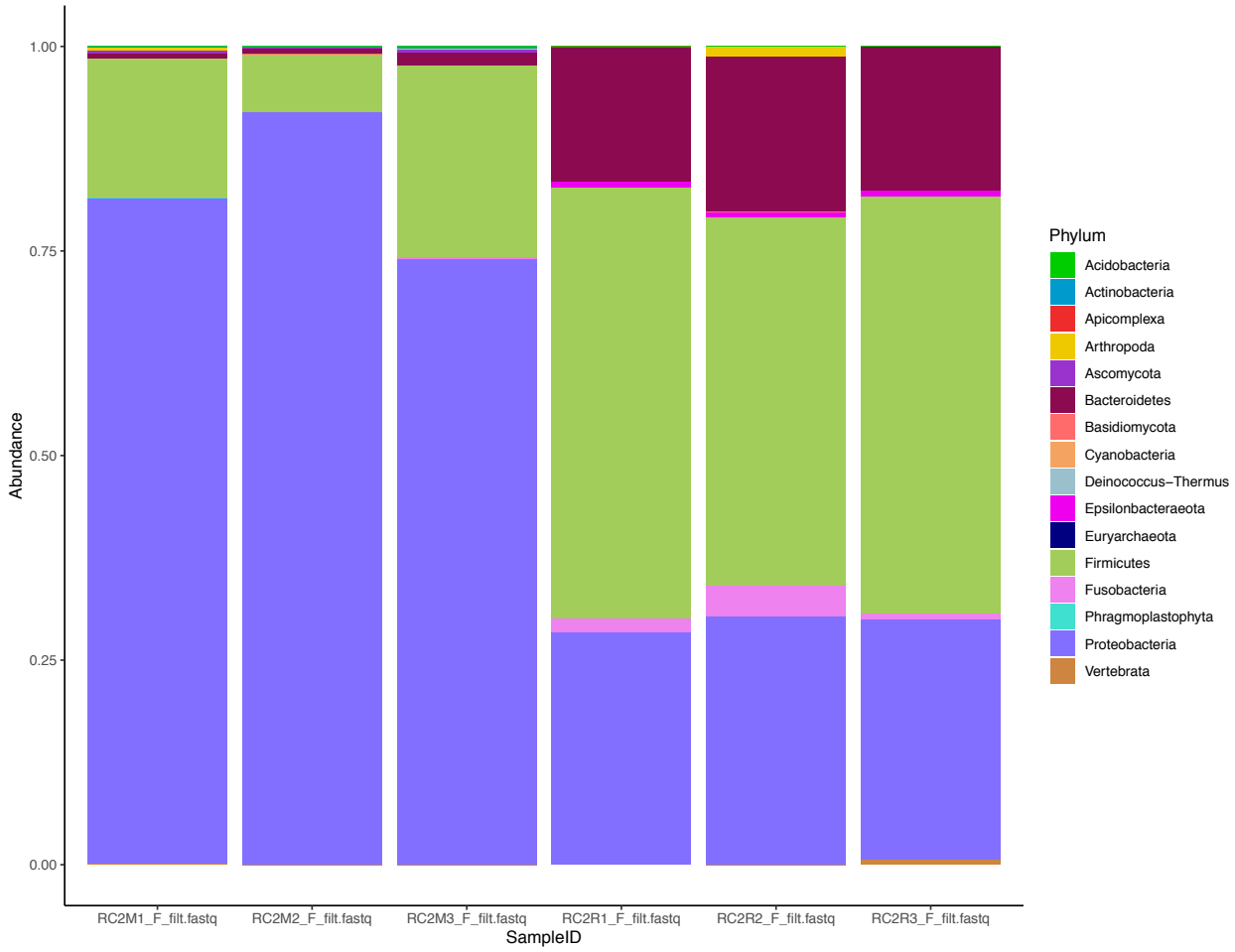
Supplemental Figure 1: Bioassay plate showing three isolates generating a zone of inhibition within pathogen overlay.



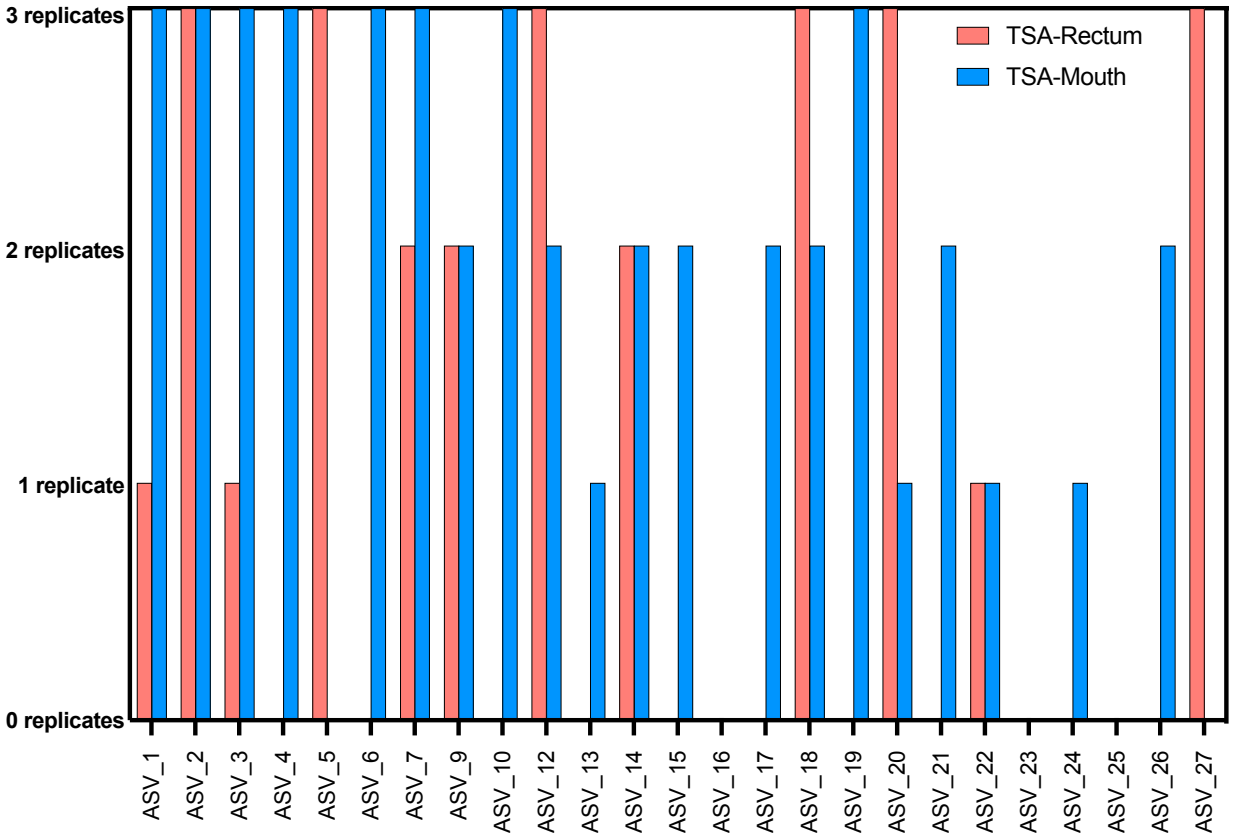
Supplemental Figure 2: Pie chart of 117 identified isolates at the genus level.

Supplemental Table 1. A quasi-factorial approach using targeted media and treatments for gut microorganisms.

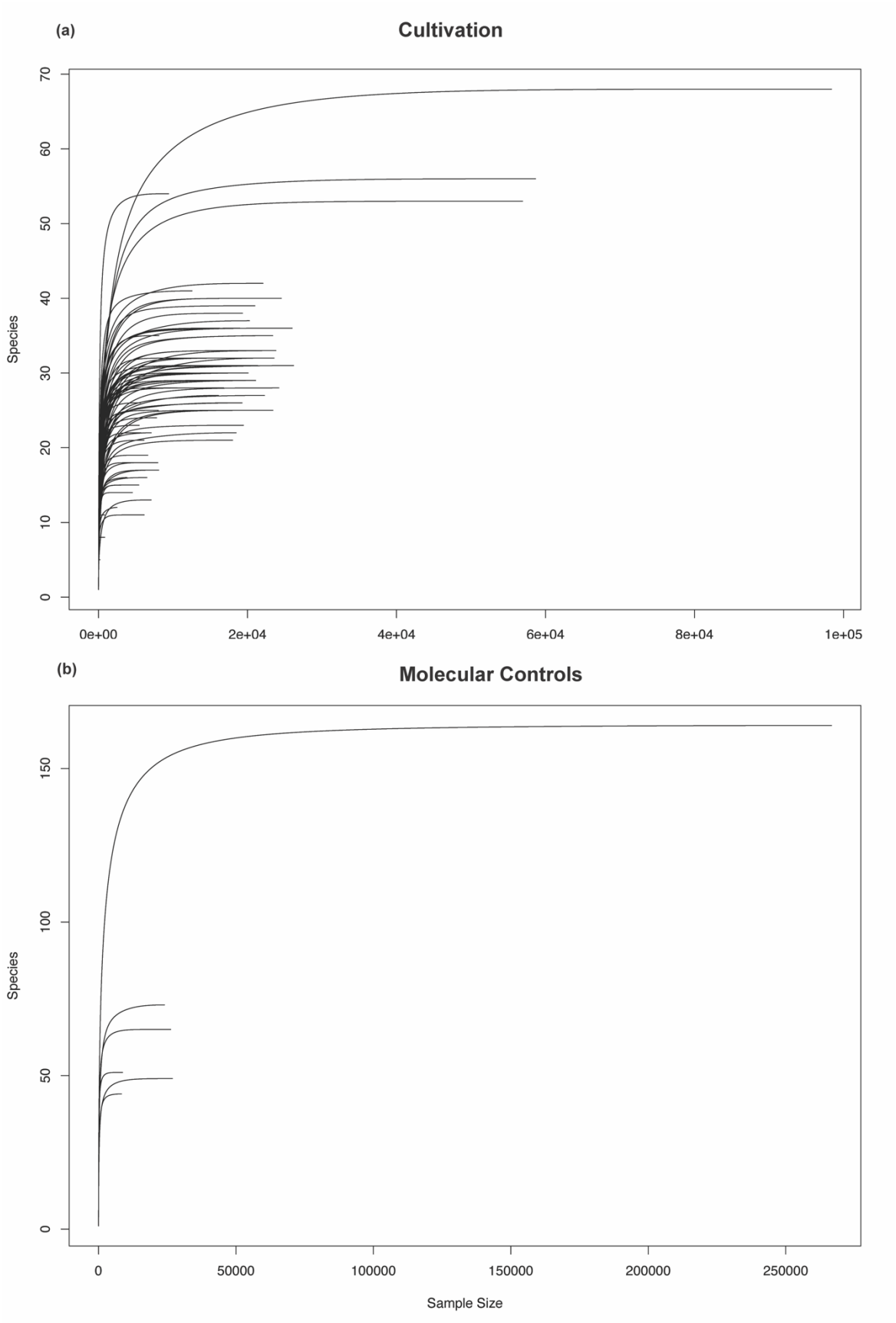
	Media			
Treatment/ Selection agents	0.1 TSA	ROXY	0.25 R2A	Blood Agar
None	X	X	X	X
Catalase	X	X	X	
Hemin and alpha-ketoglutarate		X		
70% ethanol	X		X	X
Streptomycin	X			X



Supplemental Figure 3: Phylum-level characterization of molecular controls, triplicates shown for mouth (RC2M) and rectum (RC2R).



Supplemental Figure 4: Bar chart describing the variability of consistent recovery of ASVs, data subset. Y axis indicates how many replicates of TSA plates in which the top 25 ASV were observed. Color corresponds to orifice.


(a)
Cultivation
Species
0
10
20
30
40
50
60
70
0e+00
2e+04
4e+04
6e+04
8e+04
1e+05
(b)
Molecular Controls
Species
0
50
100
150
0
50000
100000
150000
200000
250000
Sample Size

Supplemental Figure 5: Rarefaction curves showing plateaued sequencing depth of (A) cultivation samples and (B) molecular controls.

References

2019. Antibiotic resistance and threats in the United States 2019. *Center for Disease Control (CDC)*.

Albright, M.B.N., Chase, A.B., & Martiny, J.B.H. (2019). Experimental evidence that stochasticity contributes to bacterial composition and functioning in a decomposer community. *mBio 10*, 247–13. DOI: 10.1128/mBio.00568-19

Arora-Williams, K., Olesen, S.W., Scandella, B.P., Delwiche, K., Spencer, S.J., Myers, E.M., Abraham, S., Sooklal, A., & Preheim, S.P. (2018). Dynamics of microbial populations mediating biogeochemical cycling in a freshwater lake. *Microbiome 6*, 165–16. DOI: 10.1186/s40168-018-0556-7

Bartelme, R.P., Custer, J.M., Dupont, C.L., Espinoza, J.L., Torralba, M., Khalili, B., & Carini, P. (2020). Influence of substrate concentration on the culturability of heterotrophic soil microbes isolated by high-throughput dilution-to-extinction cultivation. *mSphere 5*, 489–15. DOI: 10.1128/mSphere.00024-20

Berdy, B., Spoering, A.L., Ling, L.L., & Epstein, S.S. (2017). In situ cultivation of previously uncultivable microorganisms using the ichip. *Nat Protoc 12*, 2232–2242. DOI: 10.1038/nprot.2017.074

Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., Goulding, D., & Lawley, T.D. (2016). Culturing of "unculturable" human microbiota reveals novel taxa and extensive sporulation. *Nature 533*, 543–546. DOI: 10.1038/nature17645

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., & 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., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods 7*, 335–336. DOI: 10.1038/nmeth.f.303

Chao, L. & Levin, B.R. (1981). Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proceedings of the National Academy of Sciences 78*, 6324–6328. DOI: 10.1073/pnas.78.10.6324

Clardy, J., Fischbach, M.A., & Walsh, C.T. (2006). New antibiotics from bacterial natural products. *Nat Biotech 24*, 1541–1550. DOI: 10.1038/nbt1266

Connon, S.A. & Giovannoni, S.J. (2002). High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl Environ Microbiol 68*, 3878–3885. DOI: 10.1128/AEM.68.8.3878-3885.2002

Cummings, L.A., Hoogestraat, D.R., Rassoulian-Barrett, S.L., Rosenthal, C.A., Salipante, S.J., Cookson, B.T., & Hoffman, N.G. (2020). Comprehensive evaluation of complex polymicrobial specimens using next generation sequencing and standard microbiological culture. *Sci Rep 10*, 1–12. DOI: 10.1038/s41598-020-62424-x

Davis, K.E.R., Joseph, S.J., & Janssen, P.H. (2005). Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. *Appl Environ Microbiol 71*, 826–834. DOI: 10.1128/AEM.71.2.826-834.2005

De Cáceres, M. & Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology 90*, 3566–3574. DOI: 10.1890/08-1823.1

De Cáceres, M., Sol, D., Lapiedra, O., & Legendre, P. (2011). A framework for estimating niche metrics using the resemblance between qualitative resources. *Oikos 120*, 1341–1350. DOI: 10.1111/j.1600-0706.2011. 19679.x

de Vos, M.G.J., Zagorski, M., McNally, A., & Bollenbach, T. (2017). Interaction networks, ecological stability, and collective antibiotic tolerance in polymicrobial infections. *Proceedings of the National Academy of Sciences 114*, 10666–10671. DOI: 10.1073/pnas.1713372114

Ding, Z-G., Li, M-G., Zhao, J-Y., Ren, J., Huang, R., Xie, M-J., Cui, X-L., Zhu, H-J., & Wen, M-L. (2010). Naphthospironone A: An unprecedented and highly functionalized polycyclic metabolite from an alkaline mine waste extremophile. *Chem Eur J 16*, 3902–3905. DOI: 10.1002/chem.200903198

Dione, N., Khelaifia, S., La Scola, B., Lagier, J.C., & Raoult, D. (2016). A quasi-universal medium to break the aerobic/anaerobic bacterial culture dichotomy in clinical microbiology. *Clinical Microbiology and Infection 22*, 53–58. DOI: 10.1016/j.cmi.2015.10.032

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science 14*, 927–930. DOI:10.1111/j.1654-1103.2003.tb02228.x

Donia, M.S., Cimermancic, P., Schulze, C.J., Brown, L.C.W., Martin, J., Mitreva, M., Clardy, J., Linington, R.G., & Fischbach, M.A. (2014). A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell 158*, 1402–1414. DOI: 10.1016/j.cell.2014.08.032

Eloe-Fadrosh, E.A., Ivanova, N.N., Woyke, T., & Kyrpides, N.C. (2016). Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nature Microbiology 1*, 15032–4. DOI: 10.1038/nmicrobiol.2015.32

Fuller, A.T., Mellows, G., Woolford, M., Banks, G.T., Barrow, K.D., & Chain, E.B. (1971). Pseudomonic Acid: An antibiotic produced by *Pseudomonas fluorescens*. *Nature 234*, 416–417. DOI: 10.1038/234416a0

Gilbert, J.A., Quinn, R.A., Debelius, J., Xu, Z.Z., Morton, J., Garg, N., Jansson, J.K., Dorrestein, P.C., & Knight. R. (2016). Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature 535*, 94–103. DOI: 10.1038/nature18850

Guo, J., Fu, X., Liao, H., Hu, Z., Long, L., Yan, W., Ding, Y., Zha, L., Guo, Y., Yan, J., Chang, Y., & Cai, J. (2016). Potential use of bacterial community succession for estimating post-mortem interval as revealed by high-throughput sequencing. *Sci Rep 6*, 1–11. DOI: 10.1038/srep24197

Halfvarson, J., Brislawn, C.J., Lamendella, R., Vázquez-Baeza, Y., Walters, W.A., Bramer, L.M., D'Amato, M., Bonfiglio, F., McDonald, D., González, A., McClure, E.E., Dunklebarger, M.F., Knight, R., & Jansson, J.K. (2017). Dynamics of the human gut microbiome in inflammatory bowel disease. *Nature Microbiology 2*, 1–7. DOI: 10.1038/nmicrobiol.2017.4

Hamdache, A., Lamarti, A., Aleu, J., & Collado, I.G. (2011). Non-peptide metabolites from the genus *Bacillus*. *J Nat Prod 74*, 893–899. DOI:10.1021/np100853e

Helfrich, E.J.N., Vogel, C.M., Ueoka, R., Schafer, M., Ryffel, F., Muller, D.B., Probst, S., Kreuzer, M., Piel, J., & Vorholt, J.A. (2018). Bipartite interactions, antibiotic production and biosynthetic potential of the *Arabidopsis* leaf microbiome. *Nature Microbiology 3*, 909–919. DOI: 10.1038/s41564-018-0200-0

Henson, M.W., Pitre, D.M., Weckhorst, J.L., Lanclos, V.C., Webber, A.T., & Thrash, J.C. (2016). Artificial seawater media facilitate cultivating members of the microbial majority from the Gulf of Mexico. *mSphere 1*, e00028–16. DOI: 10.1128/mSphere.00028-16

Hibbing, M.E., Fuqua, C., Parsek, M.R., & Peterson, S.B. (2009). Bacterial competition: Surviving and thriving in the microbial jungle. *Nature Reviews Microbiology 8*, 15–25. DOI: 10.1038/nrmicro2259

Hoffmann, T., Krug, D., Bozkurt, N., Duddela, S., Jansen, R., Garcia, R., Gerth, K., Steinmetz, H., & Muller, R,. (2018). Correlating chemical diversity with taxonomic distance for discovery of natural products in myxobacteria. *Nature Communications 9*, 1–10. DOI: 10.1038/s41467-018-03184-1

Jensen, V. (1968). The plate count technique. In TRG Gray & D Parkinson (Eds.), The ecology of soil bacteria (pp. 158–170). Liverpool University Press.

Justice, N.B., Sczesnak, A., Hazen, T.C., & Arkin, A.P. (2017). Environmental selection, dispersal, and organism interactions shape community assembly in high-throughput enrichment culturing. *Appl Environ Microbiol 83*, e01253–17–16. DOI: 10.1128/AEM.01253-17

Kawasaki, K. & Kamagata, Y. (2017). Phosphate-catalyzed hydrogen peroxide formation from agar, gellan, and κ-carrageenan and recovery of microbial cultivability via catalase and pyruvate. *Appl Environ Microbiol 83*, e01366–17. DOI: 10.1128/AEM.01366-17

Kim, S., Kang, I., Seo, J-H., & Cho, J-C. (2019). Culturing the ubiquitous freshwater actinobacterial acI lineage by supplying a biochemical "helper" catalase. *The ISME Journal 13*, 2252–2263. \DOI: 10.1038/s41396-019-0432-x

Kopke, B., Wilms, R., Engelen, B., Cypionka, H., & Sass, H. (2005). Microbial diversity in coastal subsurface sediments: A cultivation approach using various electron acceptors and substrate gradients. *Appl Environ Microbiol 71*, 7819–7830. DOI: 10.1128/AEM.71.12.7819-7830.2005

Kraus, E.A., Beeler, S.R., Mors, R.A., Floyd, J.G., GeoBiology 2016, Stamps, B.W., Nunn, H.S., Stevenson, B.S., Johnson, H.A., Shapiro, R.S., Loyd, S.J., Spear, J.R., & Corsetti, F.A. (2018). Microscale biosignatures and abiotic mineral authigenesis in Little Hot Creek, California. *Front Microbiol 9*, 544–13. DOI: 10.3389/fmicb.2018.00997

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution 35*, 1547–1549. DOI: 10.1093/molbev/msy096

La Scola, B., Khelaifia, S., Lagier, J.C., & Raoult, D. (2014). Aerobic culture of anaerobic bacteria using antioxidants: a preliminary report. *Eur J Clin Microbiol Infect Dis 33*, 1781–1783. DOI: 10.1007/s10096-014-2137-4

Lagier, J-C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., Bittar, F., Fournous, G., Gimenez, G., Maraninchi, M., Trape, J-F., Koonin, E.V., La Scola, B., & Raoult, D.

(2012). Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect 18*, 1185-1193. DOI: 10.1111/1469-0691.12023

Lagier, J-C., Hugon, P., Khelaifia, S., Fournier, P-E., La Scola, B., & Raoult, D. (2015). The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev 28*, 237–264. DOI: 10.1128/CMR.00014-14

Lloyd, K.G., Steen, A.D., Ladau, J., Yin, J., & Crosby, L. (2018). Phylogenetically novel uncultured microbial cells dominate Earth microbiomes. *mSystems 3*, 431–12. DOI: 10.1128/mSystems.00055-18

Luther, M.K., Bilida, S., Mermel, L.A., & LaPlante, K.L. (2015). Ethanol and isopropyl alcohol exposure increases biofilm formation in *Staphylococcus aureus* and *Staphylococcus epidermidis. Infect Dis Ther 4*, 219–226. DOI: 10.1007/s40121-015-0065-y

Marmann, A., Aly, A., Lin, W., Wang, B., & Proksch, P. (2014). Co-cultivation—A powerful emerging tool for enhancing the chemical diversity of microorganisms. *Marine Drugs 12*, 1043–1065. DOI: 10.3390/md12021043

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

Medina, D., Walke, J.B., Gajewski, Z., Becker, M.H., Swartwout, M.C., & Belden, L.K. (2017). Culture media and individual hosts affect the recovery of culturable bacterial diversity from amphibian skin. *Front Microbiol 8*, 32–15. DOI: 10.3389/fmicb.2017.01574

Metcalf, J.L., Wegener Parfrey, L., González, A., Lauber, C.L., Knights, D., Ackermann, G., Humphrey, G.C., Gebert, M.J., Van Treuren, W., Berg-Lyons, D., Keepers, K., Guo, Y., Bullard, J., Fierer, N., Carter, D.O., & Knight, R. (2013). A microbial clock provides an

accurate estimate of the postmortem interval in a mouse model system. *eLife 2*, 403–19. DOI: 10.7554/eLife.01104

Motley, J.L., Stamps, B.W., Mitchell, C.A., Thompson, A.T., Cross, J., You, J., Powell, D.R., Stevenson, B.S., & Cichewicz, R.H. (2017). Opportunistic sampling of roadkill as an entry point to accessing natural products assembled by bacteria associated with non-anthropoidal mammalian microbiomes. *J Nat Prod 80*, 598–608. DOI: 10.1021/acs.jnatprod.6b00772

Newman, D.J. & Cragg, G.M. (2016). Natural products as sources of new drugs from 1981 to 2014. *J Nat Prod 79*, 629–661. DOI: 10.1021/acs.jnatprod.5b01055

Nichols, D., Cahoon, N., Trakhtenberg, E.M., Pham, L., Mehta, A., Belanger, A., Kanigan, T., Lewis, K., & Epstein, S.S. (2010). Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial species. *Appl Environ Microbiol 76*, 2445–2450. DOI: 10.1128/AEM.01754-09

Oberhardt, M.A., Zarecki, R., Gronow, S., Lang, E., Klenk, H-P., Gophna, U., & Ruppin, E. (2015). Harnessing the landscape of microbial culture media to predict new organism-media pairings. *Nature Communications 6*, 1–14. DOI: 10.1038/ncomms9493

Parada, A.E., Needham, D.M., & Fuhrman, J.A. (2015). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol 18*, 1403–1414. DOI: 10.1111/1462-2920.13023

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research 41*, D590–6. DOI: 10.1093/nar/gks1219

Raaijmakers, J.M., Weller, D.M., & Thomashow, L.S. (1997). Frequency of antibiotic-producing *Pseudomonas* spp. in natural environments. *Appl Environ Microbiol 63*, 881–887. DOI: 10.1128/AEM.63.3.881-887.1997

Rangseekaew, P. & Pathom-aree, W. (2019). Cave actinobacteria as producers of bioactive metabolites. *Front Microbiol 10*, 381–11. DOI: 10.3389/fmicb.2019.00387

Reasoner, D.J. & Geldreich, E.E. (1985). A new medium for the enumeration and subculture of bacteria from potable water. *Appl Environ Microbiol 49*, 1–7. DOI: 10.1128/AEM.49.1.1-7.1985

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., & Mesirov, J.P. (2011). Integrative genomics viewer. *Nature Biotechnology 29*, 24–26. DOI: 10.1038/nbt.1754

Schatz, A., Bugle, E., & Waksman, S.A. (1944). Streptomycin, a substance exhibiting antibiotic activity against Gram-positive and Gram-negative bacteria. *Experimental Biology and Medicine 55*, 66–69. DOI: 10.1097/01.blo.0000175887.98112.fe

Schloss, P.D., Girard, R.A., Martin, T., Edwards, J., & Thrash, J.C. (2016). Status of the Archaeal and Bacterial census: An update. *mBio 7*, 995–10. DOI: 10.1128/mBio.00201-16

Singh, B.K., Bardgett, R.D., Smith, P., & Reay, D.S. (2010). Microorganisms and climate change: terrestrial feedbacks and mitigation options. *Nature Reviews 8*, 779–790. DOI: 10.1038/nrmicro2439

Smanski, M.J., Schlatter, D.C., & Kinkel, L.L. (2015). Leveraging ecological theory to guide natural product discovery. *Journal of Industrial Microbiology & Biotechnology 43*, 115–128. DOI: 10.1007/s10295-015-1683-9

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., Arrieta, J.M., & Herndl, G.J. (2006). Microbial diversity in the deep sea and the underexplored ‘‘rare biosphere’’. *Proceedings of the National Academy of Sciences 103*, 12115–12120. DOI: 10.1073/pnas.0605127103

Staley, J.T. & Konopka, A. (1985). Meaurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol 39*, 321–346. DOI: 10.1146/annurev.mi.39.100185.001541

Stevenson, B.S., Eichorst, S.A., Wertz, J.T,, Schmidt, T.M., & Breznak, J,A. (2004). New strategies for cultivation and detection of previously uncultured microbes. *Appl Environ Microbiol 70*, 4748–4755. DOI: 10.1128/AEM.70.8.4748-4755.2004

Stierle, A.A., Stierle, D.B., & Kelley, K. (2006). Berkelic acid, a novel spiroketal with selective anticancer activity from an acid mine waste fungal extremophile. *Journal of Organic Chemistry 71*, 5357–5360. DOI: 10.1021/jo060018d

Sumi, C.D., Yang, B.W., Yeo, I-C., & Hahm, Y.T. (2015). Antimicrobial peptides of the genus *Bacillus*: a new era for antibiotics. *Can J Microbiol 61*, 93–103. DOI: 10.1139/cjm-2014-0613

Tortorella, E., Tedesco, P., Palma Esposito, F., January, G., Fani, R., Jaspars, M., & de Pascale, D. (2018). Antibiotics from deep-sea microorganisms: Current discoveries and perspectives. *Marine Drugs 16*, 355–16. DOI: 10.3390/md16100355

Traxler, M.F., Watrous, J.D., Alexandrov, T., Dorrestein, P.C., & Kolter, R. (2013). Interspecies interactions stimulate diversification of the *Streptomyces coelicolor* secreted metabolome. *mBio 4*, e00459–13. DOI: 10.1128/mBio.00459-13

van der Meij, A., Worsley, S.F., Hutchings, M.I., & van Wezel, G.P. (2017). Chemical ecology of antibiotic production by actinomycetes. *FEMS Microbiology Reviews 41*, 392–416. DOI: 10.1093/femsre/fux005

Yang, S-J., Kang, I., & Cho, J-C. (2015). Expansion of cultured bacterial diversity by large-scale dilution-to-extinction culturing from a single seawater sample. *Microb Ecol 71*, 29–43. DOI: 10.1007/s00248-015-0695-3

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., & Glöckner, F.O. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research 42*, D643–8. DOI: 10.1093/nar/gkt1209

Zehavi, T., Probst, M., & Mizrahi, I. (2018). Insights into culturomics of the rumen microbiome. *Front Microbiol 9*, D46–10. DOI: 10.3389/fmicb.2018.01999

Chapter 3: Impact of Habitat Structure on Microbial Composition and Secondary Metabolism

Foreword

The research in Chapter 3 was conducted by Emily N. Junkins with the help of undergraduate researcher Joseph McWhirter who aided in sample collection, media preparation, experimental breakdown, and organic extractions. Emily N. Junkins performed all molecular work and metabolomic sample preparation. All microbiome data analysis was performed by Emily N. Junkins. All metabolome data analysis was performed by Emily N. Junkins under the guidance of Dr. Laura-Isobel McCall. Dr. Vincent Bonifay and Dr. Steven Foster at the OU Mass Spectrometry, Proteomics and Metabolomics facility provided instrumentation for LCMS/MS analysis and guidance when writing methods. Dr. Graham Wiley and staff at University of Oklahoma Consolidated Core Laboratories provided sequencing support and instrumentation.

Abstract

Determining the drivers of microbial community assembly is a central theme of microbial ecology, and chemical ecologists seek to characterize how secondary metabolites mediate the same microbial community assembly. Habitat structure affects both how communities assemble and what metabolic pathways aid in assembly. Here, we sought to bridge these two perspectives by addressing the chemical drivers of community assembly within a spatially structured landscape. We hypothesized that structured habitats would select for higher microbial diversity and metabolite diversity. We anticipated that the production of a compound would be more advantageous in a structured habitat (less mixing) compared to an unstructured habitat (more mixing), where the molecule would have a diminished local effect. We observed this to only be partially true in our experiments: structured habitats had similar microbial diversity compared to unstructured habitats but differed significantly in the metabolites produced. We also found that structured habitats selected for communities that were more evenly distributed, rather than communities with higher richness. This provides further support to the idea that when characterizing the drivers of community assembly, it matters less about *who is there* and more about *what are they doing*. Overall, these data contribute to a growing effort to approach microbial community assembly with interdisciplinary tools and perspectives.

Introduction

A central focus of microbial ecology is characterizing the forces that govern microbial community assembly. Applying theories from macro-community ecology and accounting for attributes of microbial communities that make them unique is not an easy endeavor. The complexity of microbial community assembly stems from stochastic processes, like dispersal, and the effects abiotic and biotic, deterministic forces play on a population or community. Because microbial communities exist in spatial patches (Martiny et al. 2006; Caswell and Cohen 1991) and assemble into cell aggregates (Yanni et al. 2019; Tolker-Nielsen and Molin 2000), interactions will play a major role in community assembly as populations establish complex networks that define how microbial communities function. For example, microbial interactions can drive geochemical processes (Ho et al. 2016; Falkowski, Fenchel, and Delong 2008; Overmann and van Gemerden 2000) or nutrient uptake and host physiology in gut systems (García-Bayona and Comstock 2018; Coyte, Schluter, and Foster 2015). Many microbial interactions are mediated by the production of secondary metabolites that aid in both communication (X. Wang, Li, and Ling 2017) or competition (Hotterbeekx et al. 2017). Characterizing the chemical communication network in which microorganisms exist is crucial to understanding the rules that govern interactions at the population level (Dal Co et al. 2020; Justice et al. 2017). Where microbial ecology often halts is in bridging the gap between biodiversity/community assembly and chemical ecology/chemodiversity (Hilker 2014). Because ecology is mediated by chemical agents at various temporal and spatial scales, the chemistry of a system is an important factor when characterizing a community's function and stability (Raguso et al. 2015; Hilker 2014). Since many microorganisms naturally exist in a spatially structured, heterogeneous habitat (Tolker-Nielsen and Molin 2000), like biofilms, it is important to

characterize the effect of habitat structure on microbial community composition and function.

Here, we refer to habitat as the physical space a population inhabits, the gradients that exist or are constructed, and the interactions among members of the community. A habitat can impose barriers, chemical or physical, which can then influence the structure of a microbial community and its spatiotemporal dynamics (i.e., habitat heterogeneity, niche construction, stability). A structured habitat would lead to more environmental gradients and unstructured habitat conditions would harbor fewer environmental gradients. This type of structure can be achieved in a habitat though abiotic or biotic mechanisms. Structure can result abiotically from microscale gradients in nutrients or chemicals, like oxygen or light, as well as physical barriers, like surfaces. Structure can be established biotically through the metabolic activity of microorganisms themselves via nutrient degradation and the production of metabolites involved in signaling or competition (reviewed in (McNally and Brown 2015)). This process is referred to as niche construction (Odling-Smee and Laland 1996). Overall, gradients create spatial heterogeneity which affords more niches into which diverse groups can sort. It is as populations seek to assemble or compete in these gradients that microorganisms interact with each other (Dal Co et al. 2020) and where biodiversity and chemodiversity are intrinsically linked.

Structured habitats provide a mosaic in which neighbors can affect both each other and their surroundings (Picketts and Cadenasso 1999). The effect that structured habitats (i.e., adaptation and niche differentiation) have on clonal microbial populations has been characterized for some organisms, such as colicin-producing and colicin-sensitive *Escherichia coli* strains, where the structure of the habitat and the population size determined if a producer strain could outcompete

the sensitive strain (Chao and Levin 1981). A structured habitat also selects for phenotype diversification in *Pseudomonas fluorescens* compared to a constantly mixed, unstructured habitat (Rainey and Travisano 1998). In the case of *P. fluorescens,* structure provided vacant niches which allowed for different genotypes to partition along a gradient, particularly an oxygen gradient (Rainey and Travisano 1998). Another example using a mixture of algal species showed that spatial structure afforded by heterogenous stream flow rates selected for coexistence (diversity) where a uniform flow rate (unstructured) led to competitive exclusion (Cardinale 2011). These examples highlight the importance of dispersal where microorganisms stochastically arrive and then employ deterministic mechanisms to colonize new territory. Overall, structured habitats can afford multiple niche spaces, selecting for microbial diversity and functional diversity (Loreau et al. 2001) thereby creating ecological opportunity (Wellborn and Langerhans 2015). It follows that where there is opportunity that multiple groups can take advantage of, there will be competition for resources as organisms seek to establish themselves. Unlike the previous examples, we seek to characterize community assembly and interactions in a structured environment using a complex microbial community.

We hypothesized that structured habitats would select for the production of antimicrobial metabolites as a competitive advantage during community assembly. We anticipated that the production of a compound would be more advantageous in structured habitat (limited mixing), compared to an unstructured habitat (more mixing) where the molecule would have a diminished local effect (Czárán and Hoekstra 2003). Specifically, a structured habitat would keep the producer and compounds within close range so that a producer can incur the benefit from costly production (Dal Co et al. 2020; West et al. 2006; Chao and Levin 1981). Therefore, spatial

structure will determine the effect molecules can have on a population. We also hypothesized that the static communities would be more diverse (i.e. higher richness) since differential gradients could allow for slow-growers to partition, compared to the shaking communities that had a highly oxygenated environment that could select for fast-growers that would quickly dominate. To test these hypotheses, we took a diverse wastewater community, diluted it to obtain inocula of varying cell density, and incubated them in either static conditions to promote structure (e.g., oxygen gradients) or shaking conditions to disrupt any gradients (Rainey and Travisano 1998). We sampled these communities over the course of 12 days to generate a time series. Our mesocosm communities had high levels of diversity while also allowing us to control for abiotic variability. We measured the microbial community diversity via 16S rRNA gene amplicon sequencing and the metabolite diversity through untargeted metabolomics. We found that structured environments selected for different microbial communities and different chemical profiles.

Materials and Methods

Sampling and Experimental Set Up

Wastewater was collected from the aerator basin at the Norman Water Reclamation Facility, Norman, OK on 11 February 2020, stored on ice, and transported immediately back to the laboratory (< 30 minutes). The sample (50 mL volume) was homogenized vigorously on a vortex mixer for ~2 minutes to break up flocks and produce a slurry. The sample was serially diluted (1:10) in sterile R2B liquid medium four times to generate 5 different dilutions (dilution factors of 1 to 10,000) to be used for the starting inocula (Reasoner and Geldreich 1985). Each dilution was inoculated into sterile R2B medium for a final volume of 35 mL. A total of 150 enrichments were made so that three replicates could be sampled destructively, representing each dilution (1E-00, 1E-01, 1E-02, 1E-03, 1E-04) for each sampling day (Day 0, 3, 6, 9, 12) and for each condition (shaking or static). A total of 15 uninoculated medium controls (n=3) were made to be used at each time point; these were incubated with the static samples. All cultures were incubated at 30 °C, either shaking at 250 rpm or static (Supplemental Figure 1).

On each sampling day (0, 3, 6, 9, 12) prior to DNA extraction, triplicate cultures from each inoculum size and condition were vortexed and pooled into a sterile glass beaker. Any remaining biofilm was scraped from the glass walls of each test tube using a sterile, silicon cell scraper (Sarstedt, Nümbrecht, Germany), transferred to the sterile beaker, and mixed. From this pool, three 1.0 mL aliquots were transferred to sterile micro centrifuge tubes and spun at 10,000 x g for 1 min. The supernatant was removed, cells were resuspended 1 mL of sterile 1x PBS for one wash and spun at 10,000 x g for 1 min to generate a cell pellet. After the 1x PBS supernatant was removed, cells were resuspended in 750 uL Zymo bashing bead buffer (per Zymo Quick

DNA manufacturer protocol), transferred to Zymo bead tubes (Zymo Research Corp., Irvine, CA), homogenized for 45 seconds using a hand-held reciprocating saw with custom attachment (RYOBI, Anderson, SC), and stored at -20 °C until DNA extraction. Lastly, 103 mL of ethyl acetate was added to the remaining volume of the pooled samples (~103 mL) and mixed to generate a 1:1 ratio of culture and solvent for an overnight, organic extraction at room temperature. Medium controls for each time point, described above, were pooled, sampled for DNA extraction, and prepared for ethyl acetate organic extraction as described above for the cultures to be used as blanks for both sequencing and metabolomic data.

DNA Extraction and Sequencing

Before DNA extraction, each sample was thawed on ice and homogenized at maximum speed for 45 seconds using a hand-held reciprocating saw. DNA was extracted according to manufacturer specifications using the Zymo *Quick*-DNA Fungal/Bacterial Miniprep kit (Cat# D6005, Zymo Research Corp., Irvine, CA). Samples were stored at -20°C until extractions could be processed at once. For community characterization, a conserved region of the SSU rRNA gene of most bacteria, archaea, and eukarya was amplified using primers 515F-Y and 926R (Parada, Needham, and Fuhrman 2015) via the following PCR protocol: initial denaturation at 94˚C for 2 min, followed by 30 cycles of denaturation at 94˚C for 45 s, annealing at 50˚C for 45 s, and extension at 68˚C for 90 s, with a final extension at 68˚C for 5 min. These primers produced two amplicons, a ~400 bp fragment for bacteria and archaea, and a 600 bp fragment for eukaryotes. The forward primer 515F-Y (5'-GTA AAA CGA CGG CCA G CCG TGY CAG CMG CCG CGG TAA-3') contains the M13 tag (underlined) fused to the 5' end of the forward primer (Kraus et al. 2018). The reverse primer 926R (5'-CCG YCA ATT YMT TTR AGT TT-

3') was unmodified from Parada et. al 2015. Each PCR contained 5 PRIME HOT master mix (1X; 5 PRIME Inc., Gaithersburg, MD), 0.2 μM of each primer, and 3.0 μL of extracted DNA at a final volume of 50 μL. The amplified fragments were purified using Sera-Mag magnetic beads (GE) with the AmPureXP (Beckman Coulter) protocol at a final concentration of 1.8x v/v. After purification, 3 μL of each PCR product, 1x 5 PRIME HOT master mix (Quantabio, Massachusetts, USA), 0.2μM of the 926R primer, and 0.2 μM of a specific 12 bp oligonucleotide was used in a separate barcoding PCR (6 cycles) in 50 μL reactions to attach a unique barcode to amplicons of each library. The same thermocycler protocol was used as above but only run for 6 cycles. The now barcoded amplicons were purified using Sera-Mag (GE) beads with the AmPureXP (Beckman Coulter) protocol to a final volume of 40 μL, quantified using the QuBit HS DS DNA assay kit (Thermo Fisher Scientific Inc., Waltham, MA), and pooled in equimolar amounts. The pooled, barcoded amplicon libraries were then concentrated to a final volume of 100 μL (16.6 ng/ μL) with an Amicon-Ultra filter (Millipore, Burlington, MA, USA) following manufacturer's protocol. The combined amplicon libraries were denatured according to Miseq library preparations protocol (Illumina, San Diego, CA, USA). The sample was loaded at a concentration of 10 pM and sequenced using 2x250 paired-end strategy on the Miseq (Illumina San Diego, CA, USA) platform for 251 cycles.

Organic Extractions for LCMS/MS

After overnight ethyl acetate extractions, the organic layer was transferred to round-bottom flasks and evaporated on a Heidolph Hei-Vap Value (Heidolph, Schwabach, Germany) rotovap. Samples were resuspended in 100% methanol and transferred to pre-weighed scintillation vials. Samples were dried under nitrogen gas at room temperature and weighed for

final yield (mg). For this experiment, all beakers were brand-new from the same lot, rinsed with 100 % methanol prior to use, and one beaker was incubated overnight with 103 mL of ethyl acetate to generate a glassware blank to attribute any features in the metabolomics data to glassware alone. Once samples were dry, they were stored at -80°C until experiment was complete.

After all metabolite samples were dried and stored at -80°C, each sample was resuspended to a final concentration of 11 mg/mL in 100% isopropanol containing 0.1 μM of sulfadimethoxine as an internal standard. A volume of 100 μL of each sample was transferred to a LCMS/MS 96-well plate and stored at -80 °C until it was run on the instrument. In addition to the samples contain in internal control, a pooled sample control containing 2 μL of every sample and a 100% isopropanol (IPA) with internal standard was also prepared. Lastly, 100% methanol blanks were run between each experimental sample.

HPLC-ESI Mass Spectrometry

HPLC separation of the samples (5μL injection volume at 11 mg/mL) was performed on an Agilent 1290 HPLC system with a Waters ACQUITY UPLC BEH column (ODS-18; 2.1 x 100 mm; 1.7 μm particle size). Mobile phase A was comprised of 0.1% formic acid in Optima LC/MS water (Fisher), while mobile phase B was comprised of 0.1% formic acid in 100% HPLC-grade acetonitrile. A binary gradient at 0.4 mL/min flow rate was used under the following steps: 90% solvent A and 10% solvent B from 0 to 1 min, linear gradient to 20% solvent B from 1-2 min, linear gradient to 80% solvent B from 2-16 min, linear gradient to 100% solvent B from 16-18 min, and 100% solvent B from 20 – 21 min, which was maintained for an

additional 2 min before the next sample injection. The resulting eluent from the HPLC column was run on an Agilent 6545 accurate mass Q-TOF mass spectrometer with an electrospray ionization source operating in the positive mode. Nitrogen was used as a nebulizing gas (40 lbs/in^2) and as a drying gas (325 °C; 10 L/min flow rate) (Bartley et al., 2013). Fragmentor voltage was 180 V, skimmer voltage was 45 V, and capillary voltage was 4,000 V. Untargeted MS/MS was used with mass-dependent collision energy ramp from 20-35 V using UHP Nitrogen. Data was collected with Mass Hunter Acquisition software (B.08.00) and analyzed with Mass Hunter Quantitative Analysis software (B.06.00).

ASV Detection

Barcodes from raw SSU rRNA amplicon sequences were trimmed and sequences were dereplicated using cutadapt v 3.0 (M. Martin 2011). Demultiplexed reads were quality filtered, forward and reverse reads were merged (~372 bp) and amplicon sequence variants (ASVs) were inferred as a part of the *dada2* pipeline (Callahan, McMurdie, and Holmes 2017). Taxonomy was assigned using the SILVA database v32 (Yilmaz et al. 2014; Quast et al. 2013;). Contaminants were removed via the r package *decontam* v. 1.10 using the frequency and prevalence based filtering with a threshold of 0.5 (Davis et al. 2017). The final, experimental dataset consisted of ~1.4 million reads resulting in 1923 ASVs with a median of 9054 reads per samples.

Tree Building and Community Analysis

Sequences were aligned using the r package *DECIPHER* v 2.0 (Wright 2016; Wright 2015) and the r package *phangorn* v2.5.5 (Schliep 2011) was used to first construct a neighbor-joining tree to use as the starting point for a GTR+G+I maximum likelihood tree. Community

analysis was done using environments created from the r package *phyloseq* v1.26.1 (McMurdie and Holmes 2013). Richness was calculated using the r package *breakaway* which calculates error in richness estimation (Willis and Bunge 2015). Evenness was calculated using Pielou's index (Pielou 1966). To test whether habitat structure had an effect on the richness and evenness of the microbial communities for each dilution, a Levene's test was used to confirm homogeneity of variance (Levene 1960) using the r package *car* (Fox and Weisberg 2019) before a two-way ANOVA coupled with a Tukey's honest significant difference test was used to compare differences between shaking and static cultures for each time point in PRISM v9 (GraphPad Software, San Diego, California USA).

To characterize beta diversity, distances were calculated using weighted UniFrac distances (Lozupone and Knight 2005) and ordinated using Principal Coordinate Analysis (PCoA) calculated with 95% confidence ellipses. Statistical comparisons across cultivation conditions, dilution, and day were done using PERMANOVA with the adonis function in the *vegan* r package (~ Dilution*Condition*Day) (Oksanen et al. 2019). To fully understand the significance described in the PERMANOVA, a PERMDISP with the betadisper function in *vegan* was used to describe any within sample variance (i.e. across replicates) that could explain any significant differences detected in the PERMANOVA . Hypothesis testing via ANOVA with permutations (n=999) was used with PERMDISP to determine any significant differences in variation within samples.

To further investigate changes in abundance of certain taxa between shaking and static conditions, the r package *corncob* was used to calculate differential abundances using a beta

binomial model at the genus level (B. D. Martin, Witten, and Willis 2020). This model controlled for covariates (day or dilution).

Metabolomics Data Processing

Raw data obtained from the LCMS/MS instrument was converted to .mZXML files using msConvert within ProteoWizard v3 (Chambers et al. 2012). The resulting .mzXML were uploaded to MassIVE (massive.ucsd.edu) as public dataset (MSV000086507).

Molecular Networking and Spectral Library Search in GNPS

The mass spectrometry data were first processed with MZmine2 v2.51 (Pluskal et al. 2010). Peaks were identified (MS1 noise cutoff: 1.0E2, MS2 noise cutoff: 1.0E1). Chromatograms were built under the following criteria using the ADAP chromatogram builder (Myers et al. 2017): MS1 scans; Minimum Group Number: 5; Group Intensity: 2.0E2; Minimum High Intensity: 1.0E3; *m/z* tolerance: 0.02 *m/z* or 0.0 ppm. Deconvolution was achieved through the following parameters via the local minimum algorithm: Chromatographic Threshold: 5.0%; search minimum in RT range: 0.4 min; Minimum Relative Height: 5.0%; Minimum Absolute Height: 2.0E3; Minimum Ratio Peak Top/Edge: 1; Duration: 0.01-3.50 min; Algorithm: Median; *m/z* MS2: 0.01Da; RT Range: 0.2 min. Isotopes were grouped under the following criteria: *m/z* total: 0.02; RT total: 0.10 min; Maximum charge: 3; Representative Isotope: Max Intensity. The feature table was constructed with using the following alignment parameters: *m/z* tolerance: 0.02 *m/z* or 200 ppm; Weight for *m/z*: 75; RT tolerance: 0.1 min; Weight for RT: 25 and filtered using the following criteria: Min peaks in per row : 3; Keep only peaks with MS2: yes. Poor peaks were removed from the peak list (long tails or plateaus) and any feature found in a blank

(methanol, glassware, or isopropanol) was removed from the feature table. Results were exported to GNPS (M. Wang et al. 2017), for feature based molecular networking (FBMN) analysis (Nothias et al. 2020).

In GNPS the data were analyzed using FBMN workflow (Nothias et al. 2020). The data were filtered by removing all MS/MS fragment ions within +/- 17 Da of the precursor *m/z*. MS/MS spectra were filtered by choosing the top 6 fragment ions in the +/- 50 Da window. The precursor ion mass tolerance was set to 0.02 Da and the MS/MS fragment ion tolerance to 0.02 Da. A molecular network was created with edges (filtered via cosine score > 0.7 and > 4 matched peaks). Also, edges between nodes were kept only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100. The analogue search mode searched against MS/MS spectra with a maximum difference of 100 Da in the precursor ion value. The library spectra were filtered as the input data was filtered, as discussed above, cosine score above 0.7 and at least 4 matched peaks.

Data Deposition and Job Accessibility

Sequence data has been deposited at NCBI's Sequence Read Archive (SRA) database under accession number PRJNA732431.

The mass spectrometry data were deposited on MassIVE (MSV000086507). The molecular networking job can be publicly accessed at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=e96bf8001f6d4518834ac6ba3742bd81. (Supplemental run containing samples from Day 0 can be accessed at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=807cc96f3c5d4898aca64659484d86ea).

Metabolomics Data Analysis and Hypothesis Testing

Principal Coordinate Analysis using Bray-Curtis distances and 95 % confidence ellipses calculated from total ion current (TIC)-normalized feature table was used to visualize the composition of metabolites using the r package *phyloseq* v1.26.1 (McMurdie and Holmes 2013). The molecular networks were visualized using Cytoscape software v3.8.1 (Shannon et al. 2003). To test for differences, a PERMANOVA and PERMDISP were used (as described above) to characterize differences between metabolite composition among samples. To test the hypothesis that structured habitats yield a community that produces a richer pool of metabolites, we first normalized the feature table to the total signal total ion current (TIC) then performed a Mann-Whitney U test (Wilcoxon rank sum test) using the r package *stats* (R Core Team 2018). We also used a Fisher's Exact Test to confirm differences in unique/total metabolites for each condition. To characterize how metabolite diversity changes over time, a liner mixed effects model was fit using REML via the r package *lme4* lmer function (Bates et al. 2015). The formula *Features ~ Condition * Day + (1 | Dilution)* was used to compare the change in metabolite diversity (richness) over time between shaking and static conditions while treating inoculum size as a random variable (e.g. dispersion).

To investigate the relationship between microbial and metabolite alpha diversity measures, we performed a Spearman's rank-based correlation comparing the metabolite richness to each alpha diversity measure using the r packages *Hmisc* and *corrplot* (Frank E Harrell Jr 2020; Wei and Simko 2017).

Results

Biofilm Formation in Shaking and Static Cultures

Despite the intention to remove structure through shaking and homogenization, many shaking cultures contained visible biofilm on the walls of the culture tube at the air-media interface (Supplemental Figure 2). The speed of shaking was maximized, limited only by the need to prevent spillage. We speculate that communities exploited this available structure, as microorganisms live in structured environments naturally (Tolker-Nielsen and Molin 2000; Picketts and Cadenasso 1999), leading to the observed biofilm. However, there was nevertheless less structure than in the static cultures.

Differences in Microbial Community Structure Were Primarily Driven by Habitat Structure

To focus on those communities being selected for through this experiment, Day 0 samples were removed from the analysis and only samples that result from cultivation were further analyzed (See Supplemental Text). After removing Day 0, the most obvious pattern was the separation of static communities and shaking communities (PERMANOVA $F= 75.12$, $r^2=0.24$, $p=0.001$; PERMDISP $F=1.46$, $p=0.218$) (Figure 1). Microbial communities also clustered according to inoculum size where samples partitioned from low to high dilutions for both static and shaking conditions (PERMANOVA $F= 13.89$, $r^2=0.17$, $p=0.001$; PERMDISP $F=2.71$, $p=0.0.038$) (Figure 1B). Lastly, microbial community structure changed little over time (PERMANOVA $F= 6.42$, $r^2 = 0.06$, $p=0.001$; PERMDISP $F= 4.80$, $p=0.007$) (Figure 1A). Interaction effects between variables suggest that main effects alone do not drive differences, but rather combinations of factors like starting inoculum size, time, and condition jointly have an

effect (Dilution:Condition F = 4.242, r^2= 0.05, p= 0.001; Dilution:Day F= 2.67, r^2= 0.10, p= 0.001; Condition:Day F= 3.7, r^2= 0.04, p=0.001; Dilution:Condition:Day F= 2.84, r^2= 0.11, p=0.001). Interactions between treatments and the significant dispersion within treatments, like dilution or time, highlight important factors that will affect microbial community assembly. For most conditions, the beta diversity between replicate samples became more variable as inoculum size decreased (i.e. high ranges for distance to centroid) (Supplemental Figure 4B). This would suggest that community assembly follows a predictable trend when the inoculum is larger, but as populations are diluted, the outcome of community assembly becomes more difficult to predict. Conversely, little dispersion between static and shaking samples suggest that the effect of habitat structure is acting in a predictable manner (Supplemental Figure 4A). Overall, beta diversity indicates that structure, afforded by static cultivation or removed by shaking cultivation, was the primary driver of diversity and that the size of the starting community and time can have an effect on community assembly in structured and non-structured habitats.

Structured Habitats Selected for Different Community Compositions

Richness was similar between shaking and static communities, and richness changed little over time (Figure 2A) (Structure; Dilution 1: F = 3.7, p=0.07; Dilution 0.1: F = 2.6, p=0.13; , Dilution 0.01: F = 2.6, p=0.13; Dilution 0.001: F = 1.6, p=0.23; Dilution 0.0001: F = 0.28, p=0.60) (Time; Dilution 1: F = 1.5, p=0.27; Dilution 0.1: F = 1.8, p=0.20; Dilution 0.01: F = 1.5, p=0.25; Dilution 0.001: F = 2.5, p=0.10; Dilution 0.0001: F = 1.0, p=0.42). In contrast, evenness significantly differed between some shaking and static communities (Figure 2B) (Structure; Dilution 1: F = 0.32, p=0.58; Dilution 0.1: F = 18, p=0.0008; Dilution 0.01: F = 56, p<0.0001; Dilution 0.001: F = 123, p<0.0001; Dilution 0.0001: F = 17, p=0.0011). Evenness also

significantly changed over time (Time; Dilution 1: F = 7.1, p=0.004; Dilution 0.1: F = 2.8, p=0.07; Dilution 0.01: F = 15, p<0.0001; Dilution 0.001: F = 25, p<0.0001; Dilution 0.0001: F = 3.4, p=0.05) (Figure 2B). Further, structured habitats selected for higher abundances of certain taxa (Figure 3), despite some taxa being shared between shaking and static conditions (Supplemental Figure 5). Specifically, static cultures had significantly higher relative abundances of *Acrobacter* spp., *Prevotella* spp., *Pseudomonas* spp., and *Delftia sp*p. compared to shaking. We also observed lower abundant taxa, like *Desulfovibrio* spp., to have significantly higher relative abundances in static cultures compared to shaking cultures (Figure 3). In contrast, shaking cultures were dominated by *Aeromonas* spp. with similar relative abundance across inoculum size (Supplemental Figure 5). Considering structured environments, static cultures could have provided different niches, particularly along an oxygen gradient, that may allow different taxa to grow equally well, resulting in more even communities. This is contrasted with unstructured environments, that possibly selected for more oxygen-adapted taxa that were able to quickly outcompete other taxa, resulting in communities with a dominate taxon, like *Aeromonas* sp. Overall, communities resulting from shaking or static samples were composed of similar populations but were structurally different.

Cultivation Conditions Produced Communities with Different Metabolic Profiles

We observed the structure of the metabolome to be significantly different between shaking and static cultures (PERMANOVA F= 64.83, r^2= 0.07, p=0.001, PERMDISP F= 18.44, p = 0.001) (Figure 4A, 4C) indicating that these conditions resulted in different metabolic profiles. The metabolome structure also differed among inoculum sizes (PERMANOVA F= 26.74, r^2= 0.12, p=0.001, PERMDISP F= 1.04, p = 0.395) and over time (PERMANOVA F=

30.85, r^2= 0.11, p=0.001, PERMDISP F= 2.85, p = 0.041) (Figure 4A, 4C). More so than was observed in the microbial community structure, interactions between cultivation conditions, inoculum size, and time, explained much of the variation between metabolite compositions (PERMANOVA Dilution:Condition F= 24.26, r^2= 0.11, p= 0.001; Dilution:Day F= 17.56, r^2 = 0.23 and p = 0.001; Condition:Day F= 23.35, r^2 = 0.08, p = 0.001, Dilution:Condition:Day F= 0.21, r^2 = 0.21 and p=0.001). A total of 391 metabolites were shared between shaking and static cultures, while static cultures had 1204 unique metabolites and shaking cultures had 263 unique metabolites (Fisher's exact test, p < 0.0001) (Figure 4B). To test the hypothesis that structured environments foster a community producing a higher richness of metabolites, we performed a Mann-Whitney U Test on feature counts (TIC normalized) with Benjamini-Hochberg correction and found static cultures to have significantly higher richness than shaking cultures (W = 952.5, p-value = 3.657e-05), though differences in richness between shaking and static was dependent on starting inoculum sizes (Figure 3C) (Dilution 1: F = 84, p<0.0001; Dilution 0.1: F = 2.0, p=0.18; Dilution 0.01: F = 72, p<0.0001; Dilution 0.001: F = 1.9, p=0.19; Dilution 0.0001: F = 48, p<0.0001). We also found an overall greater diversity of metabolite features over time in static cultures (anova(lmer):Chi-sq= 22.4, p<0.0001)(Figure 3D), but not shaking cultures, though changes over time varied between dilutions (Dilution 1: F = 51, p<0.0001; Dilution 0.1: F = 15, p<0.0001; Dilution 0.01: F = 26, p<0.0001; Dilution 0.001: F = 9.0, p=0.0012; Dilution 0.0001: F = 17, p=0.0011). Because we normalized by weight, the increase in feature counts is not just due to increases in biomass as cultures grew.

This increase in metabolite diversity could potentially highlight different metabolic mechanisms aiding populations during community assembly. For example, of the metabolites

that could be annotated, the majority belonged to families of compounds involved in quorum-sensing and signaling, particularly observed in *Pseudomonas* spp. physiology (Figure 4D; Supplemental Figure 6). The production of these molecules and the ability to rapidly adapt make pseudomonads successful colonizers and aggressors (Thierbach et al. 2019; Morales-Soto et al. 2018; Heeb et al. 2011; Rainey and Travisano 1998). Lastly, compounds known to have antagonistic activity were more common in static cultures compared to shaking cultures, like anisomycin (Grollman 1967; Sobin and Tanner Jr. 1954) or orfamides (Ma et al. 2016; Jang et al. 2013; Gross et al. 2007) (Supplemental Figure 6-7). This observation suggests that structured habitats afford more opportunity to interact and may select for organisms that use metabolite-mediated strategies to be successful during community assembly, compared to shaking cultures where the advantage was most probably due to increased growth rate (e.g. *Aeromonas* sp.).

We initially hypothesized that structured habitats would have higher microbial richness as a function of environmental gradients creating diversified niche-availability and would also have higher richness of metabolites as a function of higher richness of community members. To investigate this relationship, we performed a Spearman's correlation on the richness and evenness of the microbial community and the metabolomes. We observed that metabolite richness had no significant correlations with microbial alpha diversity measures (breakaway/microbial richness: r^2= -0.02, p= 0.88; microbial evenness: r^2= 0.23, p= 0.16, metabolite evenness : r^2= -0.04, p = 0.83) (Supplemental Figure 8). This suggests that it may not the number of organisms in a community that matter for metabolite diversity, but rather who they are and the strategies they employ during community assembly, like pseudomonads and the fleet of metabolites they produce, compared to *Aeromonas* sp., that quickly outgrew other

populations.

Similar to the microbiome data, some treatments had significant variability in terms of metabolome (PERMDISP Condition F= 18.44, p = 0.001; Dilution F= 1.04, p= 0.395; Day F= 2.85, p=0.041). However, the effect of those treatments on the structure of the metabolome varied differently than the structure of the microbiome. First, habitat structure (shaking or static) produced variable metabolite compositions among samples compared to its effect on the microbiome, p =0.001 and p=0.281, respectively. Further, the metabolomes from shaking cultures were more variable (Figure 4A, 4C) compared to the metabolome structure from static cultures suggesting that the structure afforded by static conditions may affect the metabolome in a more predictable way, in the context of our starting microbial communities. Secondly, the metabolomes resulting from each inoculum size did not vary in structure compared the microbial communities, p= 0.395 and p=0.038, respectively. This could indicate that changes in the inoculum size may have a stochastic effect on the microbial community (e.g. dispersal) but a deterministic effect on the metabolome (e.g. frequency-dependent, metabolite-mediated interactions), such that community assembly is less about predicting the answer to the question 'who will be there?', but rather predicting the answer to the question 'what will they be doing'? Overall, habitat structure may select for predictable compositions in the microbial community and the metabolome in structured communities, but the metabolites associated with community assembly in unstructured habitats can vary.

Discussion

By manipulating the physical environment of a liquid ecosystem through shaking or static cultivation, we characterized the microbial community response through changes in microbial diversity and chemodiverisity. Our data show that (1) habitat structure drives differences in the metabolome and microbiome, (2) habitat structure selected for more evenness in microbial communities compared to unstructured environments, (3) structured habitats selected for higher metabolite richness, and (4) the microbial richness metrics failed to capture the metabolic capabilities of microbial communities. We show that membership was not the only determinant of community assembly. Instead, habitat structure plays a significant role in secondary metabolic processes that drive community assembly. These findings add to efforts seeking to determine what drives microbial community assembly and how the metabolome can be linked to those drivers.

Community structure was primarily driven by cultivation condition, suggesting that selection was driven by the structure of the habitat. Because there is no pattern of change in structure from day 3 through day 12 (Figure 1), we suggest that cultivation condition has the greatest selective effect early and benefited those that could grow quickly. Constant mixing of the shaking cultures reduced any oxygen gradients and potentially selected for fast-growing taxa that would allocate resources to increase growth rate over costly secondary metabolite production, which would have a diminished local effect in well-mixed habitats (Dal Co et al. 2020). Growth rate, in this case an organism's ability to quickly establish abundance, was perhaps the driver of community difference and an important strategy during community assembly. The suggestion is supported by the fact that *Aeromonas* sp. quickly establish

dominance (Day 3) in shaking cultures compared to static cultures that were more even. However, this ubiquitous taxon, commonly found in wastewater (Skwor et al. 2020; Figueira et al. 2011) and gastrointestinal tracts (Janda and Abbott 2010), is a known biofilm producer (Liébana et al. 2016). Despite the intention to remove structure, the shaking cultures yielded visible biofilms on the walls of the tubes where medium was washed through shaking (Supplemental Figure 2). Although we did not identify isolates from the biofilms, one could expect *Aeromonas* spp. to be involved in what is probably a multi species-biofilm in these cultures. Since most microbial populations exist in spatial structure (Tolker-Nielsen and Molin 2000), particularly in biofilms, it follows that populations would seek to establish structure in a new environment to reap benefits of secreted molecules as cooperative or public goods (Yanni et al. 2019).

In contrast, static cultures exhibited significant biofilm formation but did not contain abundant *Aeromonas* spp. populations (Figure 3). Of the taxa that were significantly associated with static cultures, *Arcobacter* spp., *Delftia* spp., and *Pseudomonas* spp. were the most abundant in those cultures. Each of these genera contain members involved in biofilm formation (Rema et al. 2014; Assanta et al. 2002; Costerton, Stewart, and Greenberg 1999). Further, *Pseudomonas* spp. are particularly associated with competitive interactions in structured habitats, like biofilms (Inglis et al. 2011; Harrison et al. 2008). In well studied systems, like the cystic fibrosis lung or wound models, interactions between *Pseudomonas aeruginosa* and *Staphylococcus* spp. are largely antagonistic, but when afforded structure by a biofilm extracellular matrix, populations can co-exist. This type of antagonism will eventually generate structure in which interactions take place among relatives, rather than competitors, leading to

coexistence and cooperation (Yanni et al. 2019; Ghoul and Mitri 2016). However, these types of interactions will be dependent on what Dal Co et al refer to as the 'interaction range', or the spatial range and arrangement between interacting cells (Dal Co et al. 2020). The ability for diffusible molecules to influence a population will depend on density of cells and the uptake rate of that molecule (Dal Co et al. 2020). They will work similarly to local killing. The act of a population outcompeting and establishing itself will create spatial structure (Yanni et al. 2019), since a molecule acting locally will help generate structure in chemical space (Yanni et al. 2019; Czárán and Hoekstra 2003) by establishing a gradient. In our experiment, we observed *Pseudomonas sp.* to be significantly associated with structured conditions and elicitation of multiple families of metabolites associated with pseudomonad quorum sensing and signaling.

Overall, cultivation conditions selected for differential abundances of shared taxa resulting in two biofilm forming communities that either had a dominant taxon (shaking) or a more even community composition (static), highlighting the growth mechanism of microorganisms seeking to establish or thriving in structure (Tolker-Nielsen and Molin 2000; Picketts and Cadenasso 1999) through biomass and/or metabolite production (van der Meij et al. 2017; Traxler and Kolter 2015; Kinkel et al. 2014; Pacala and Levin 1997). In short, spatial structure will play a key role in community assembly as it both defines the initial constraints by limiting nutrients or physical space, and adapts as the mechanisms by which microorganisms proliferate can alter the spatial structure of the habitat.

Differences in community structure were also affected by the size of the starting inoculum (i.e. dilution). Diluting the inoculum reduced the number of cultivable organisms and

introduced variability at high dilutions. The effect of decreased inoculum sizes has been demonstrated in dilution-to-extinction cultivation efforts (Justice et al. 2017). We observed this in our data where, despite having differences in community structure based on dilution, we observed significant dispersion within each dilution (PERMDISP). When samples vary differentially, this could indicate that the cultivation conditions may not be selecting for a single type of community or taxon through deterministic processes (e.g. interactions or environmental filtering) where they are acting equally on each dilution, but rather creating stochastic assembly opportunities, where the establishment of an organism is based on random, non-niche based chance (Zhou and Ning 2017) introduced through altering the inoculum size. Dilution can vary the proportions of organisms most suited to be successful which in turn will alter the effect of deterministic processes during community assembly, like interactions and niche-specific traits. In short, the success of an organism will be a function of its abundance in the inoculum where both stochastic and deterministic processes can affect that success.

Though habitat structure did not influence the richness of a microbial community, we did observe these cultivation conditions to select for a more even community. Evenness has been implicated in the resilience, stability, and productivity of microbial communities (Haig et al. 2015; Wittebolle et al. 2009). We observed that more even communities had higher richness of metabolites. Further, competitive interactions have been shown to increase stability by limiting the inefficient feedback loops created by cooperation (Coyte, Schluter, and Foster 2015; Oliveira, Niehus, and Foster 2014). It follows that communities in structured environments with high cell density have more opportunities for competitive interactions mediated by secondary metabolites, especially when resources become limited.

Between structured and unstructured habitats, we observed that these communities were composed of similar populations (richness) but were structurally different (Figure 2B, Figure 1). Further, we observed structured habitats to select for higher richness of metabolites, but not higher richness of organisms, highlighting that microbiome-level metrics of richness failed to capture the metabolic capabilities of a community, particularly those related to interactions and the production of metabolites. Our assumption that a rich community will also be a productive community (e.g., interacting via small molecules) assumes that members of a community (1) possess the ability and strategy of compound production equally and (2) the method of extraction captures all metabolites equally, but this was not observed. As such, one explanation for the differences in the metabolome between structured and non-structured habitats was driven by the bacterial strains for which the conditions are selecting and the method through which the metabolites were extracted. For example, structured habitats could be selecting for organisms that use structure and the production of small molecules as an advantage (Raguso et al. 2015; Davies and Ryan 2012), like *Pseudomonas* spp. during biofilm production (Bassler and Losick 2006; Venturi 2006). But it is important to acknowledge that non-polar extraction methods, like ethyl acetate used here, will favor the recovery of less polar molecules, like *Pseudomonas* spp. metabolites. In this case, both habitat structure and the type of molecules extracted would describe differences in the metabolome, rather than an effect driven by the microbial richness of that community.

Metabolome structure was primarily driven by inoculum size, time, and cultivation condition (PERMANOVA). Inoculum size had the greatest effect on metabolite composition

either alone ($r^2 = 0.11$) or interacting with cultivation conditions ($r^2 = 0.11$) or time ($r^2 = 0.23$) or both ($r^2 = 0.21$). Here, the size of the establishing populations mattered for subsequent metabolite production and thus the structure of the metabolome is an important factor to consider when trying to explain drivers and strategies in microbial community assembly. For example, differences in the composition of the metabolome could be related to strategies for increased growth rates during community assembly where many of the detected metabolites are by-products of rapid resource consumption. Alternatively, in the case of our structured habitat data, the composition of the metabolome had many molecules used during interactions (Figure 4D, Supplementary Figure 6) and/or had antimicrobial properties (Supplementary Figure 6). This suggests that biotic interaction played a role during community assembly in structured habitats. Consider a microbial community as a neighborhood where smaller, shorter range interactions can govern microbial community dynamics (Dal Co et al. 2020; Ladau and Eloe-Fadrosh 2019; Oliveira, Niehus, and Foster 2014). The type of interaction can vary (Granato, Meiller-Legrand, and Foster 2019; Estrela and Brown 2018; Hibbing et al. 2010), but for those mediated by small diffusible molecules, the effects of those molecules will have greater effect in shorter ranges (Dal Co et al. 2020; Garcia-Garcera and Rocha 2020; Chao and Levin 1981), something a structured habitat would provide.

Conclusion

Merging microbial ecology and chemical ecology perspectives can put into context the abiotic and biotic drivers of community assembly by combining chemical and molecular approaches. Here, we showed that community assembly patterns are dependent on habitat structure, and that a structured environment selected for communities with higher chemical diversity. Not only do these findings seek to connect the 'who is there' and 'what are they doing' paradigm of ecology, but also highlights the role cultivation can play in targeting new metabolisms for natural product discovery.

Figures

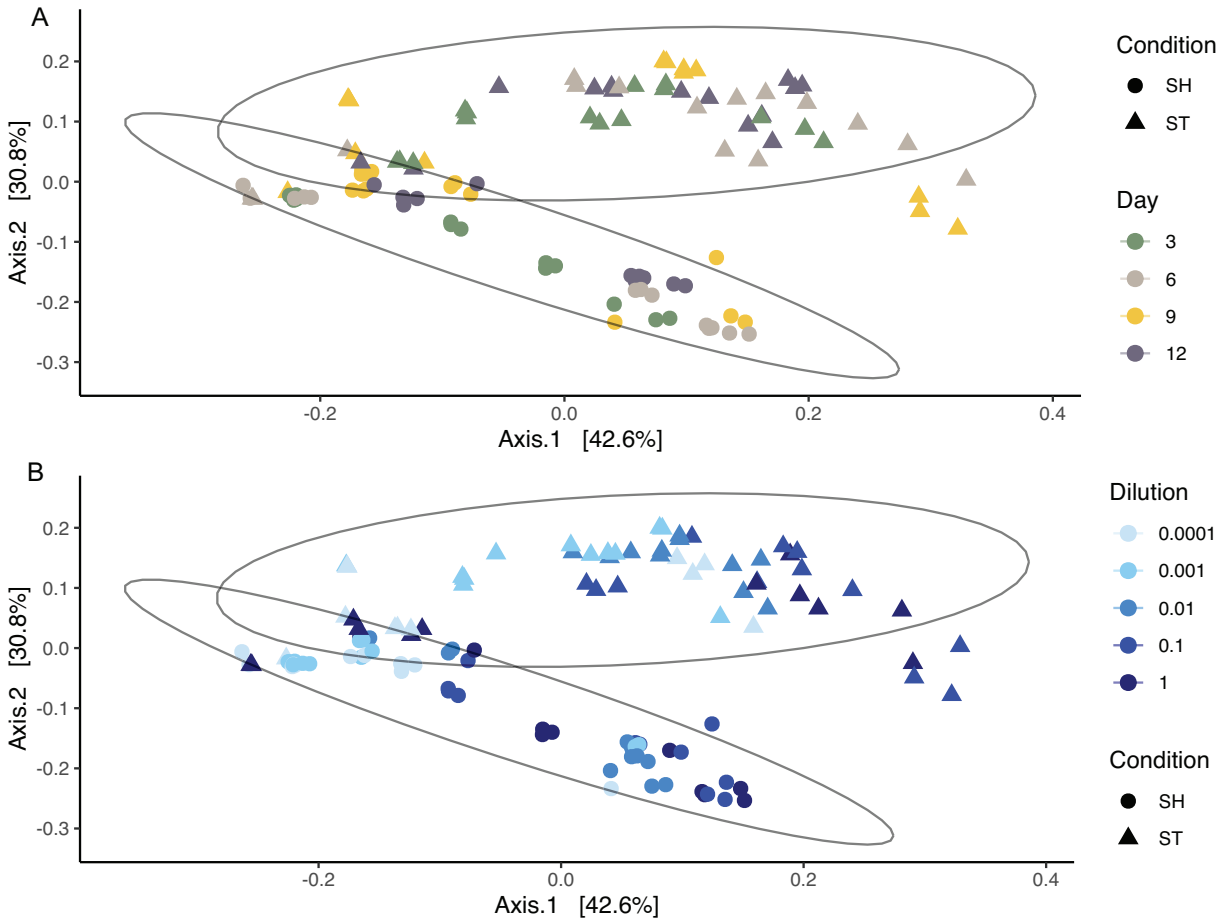


Figure 1: Principal Coordinate Analysis of microbial communities using weighted UniFrac distances. SH refers to shaking cultures, ST refers to static cultures. Shape and 95% confidence ellipses refer to cultivation condition, color refers to (A) time and (B) inoculum size.

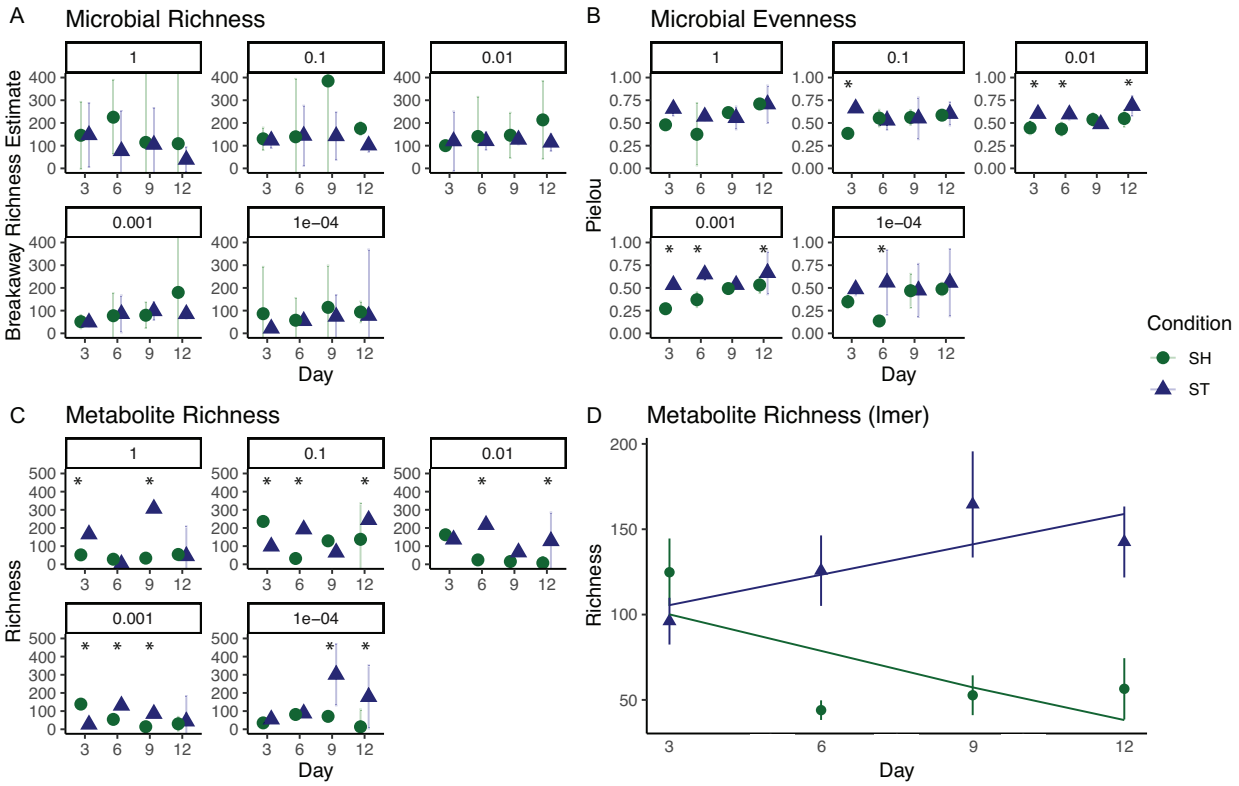


Figure 2: Alpha diversity metrics for (A) microbial richness, (B) microbial evenness, (C) metabolite richness, and (D) a mixed-effects model of metabolite richness over time. Shape and color refer to cultivation condition. Data (A-C) were handled as independent samples based on inoculum size. Significant differences from Tukey's HSD test ($p>0.05$) between shaking and static cultures designated with (*).

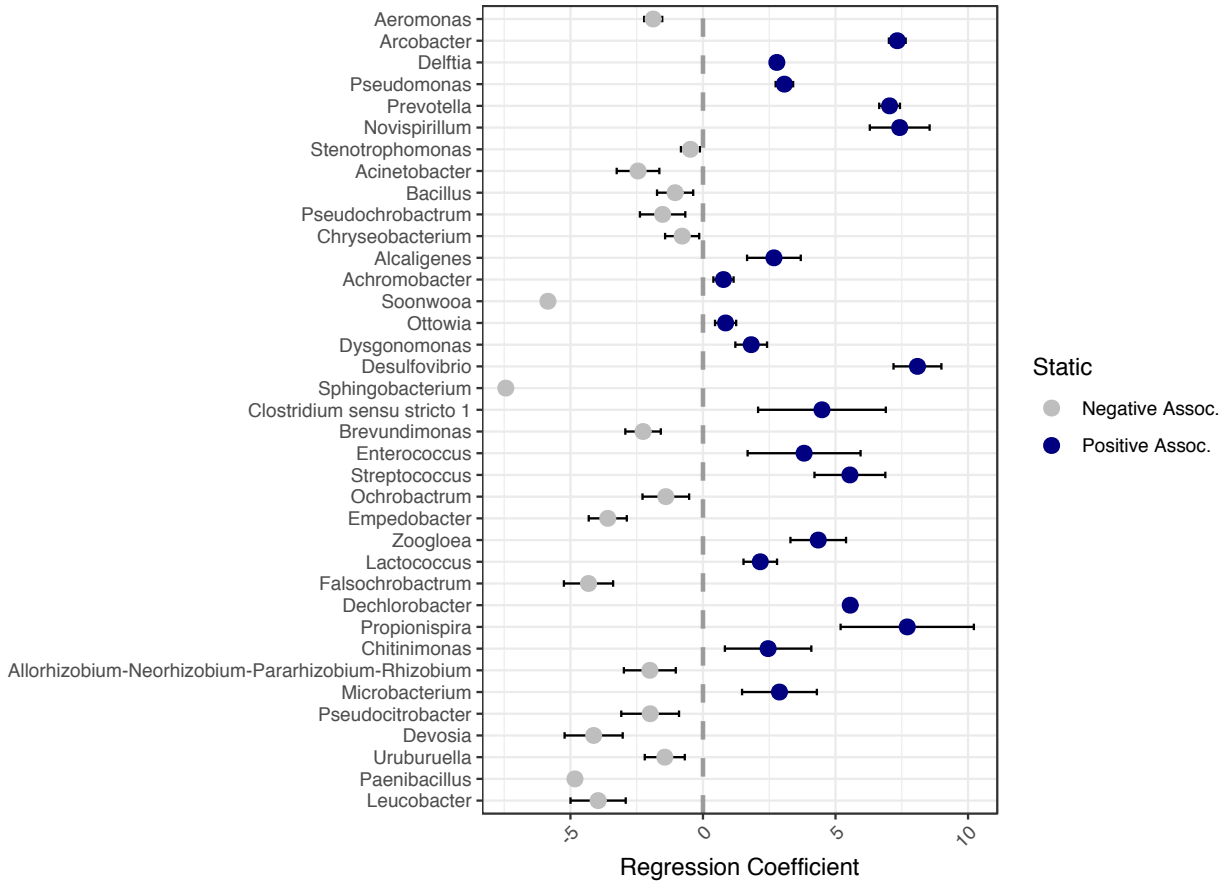


Figure 3: Differential abundance calculated using beta binomial regression at genus level. Blue refers to genera positively associated with static conditions; grey refers to genera negatively associated with static conditions. Order of genera based on relative abundance across experiment.

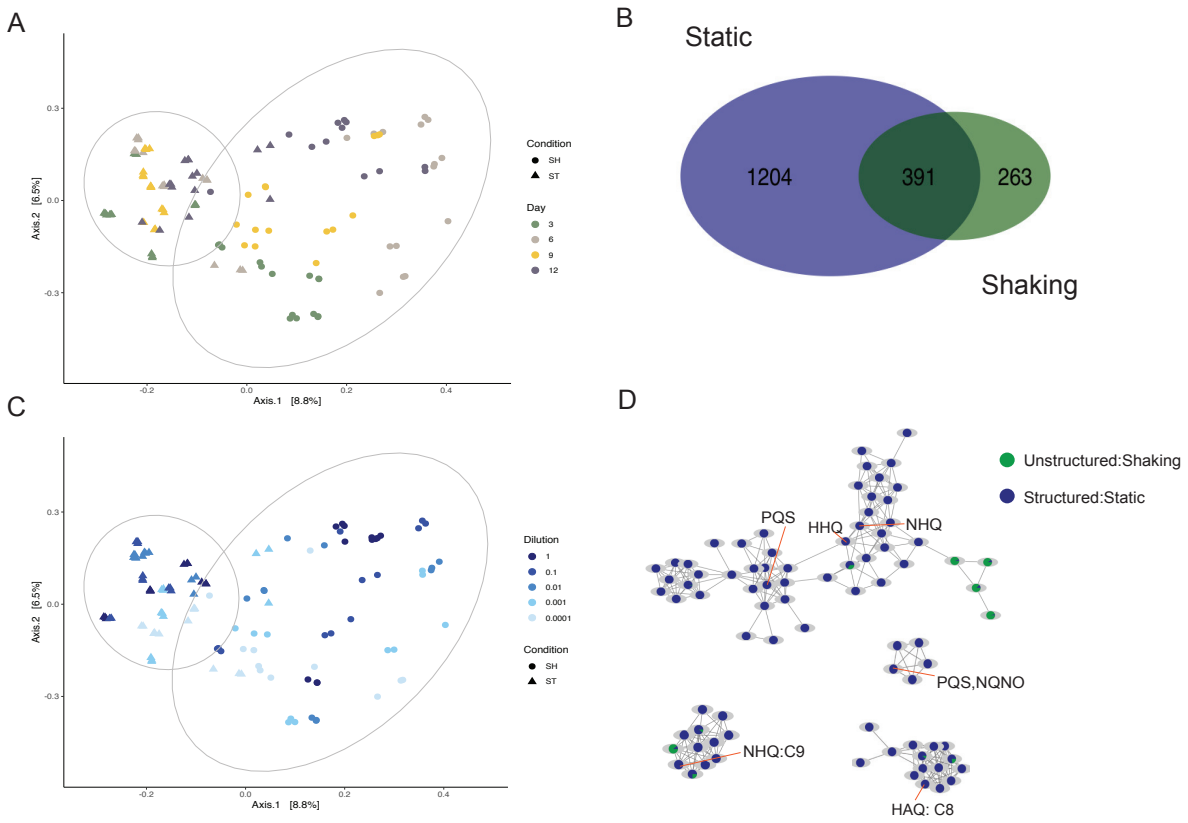
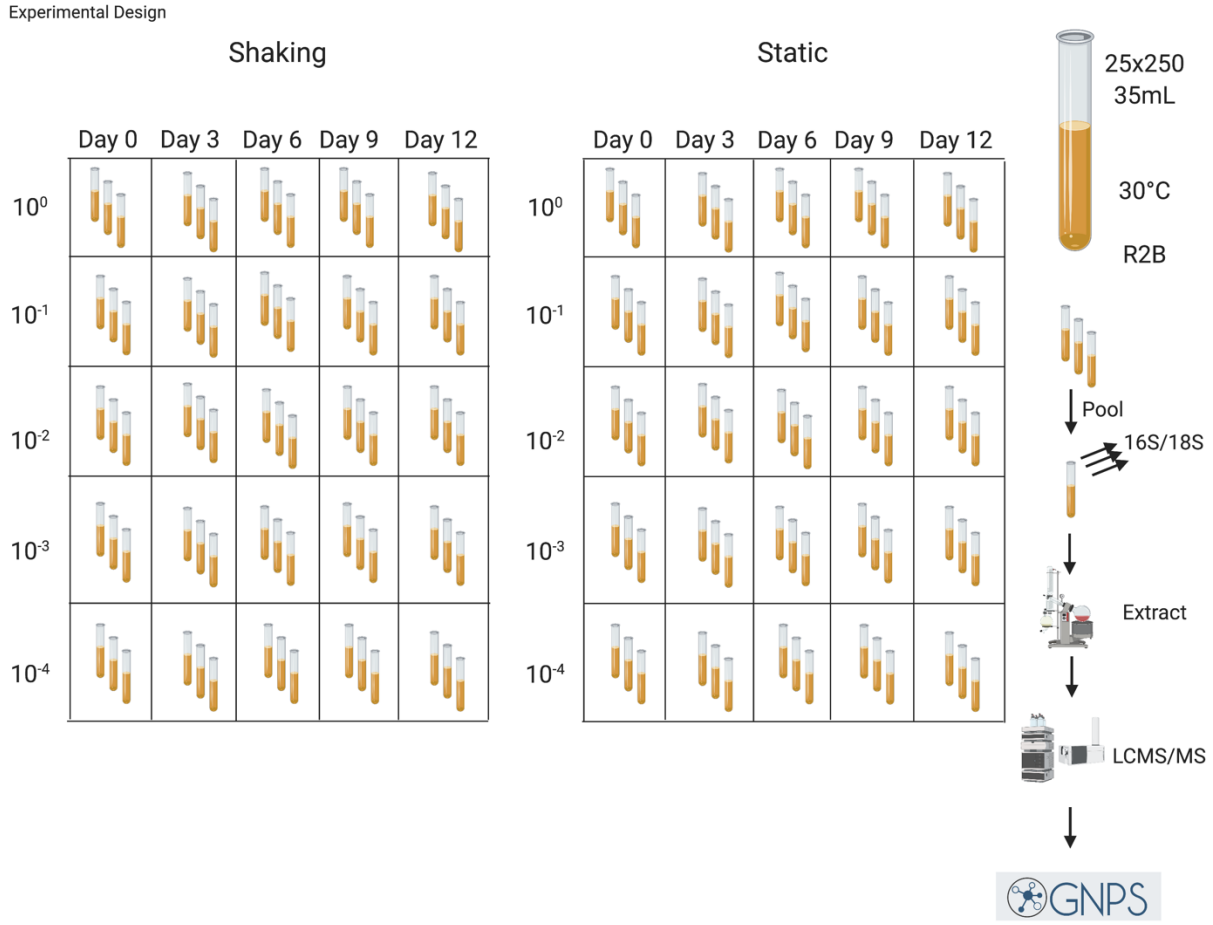


Figure 4: Principal Coordinate analysis of metabolites using Bray-Curtis distances colored by (A) time and (C) inoculum size; shape and 95% CI refer to shaking (SH) or static (ST). (B) Venn diagram of features shared or unique between cultivation conditions. (D) Networks of features Pseudomonad signaling molecules. Nodes are pie charts showing the proportion of which cultivation type a certain feature was observed. Orange call-out lines indicate an annotated node (see Supplemental Figure 6 for mirror plots). HAQ refers to 4-hydroxy-2-octylquinolone 1-oxide. NHQ is 2-nonylquinolin-4(1H)-one. HHQ is 2-heptylquinolin-4(1H)-one. PQS is *Pseudomonas* quinolone signal. NQNO is 2-nonyl-4-hydroxyquinolone *N*-oxide (Heeb et al. 2011).

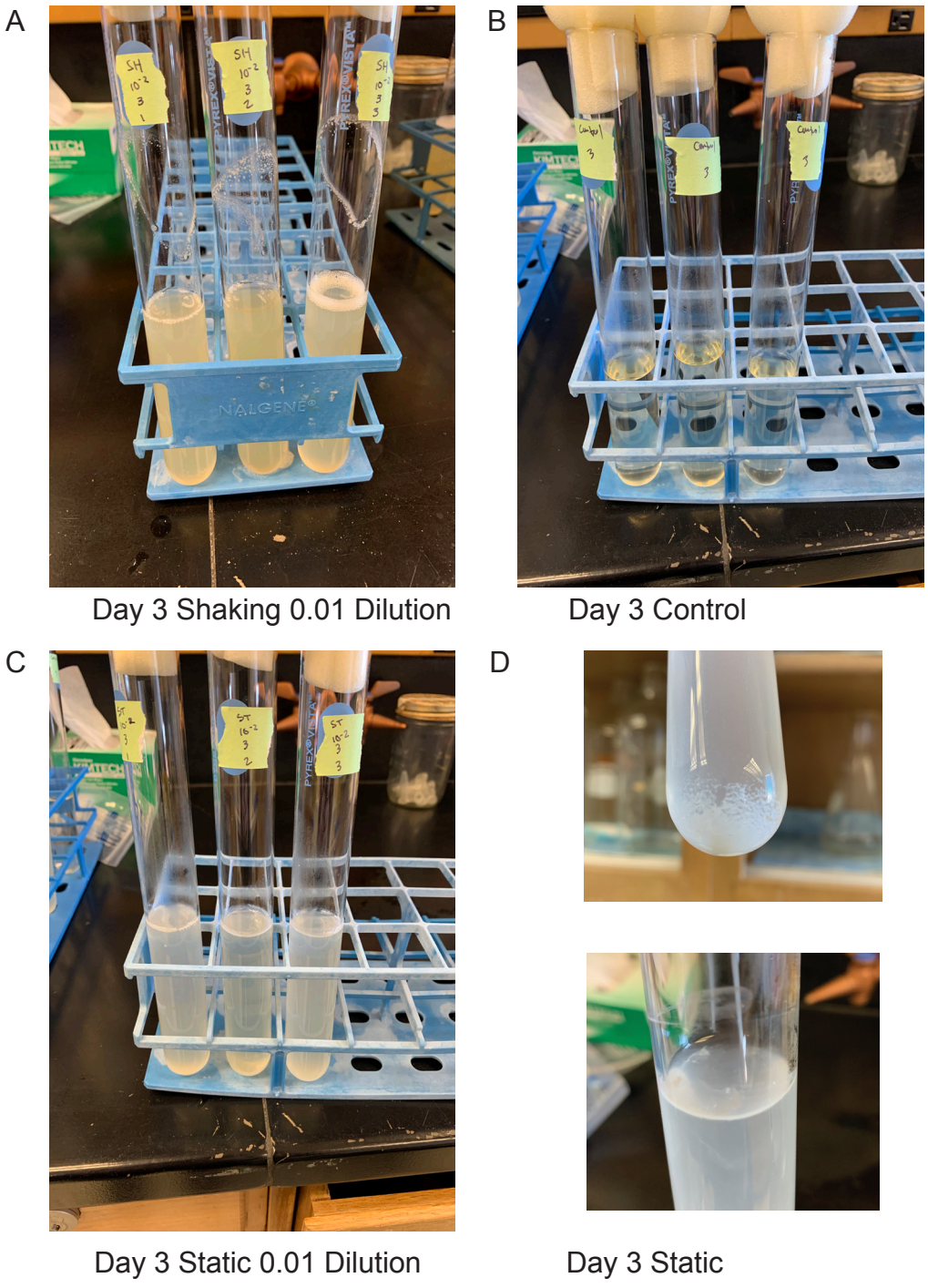
Supplemental Text

Differences in community structure were primarily driven by inoculum size (PERMANOVA F= 18.97, r2= 0.18, p=0.001; PERMDISP F=0.71, p=0.599). Both cultivation condition and day significantly contributed to community structure. However, within-group variance was high and could affect the significance contribution to community structure. Day 0 samples were tightly clustered, indicating high similarity between Day 0 samples, which was expected (Supplemental Figure 3). Because Day 0 samples represented the starting community that had not undergone any selection for which we were testing and would thus skew the distance matrix, they were removed for subsequent analyses. Like the sequence data, Day 0 metabolites clustered tightly and away from the rest of the data set. Because these metabolites represent what was in the inoculum before any selection would take place based on our hypothesis, these samples were removed from subsequent analysis (Supplemental Figure 9).

Supplemental Figures

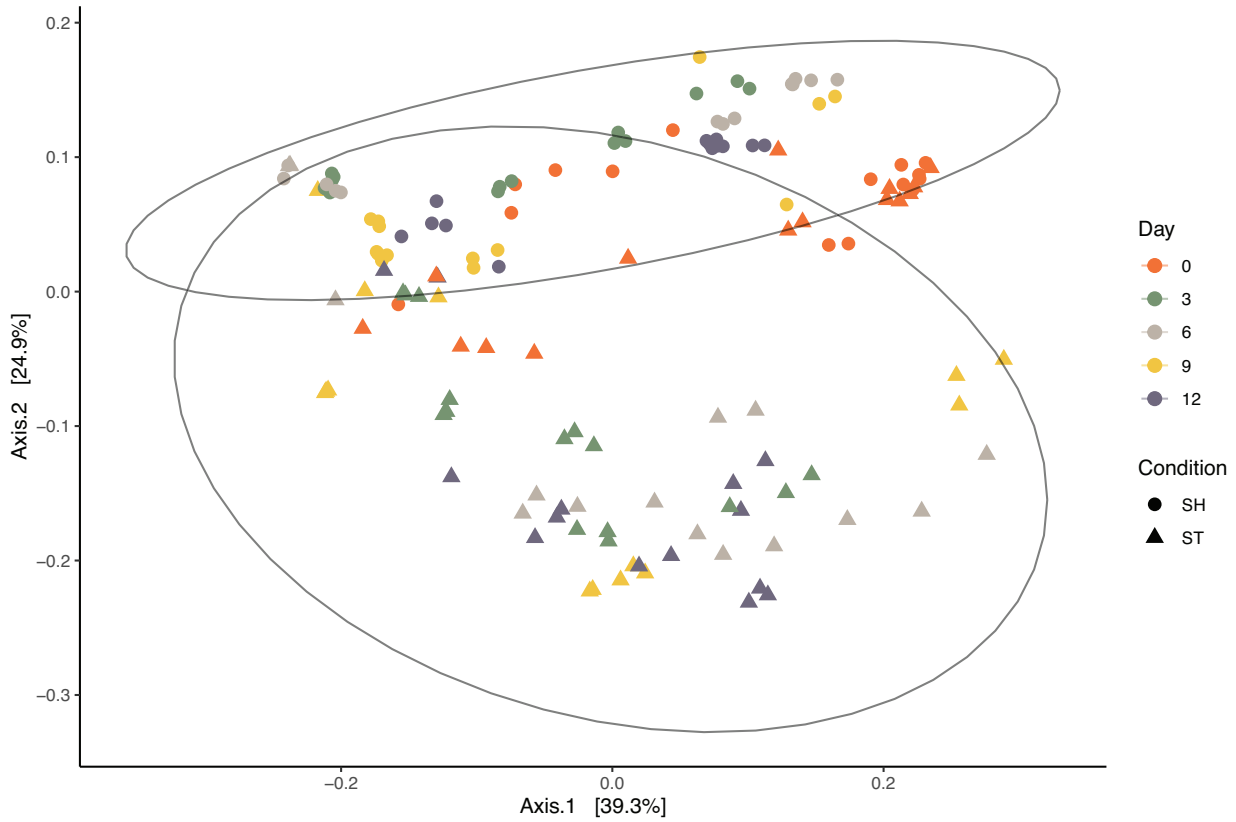


Supplemental Figure 1: Schematic of experimental setup. Created with BioRender.com.

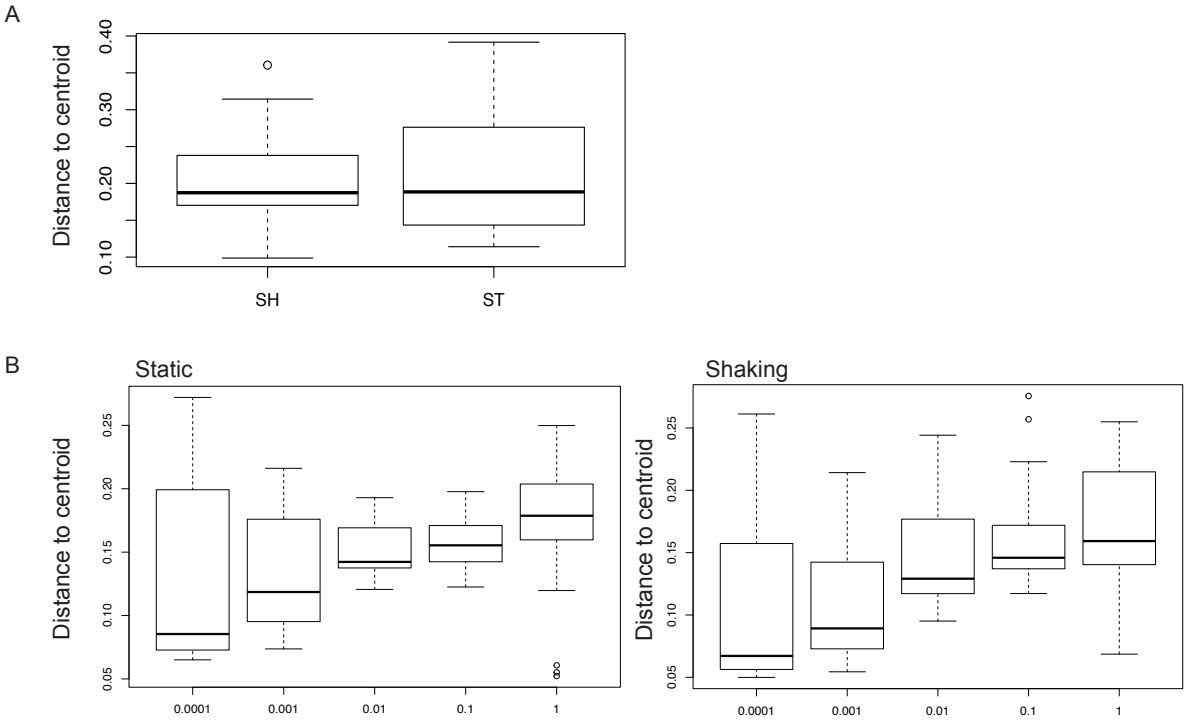


Supplemental Figure 2: Images showing biofilm formation during cultivation for (A) shaking, (B) controls, and (C) static cultures. Also shown are biofilms forming at the top or bottom of a

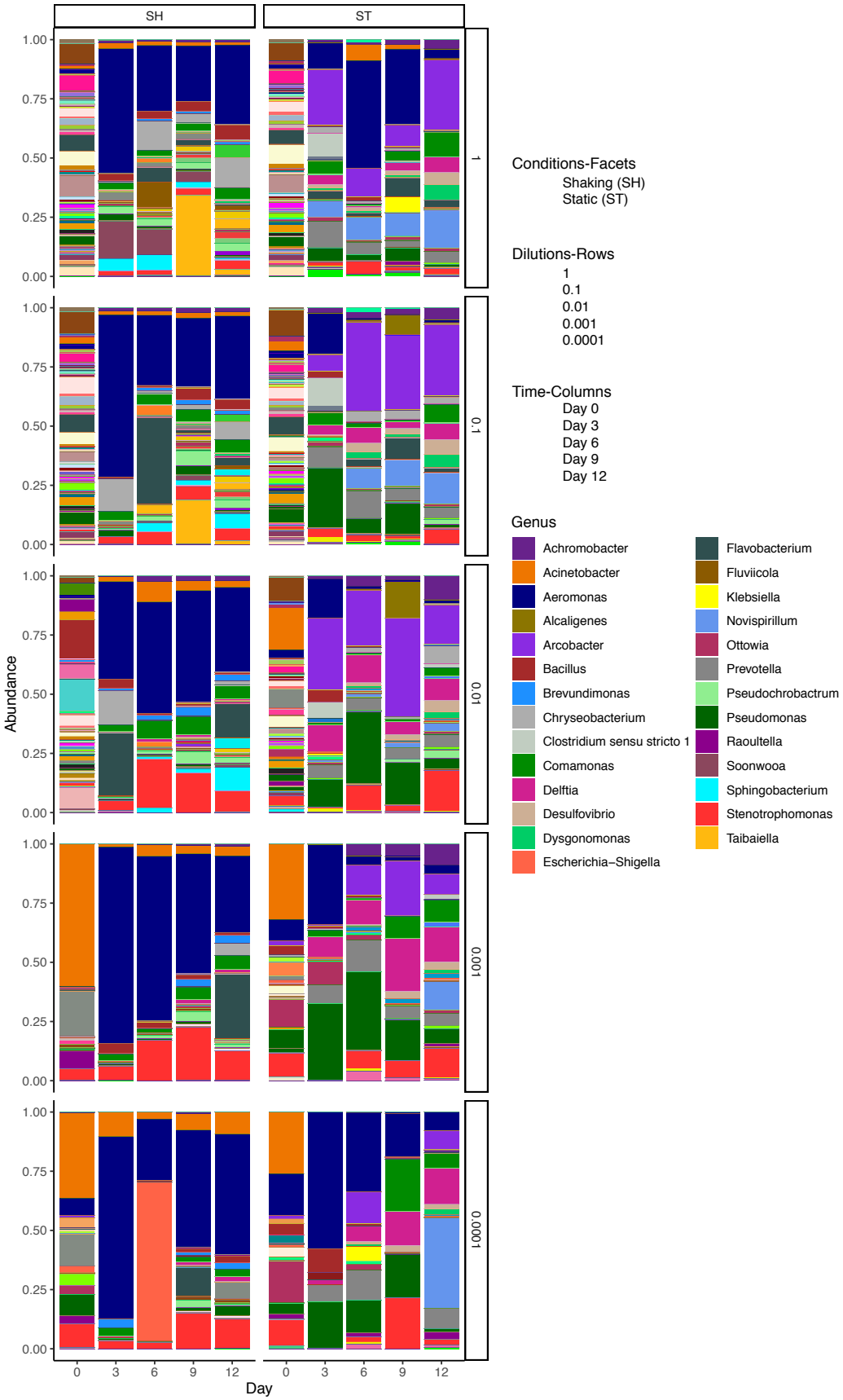
static culture (D).



Supplemental Figure 3: Principal Coordinate Analysis of microbial communities including Day 0 samples using weighted UniFrac distances. Shape and 95% confidence interval ellipses refer to cultivation condition, color refers to time. SH refers to shaking, ST refers to static.

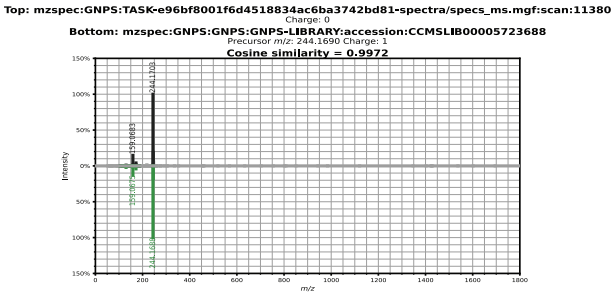


Supplemental Figure 4: Boxplots showing dispersion of microbial communities by (A) conditions or (B) inoculum size. Large ranges in boxplots indicated high variability between samples. SH refers to shaking, ST refers to static.

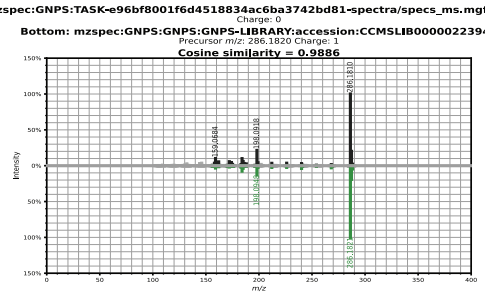

SH
ST
1
0.1
0.01
0.001
0.0001
Abundance
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Day
Conditions-Facets
Shaking (SH)
Static (ST)
Dilutions-Rows
1
0.1
0.01
0.001
0.0001
Time-Columns
Day 0
Day 3
Day 6
Day 9
Day 12
Genus
Achromobacter
Acinetobacter
Aeromonas
Alcaligenes
Arcobacter
Bacillus
Brevundimonas
Chryseobacterium
Clostridium sensu stricto 1
Comamonas
Delftia
Desulfovibrio
Dysgonomonas
Escherichia−Shigella
Flavobacterium
Fluviicola
Klebsiella
Novispirillum
Ottowia
Prevotella
Pseudochrobactrum
Pseudomonas
Raoultella
Soonwooa
Sphingobacterium
Stenotrophomonas
Taibaiella

Supplemental Figure 5: Taxonomic bar charts at genus level.

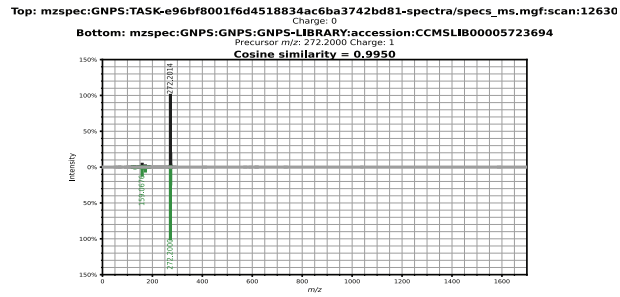
2-heptylquinolin-4(1H)-one (HHQ)



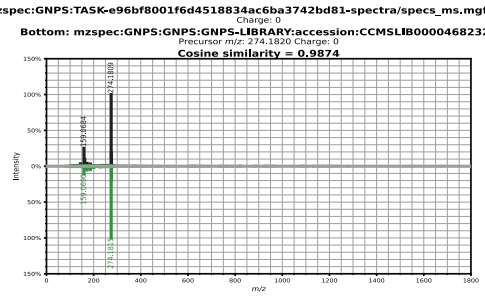
NQNO



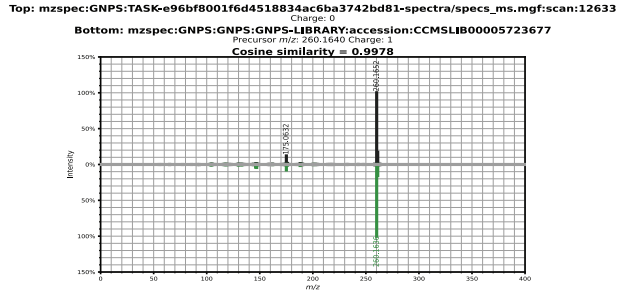
2-nonylquinolin-4(1H)-one (NHQ)



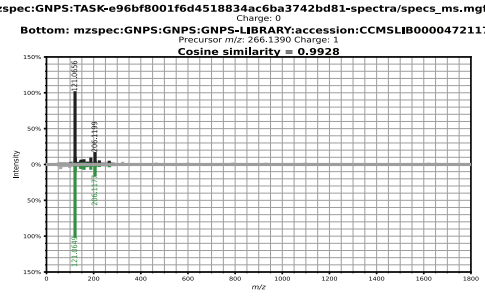
4-hydroxy-2-octylquinoline 1-oxide (HAQ)



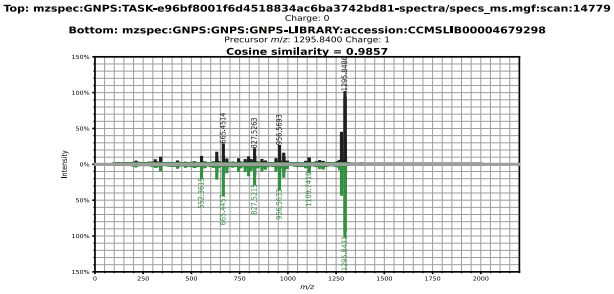
Pseudomonas quinolone signal



Anisomycin



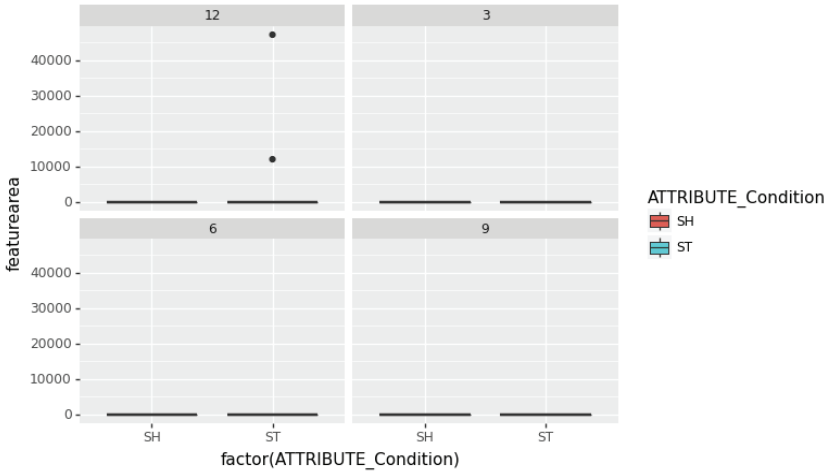
Putative Orfamide A



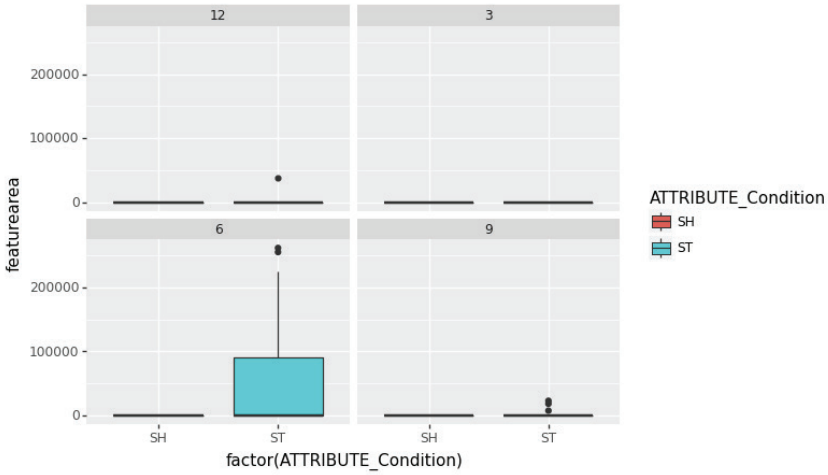
Supplemental Figure 6: Mirror plots for each feature referred to in text. Black is the experimental

spectrum while green is the database match.

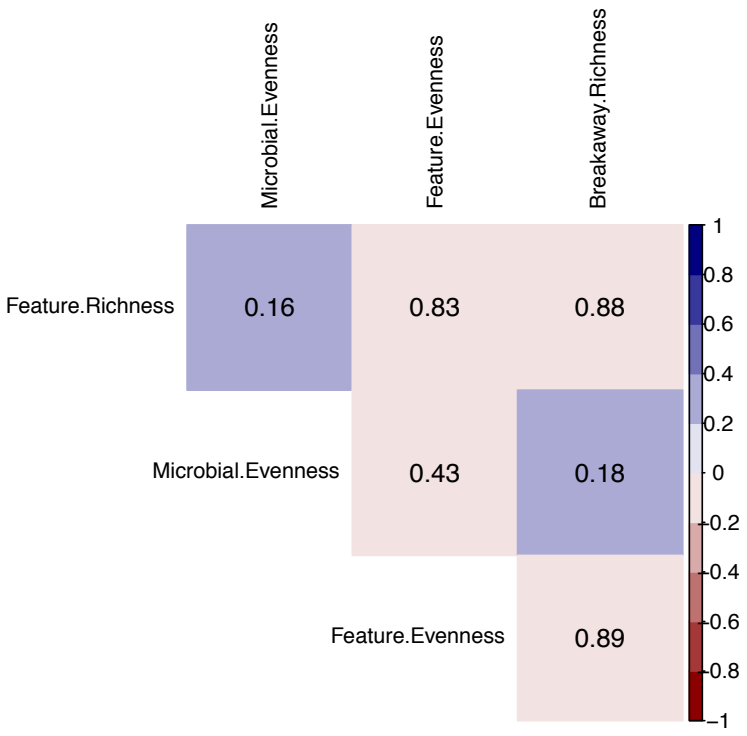
Anisomycin-19477



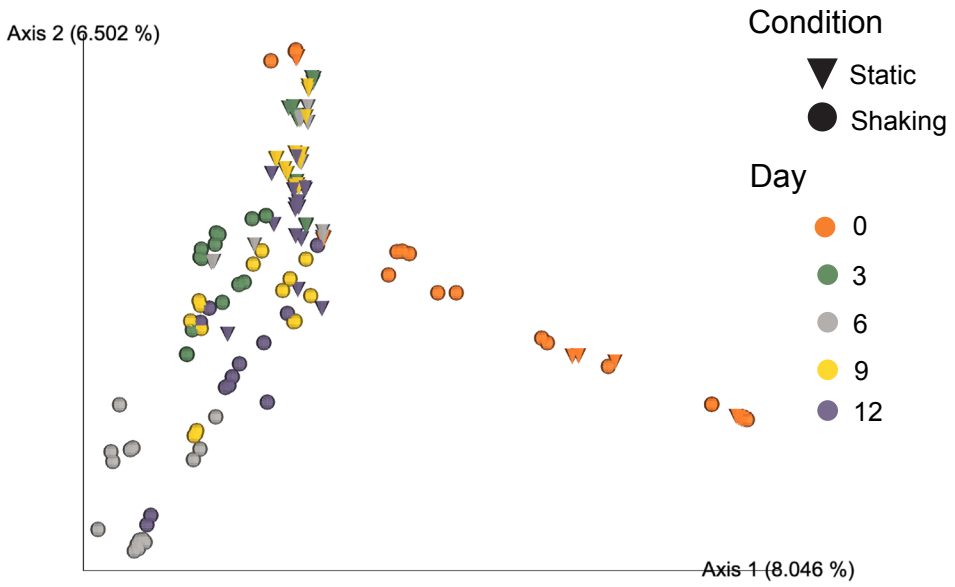
Putative Orfamide A-14779



Supplemental Figure 7: Differential abundances for features matched to anisomycin and orfamide A for shaking (SH) and static (ST) conditions. Figure generated from Feature-Based Molecular Networking (GNPS) (Nothias et al. 2020).



Supplemental Figure 8: Spearman's Rank correlation between alpha diversity metrics. Blue corresponds to positive correlations; red corresponds to negative correlations. P-values are shown in figure for each combination.



Supplemental Figure 9: Principal Coordinate Analysis of metabolomes including Day 0 samples using Bray-Curtis distances. Shape refers to cultivation condition, color refers to time. Figure was created using GNPS and QIIME2 EMPeror plotting (Bolyen et al. 2019; Vázquez-Baeza et al. 2013)

References

Assanta, M. A., D. Roy, M. J. Lemay, and D. Montpetit. 2002. "Attachment of Arcobacter Butzleri, a New Waterborne Pathogen, to Water Distribution Pipe Surfaces." *Journal of Food Protection* 65 (8): 1240–47. https://doi.org/10.4315/0362-028X-65.8.1240.

Bartley, L. E., Peck, M. L., Kim, S. R., Ebert, B., Manisseri, C., Chiniquy, D. M., Sykes, R., Gao, L., Rautengarten, C., Vega-Sánchez, M. E., Benke, P. I., Canlas, P. E., Cao, P., Brewer, S., Lin, F., Smith, W. L., Zhang, X., Keasling, J. D., Jentoff, R. E., … Ronald, P. C. (2013). Overexpression of a BAHD acyltransferase, OsAt10, alters rice cell wall hydroxycinnamic acid content and saccharification. *Plant Physiology*, *161*(4), 1615–1633. https://doi.org/10.1104/pp.112.208694

Bassler, Bonnie L., and Richard Losick. 2006. "Bacterially Speaking." *Cell* 125 (2): 237–46. https://doi.org/10.1016/j.cell.2006.04.001.

Bates, Douglas, Martin Mächler, Benjamin M. Bolker, and Steven C. Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software* 67 (1). https://doi.org/10.18637/jss.v067.i01.

Bolyen, Evan, Jai Ram Rideout, Matthew R. Dillon, Nicholas A. Bokulich, Christian C. Abnet, Gabriel A. Al-Ghalith, Harriet Alexander, et al. 2019. "Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2." *Nature Biotechnology* 37 (8): 852–57. https://doi.org/10.1038/s41587-019-0209-9.

Callahan, Benjamin J., Paul J. McMurdie, and Susan P. Holmes. 2017. "Exact Sequence Variants Should Replace Operational Taxonomic Units in Marker-Gene Data Analysis." *ISME Journal* 11 (12): 2639–43. https://doi.org/10.1038/ismej.2017.119.

Cardinale, Bradley J. 2011. “Biodiversity Improves Water Quality through Niche Partitioning.” *Nature* 472 (7341): 86–91. https://doi.org/10.1038/nature09904.

Caswell, Hal, and Joel E. Cohen. 1991. “Disturbance, Interspecific Interaction and Diversity in Metapopulations.” *Biological Journal of the Linnean Society* 42 (1–2): 193–218. https://doi.org/10.1111/j.1095-8312.1991.tb00560.x.

Chambers, Matthew C., Brendan MacLean, Robert Burke, Dario Amodei, Daniel L. Ruderman, Steffen Neumann, Laurent Gatto, et al. 2012. “A Cross-Platform Toolkit for Mass Spectrometry and Proteomics.” *Nature Biotechnology* 30 (10): 918–20. https://doi.org/10.1038/nbt.2377.

Chao, L., and B. R. Levin. 1981. “Structured Habitats and the Evolution of Anticompetitor Toxins in Bacteria.” *Proceedings of the National Academy of Sciences of the United States of America* 78 (10 I): 6324–28. https://doi.org/10.1073/pnas.78.10.6324.

Costerton, J. W., Philip S. Stewart, and E. P. Greenberg. 1999. “Bacterial Biofilms: A Common Cause of Persistent Infections.” *Science* 284 (5418): 1318–22. https://doi.org/10.1126/science.284.5418.1318.

Coyte, Katharine Z., Jonas Schluter, and Kevin R. Foster. 2015. “The Ecology of the Microbiome: Networks, Competition, and Stability.” *Science* 350 (6261): 663–66. https://doi.org/10.1126/science.aad2602.

Czárán, Tamás L., and Rolf F. Hoekstra. 2003. “Killer-Sensitive Coexistence in Metapopulations of Micro-Organisms.” *Proceedings of the Royal Society B: Biological Sciences* 270 (1522): 1373–78. https://doi.org/10.1098/rspb.2003.2338.

Dal Co, Alma, Simon van Vliet, Daniel Johannes Kiviet, Susan Schlegel, and Martin Ackermann. 2020. “Short-Range Interactions Govern the Dynamics and Functions of

Microbial Communities." *Nature Ecology and Evolution* 4 (3): 366–75. https://doi.org/10.1038/s41559-019-1080-2.

Davies, Julian, and Katherine S. Ryan. 2012. "Introducing the Parvome: Bioactive Compounds in the Microbial World." *ACS Chemical Biology* 7 (2): 252–59. https://doi.org/10.1021/cb200337h.

Davis, Nicole M., Diana M. Proctor, Susan P. Holmes, David A. Relman, and Benjamin J. Callahan. 2017. "Simple Statistical Identification and Removal of Contaminant Sequences in Marker-Gene and Metagenomics Data." *BioRxiv*, 1–14. https://doi.org/10.1101/221499.

Estrela, Sylvie, and Sam P. Brown. 2018. "Community Interactions and Spatial Structure Shape Selection on Antibiotic Resistant Lineages." *PLoS Computational Biology* 14 (6): 1–21. https://doi.org/10.1371/journal.pcbi.1006179.

Falkowski, Paul G., Tom Fenchel, and Edward F. Delong. 2008. "The Microbial Engines That Drive Earth's Biogeochemical Cycles." *Science* 320 (5879): 1034–39. https://doi.org/10.1126/science.1153213.

Figueira, Vânia, Ivone Vaz-Moreira, Márcia Silva, and Célia M. Manaia. 2011. "Diversity and Antibiotic Resistance of Aeromonas Spp. in Drinking and Waste Water Treatment Plants." *Water Research* 45 (17): 5599–5611. https://doi.org/10.1016/j.watres.2011.08.021.

Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression*. 3rd ed. Sage, Thousand Oaks CA.

Frank E Harrell Jr. 2020. "Hmisc: Harrell Miscellaneous." https://cran.r-project.org/package=Hmisc.

García-Bayona, Leonor, and Laurie E. Comstock. 2018. “Bacterial Antagonism in Host-Associated Microbial Communities.” *Science* 361 (6408). https://doi.org/10.1126/science.aat2456.

Garcia-Garcera, Marc, and Eduardo P.C. Rocha. 2020. “Community Diversity and Habitat Structure Shape the Repertoire of Extracellular Proteins in Bacteria.” *Nature Communications* 11 (1): 1–11. https://doi.org/10.1038/s41467-020-14572-x.

Ghoul, Melanie, and Sara Mitri. 2016. “The Ecology and Evolution of Microbial Competition.” *Trends in Microbiology* 24 (10): 833–45. https://doi.org/10.1016/j.tim.2016.06.011.

Granato, Elisa T., Thomas A. Meiller-Legrand, and Kevin R. Foster. 2019. “The Evolution and Ecology of Bacterial Warfare.” *Current Biology* 29 (11): R521–37. https://doi.org/10.1016/j.cub.2019.04.024.

Grollman, A. P. 1967. “Inhibitors of Protein Biosynthesis. II. Mode of Action of Anisomycin.” *Journal of Biological Chemistry* 242 (13): 3226–33. https://doi.org/10.1016/S0021-9258(18)95953-3.

Gross, Harald, Virginia O. Stockwell, Marcella D. Henkels, Brian Nowak-Thompson, Joyce E. Loper, and William H. Gerwick. 2007. “The Genomisotopic Approach: A Systematic Method to Isolate Products of Orphan Biosynthetic Gene Clusters.” *Chemistry and Biology* 14 (1): 53–63. https://doi.org/10.1016/j.chembiol.2006.11.007.

Haig, Sarah Jane, Christopher Quince, Robert L. Davies, Caetano C. Dorea, and Gavin Collinsa. 2015. “The Relationship between Microbial Community Evenness and Function in Slow Sand Filters.” *MBio* 6 (5): 1–12. https://doi.org/10.1128/mBio.00729-15.

Harrison, Freya, Jon Paul, Ruth C. Massey, and Angus Buckling. 2008. “Interspecific Competition and Siderophore-Mediated Cooperation in Pseudomonas Aeruginosa.” *ISME Journal* 2 (1): 49–55. https://doi.org/10.1038/ismej.2007.96.

Heeb, Stephan, Matthew P. Fletcher, Siri Ram Chhabra, Stephen P. Diggle, Paul Williams, and Miguel Cámara. 2011. “Quinolones: From Antibiotics to Autoinducers.” *FEMS Microbiology Reviews* 35 (2): 247–74. https://doi.org/10.1111/j.1574-6976.2010.00247.x.

Hibbing, Michael E., Clay Fuqua, Matthew R. Parsek, and S. Brook Peterson. 2010. “Bacterial Competition: Surviving and Thriving in the Microbial Jungle.” *Nature Reviews Microbiology* 8 (1): 15–25. https://doi.org/10.1038/nrmicro2259.

Hilker, Monika. 2014. “New Synthesis: Parallels Between Biodiversity and Chemodiversity.” *Journal of Chemical Ecology* 40 (3): 225–26. https://doi.org/10.1007/s10886-014-0402-8.

Ho, Adrian, Roey Angel, Annelies J. Veraart, Anne Daebeler, Zhongjun Jia, Sang Yoon Kim, Frederiek Maarten Kerckhof, Nico Boon, and Paul L.E. Bodelier. 2016. “Biotic Interactions in Microbial Communities as Modulators of Biogeochemical Processes: Methanotrophy as a Model System.” *Frontiers in Microbiology* 7 (AUG): 1–11. https://doi.org/10.3389/fmicb.2016.01285.

Hotterbeekx, An, Samir Kumar-Singh, Herman Goossens, and Surbhi Malhotra-Kumar. 2017. “In Vivo and In Vitro Interactions between Pseudomonas Aeruginosa and Staphylococcus Spp.” *Frontiers in Cellular and Infection Microbiology* 7 (APR): 1–13. https://doi.org/10.3389/fcimb.2017.00106.

Inglis, R Fredrik, Patrick Garfjeld Roberts, Andy Gardner, and Angus Buckling. 2011. “Spite and the Scale of Competition in Pseudomonas Aeruginosa.” *The American Naturalist* 178 (2): 276–85.

Janda, J. Michael, and Sharon L. Abbott. 2010. “The Genus Aeromonas: Taxonomy, Pathogenicity, and Infection.” *Clinical Microbiology Reviews* 23 (1): 35–73. https://doi.org/10.1128/CMR.00039-09.

Jang, Ja Yeong, Si Young Yang, Young Cheol Kim, Chul Won Lee, Myung Soo Park, Jin Cheol Kim, and In Seon Kim. 2013. “Identification of Orfamide A as an Insecticidal Metabolite Produced by Pseudomonas Protegens F6.” *Journal of Agricultural and Food Chemistry* 61 (28): 6786–91. https://doi.org/10.1021/jf401218w.

Justice, N. B., A. Sczesnak, T. C. Hazen, and A. P. Arkin. 2017. “Environmental Selection, Dispersal, and Organism Interactions Shape Community Assembly in High-Throughput Enrichment Culturing.” *Applied and Environmental Microbiology* 83 (20): 1–16.

Kinkel, Linda L., Daniel C. Schlatter, Kun Xiao, and Anita D. Baines. 2014. “Sympatric Inhibition and Niche Differentiation Suggest Alternative Coevolutionary Trajectories among Streptomycetes.” *ISME Journal* 8 (2): 249–56. https://doi.org/10.1038/ismej.2013.175.

Kraus, Emily A., Scott R. Beeler, R. Agustin Mors, James G. Floyd, Blake W. Stamps, Heather S. Nunn, Bradley S. Stevenson, et al. 2018. “Microscale Biosignatures and Abiotic Mineral Authigenesis in Little Hot Creek, California.” *Frontiers in Microbiology* 9 (MAY): 1–13. https://doi.org/10.3389/fmicb.2018.00997.

Ladau, Joshua, and Emiley A. Eloe-Fadrosh. 2019. “Spatial, Temporal, and Phylogenetic Scales of Microbial Ecology.” *Trends in Microbiology* 27 (8): 662–69. https://doi.org/10.1016/j.tim.2019.03.003.

Levene, H. 1960. “Robust Tests for Equality of Variances.” In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, edited by I. Olkin, 278–92. Stanford Univeristy Press.

Liébana, Raquel, Lucía Arregui, Antonio Santos, Antonio Murciano, Domingo Marquina, and Susana Serrano. 2016. “Unravelling the Interactions among Microbial Populations Found in Activated Sludge during Biofilm Formation.” *FEMS Microbiology Ecology* 92 (9): 1–13. https://doi.org/10.1093/femsec/fiw134.

Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, et al. 2001. “Ecology: Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges.” *Science* 294 (5543): 804–8. https://doi.org/10.1126/science.1064088.

Lozupone, Catherine, and Rob Knight. 2005. “UniFrac: A New Phylogenetic Method for Comparing Microbial Communities.” *Applied and Environmental Microbiology* 71 (12): 8228–35. https://doi.org/10.1128/AEM.71.12.8228-8235.2005.

Ma, Zongwang, Niels Geudens, Nam P. Kieu, Davy Sinnaeve, Marc Ongena, José C. Martins, and Monica Höfte. 2016. “Biosynthesis, Chemical Structure, and Structure-Activity Relationship of Orfamide Lipopeptides Produced by Pseudomonas Protegens and Related Species.” *Frontiers in Microbiology* 7 (MAR): 1–16. https://doi.org/10.3389/fmicb.2016.00382.

Martin, Bryan D., Daniela Witten, and Amy D. Willis. 2020. “Modeling Microbial Abundances and Dysbiosis with Beta-Binomial Regression.” *Annals of Applied Statistics* 14 (1): 94–115.

Martin, Marcel. 2011. “Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads.” *EMBnet Journal* 17 (1).

Martiny, Jennifer B.Hughes, Brendan J.M. Bohannan, James H. Brown, Robert K. Colwell, Jed A. Fuhrman, Jessica L. Green, M. Claire Horner-Devine, et al. 2006. "Microbial Biogeography: Putting Microorganisms on the Map." *Nature Reviews Microbiology* 4 (2): 102–12. https://doi.org/10.1038/nrmicro1341.

McMurdie, Paul J., and Susan Holmes. 2013. "Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data." *PLoS ONE* 8 (4). https://doi.org/10.1371/journal.pone.0061217.

McNally, Luke, and Sam P. Brown. 2015. "Building the Microbiome in Health and Disease: Niche Construction and Social Conflict in Bacteria." *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1675). https://doi.org/10.1098/rstb.2014.0298.

Meij, Anne van der, Sarah F. Worsley, Matthew I. Hutchings, and Gilles P. van Wezel. 2017. "Chemical Ecology of Antibiotic Production by Actinomycetes." *FEMS Microbiology Reviews* 41 (3): 392–416. https://doi.org/10.1093/femsre/fux005.

Morales-Soto, Nydia, Sage J.B. Dunham, Nameera F. Baig, Joseph F. Ellis, Chinedu S. Madukoma, Paul W. Bohn, Jonathan v. Sweedler, and Joshua D. Shrout. 2018. "Spatially Dependent Alkyl Quinolone Signaling Responses to Antibiotics in Pseudomonas Aeruginosa Swarms." *Journal of Biological Chemistry* 293 (24): 9544–52. https://doi.org/10.1074/jbc.RA118.002605.

Myers, Owen D., Susan J. Sumner, Shuzhao Li, Stephen Barne, and Xiuxia Du. 2017. "One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks." *Analytical Chemistry* 89 (17): 8696–8703.

Nothias, Louis Félix, Daniel Petras, Robin Schmid, Kai Dührkop, Johannes Rainer, Abinesh Sarvepalli, Ivan Protsyuk, et al. 2020. "Feature-Based Molecular Networking in the GNPS Analysis Environment." *Nature Methods* 17 (9): 905–8. https://doi.org/10.1038/s41592-020-0933-6.

Odling-Smee, John F., and Kevin N. Laland. 1996. "Niche Construction." *The American Naturalist* 147 (4): 641–48.

Oksanen, Jari, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, and Eduard Szoecs and Helene Wagner Peter Solymos, M. Henry H. Stevens. 2019. "Vegan: Community Ecology Package." R.

Oliveira, Nuno M., Rene Niehus, and Kevin R. Foster. 2014. "Evolutionary Limits to Cooperation in Microbial Communities." *Proceedings of the National Academy of Sciences of the United States of America* 111 (50): 17941–46. https://doi.org/10.1073/pnas.1412673111.

Overmann, Jörg, and Hans van Gemerden. 2000. "Microbial Interactions Involving Sulfur Bacteria: Implications for the Ecology and Evolution of Bacterial Communities." *FEMS Microbiology Reviews* 24 (5): 591–99. https://doi.org/10.1016/S0168-6445(00)00047-4.

Pacala, Stehpen W., and Simon A. Levin. 1997. "Biologically Generated Spatial Pattern and the Coexitence of Competing Species." In *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, edited by David Tilman and Peter Kareiva.

Parada, Alma E., David M. Needham, and Jed A. Fuhrman. 2015. "Every Base Matters: Assessing Small Subunit RRNA Primers for Marine Microbiomes with Mock Communities, Time Series and Global Field Samples." *Environmental Microbiology* 18 (5): 1403–14.

Picketts, S. T.A., and M. L. Cadenasso. 1999. “Landscape Ecology: Spatial Heterogeneity in Ecological Systems.” *NCASI Technical Bulletin* 2 (781): 420.

Pielou, E.C. 1966. “The Measurement of Diversity in Different Types of Biological Collections.” *Journal of Theoretical Biology* 13: 131–44.

Pluskal, Tomáš, Sandra Castillo, Alejandro Villar-Briones, and Matej Orešič. 2010. “MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data.” *BMC Bioinformatics* 11. https://doi.org/10.1186/1471-2105-11-395.

Quast, Christian, Elmar Pruesse, Pelin Yilmaz, Jan Gerken, Timmy Schweer, Pablo Yarza, Jörg Peplies, and Frank Oliver Glöckner. 2013. “The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools.” *Nucleic Acids Research* 41 (D1): 590–96. https://doi.org/10.1093/nar/gks1219.

R Core Team. 2018. “R: A Language and Environment for R Foundation for Statistical Computing,.” https://www.r-project.org/.

Raguso, RA, AA Agrawal, AE Douglas, G Jander, A Kessler, K Poveda, and JS Thaler. 2015. “The Raison d ’ e ˆ Tre of Chemical Ecology.” *Ecology* 96 (3): 617–30.

Rainey, Paul B., and Michael Travisano. 1998. “Adaptive Radiation in a Heterogeneous Environment.” *Nature* 394 (6688): 69–72. https://doi.org/10.1038/27900.

Reasoner, D. J., and E. E. Geldreich. 1985. “A New Medium for the Enumeration and Subculture of Bacteria from Potable Water.” *Applied and Environmental Microbiology* 49 (1): 1–7. https://doi.org/10.1128/aem.49.1.1-7.1985.

Rema, Tara, John R. Lawrence, James J. Dynes, Adam P. Hitchcock, and Darren R. Korber. 2014. “Microscopic and Spectroscopic Analyses of Chlorhexidine Tolerance in Delftia

Acidovorans Biofilms." *Antimicrobial Agents and Chemotherapy* 58 (10): 5673–86. https://doi.org/10.1128/AAC.02984-14.

Schliep, Klaus Peter. 2011. "Phangorn: Phylogenetic Analysis in R." *Bioinformatics* 27 (4): 592–93. https://doi.org/10.1093/bioinformatics/btq706.

Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. "Cytoscape: A Software Environment for Integrated Models." *Genome Research* 13: 2498–2504. https://doi.org/10.1101/gr.1239303.metabolite.

Skwor, Troy, Sarah Stringer, Jason Haggerty, Jenilee Johnson, Sarah Duhr, Mary Johnson, Megan Seckinger, and Maggie Stemme. 2020. "Prevalence of Potentially Pathogenic Antibiotic-Resistant Aeromonas Spp. in Treated Urban Wastewater Effluents versus Recipient Riverine Populations: A 3-Year Comparative Study." *Applied and Environmental Microbiology* 86 (3): 1–16. https://doi.org/10.1128/AEM.02053-19.

Sobin, Ben A., and Fred W. Tanner Jr. 1954. "Anisomycin, A New Anti-Protozoan Antibiotic." *Journal of American Chemical Society* 76 (15): 4053–4053.

Thierbach, Sven, Max Wienhold, Susanne Fetzner, and Ulrich Hennecke. 2019. "Synthesis and Biological Activity of Methylated Derivatives of the Pseudomonas Metabolites HHQ, HQNO and PQS." *Beilstein Journal of Organic Chemistry* 15: 187–93. https://doi.org/10.3762/bjoc.15.18.

Tolker-Nielsen, T., and S. Molin. 2000. "Spatial Organization of Microbial Biofilm Communities." *Microbial Ecology* 40 (2): 75–84. https://doi.org/10.1007/s002480000057.

Traxler, Matthew F., and Roberto Kolter. 2015. “Natural Products in Soil Microbe Interactions and Evolution.” *Natural Product Reports* 32 (7): 956–70. https://doi.org/10.1039/c5np00013k.

Vázquez-Baeza, Yoshiki, Meg Pirrung, Antonio Gonzalez, and Rob Knight. 2013. “EMPeror: A Tool for Visualizing High-Throughput Microbial Community Data.” *GigaScience* 2 (1): 2–5. https://doi.org/10.1186/2047-217X-2-16.

Venturi, Vittorio. 2006. “Regulation of Quorum Sensing in Pseudomonas.” *FEMS Microbiology Reviews* 30 (2): 274–91. https://doi.org/10.1111/j.1574-6976.2005.00012.x.

Wang, M., J. J. Carver, V. v. Phelan, L. M.; Sanchez, N.; Garg, Y.; Peng, D. D.; Nguyen, et al. 2017. “Sharing and Community Curation of Mass Spectrometry Data with GNPS.” *Nature Biotechnology* 34 (8): 828–37. https://doi.org/10.1038/nbt.3597.Sharing.

Wang, Xiao, Xiaolan Li, and Junqi Ling. 2017. “Streptococcus Gordonii LuxS/Autoinducer-2 Quorum-Sensing System Modulates the Dual-Species Biofilm Formation with Streptococcus Mutans.” *Journal of Basic Microbiology* 57 (7): 605–16. https://doi.org/10.1002/jobm.201700010.

Wei, Taiyun, and Viliam Simko. 2017. “R Package ‘Corrplot’: Visualization of a Correlation Matrix.”

Wellborn, Gary A., and R. Brian Langerhans. 2015. “Ecological Opportunity and the Adaptive Diversification of Lineages.” *Ecology and Evolution* 5 (1): 176–95. https://doi.org/10.1002/ece3.1347.

West, Stuart A., Ashleigh S. Griffin, Andy Gardner, and Stephen P. Diggle. 2006. “Social Evolution Theory for Microorganisms.” *Nature Reviews Microbiology* 4 (8): 597–607. https://doi.org/10.1038/nrmicro1461.

Willis, Amy, and John Bunge. 2015. “Estimating Diversity via Frequency Ratios.” *Biometrics* 71 (4): 1042–49. https://doi.org/10.1111/biom.12332.

Wittebolle, Lieven, Massimo Marzorati, Lieven Clement, Annalisa Balloi, Daniele Daffonchio, Kim Heylen, Paul de Vos, Willy Verstraete, and Nico Boon. 2009. “Initial Community Evenness Favours Functionality under Selective Stress.” *Nature* 458 (7238): 623–26. https://doi.org/10.1038/nature07840.

Wright, Erik S. 2015. “DECIPHER: Harnessing Local Sequence Context to Improve Protein Multiple Sequence Alignment.” *BMC Bioinformatics* 16 (1): 1–14. https://doi.org/10.1186/s12859-015-0749-z.

Wright, Erik S. 2016. “Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R.” *R Journal* 8 (1): 352–59. https://doi.org/10.32614/rj-2016-025.

Yanni, David, Pedro Márquez-Zacarías, Peter J. Yunker, and William C. Ratcliff. 2019. “Drivers of Spatial Structure in Social Microbial Communities.” *Current Biology* 29 (11): R545–50. https://doi.org/10.1016/j.cub.2019.03.068.

Yilmaz, Pelin, Laura Wegener Parfrey, Pablo Yarza, Jan Gerken, Elmar Pruesse, Christian Quast, Timmy Schweer, Jörg Peplies, Wolfgang Ludwig, and Frank Oliver Glöckner. 2014. “The SILVA and ‘All-Species Living Tree Project (LTP)’ Taxonomic Frameworks.” *Nucleic Acids Research* 42 (D1): 643–48. https://doi.org/10.1093/nar/gkt1209.

Zhou, Jizhong, and Daling Ning. 2017. “Stochastic Community Assembly: Does It Matter in Microbial Ecology?” *Microbiology and Molecular Biology Reviews* 81 (4): 1–32.

Chapter 4: Impact of Repeated Disturbance of Cultivated Microbial Communities on Community Composition, Antimicrobial Resistance, and Metabolisms

Foreword

The research in Chapter 4 was conducted by Emily N. Junkins with the help of undergraduate researcher Emma Kasahara who aided in sample collection, media preparation, and CFU data generation. Emily N. Junkins performed all molecular work and metabolomic sample preparation. All microbiome data analysis was performed by Emily N. Junkins. All metabolome data analysis was performed by Emily N. Junkins under the guidance of Dr. Laura-Isobel McCall. Dr. Vincent Bonifay and Dr. Steven Foster at the OU Mass Spectrometry, Proteomics and Metabolomics facility provided instrumentation for LCMS/MS analysis and guidance when writing methods. Dr. Graham Wiley at University of Oklahoma Consolidated Core Laboratories provided sequencing support and instrumentation.

Abstract

Disturbance is a common, shared experience of microbial communities. A disturbance will either alter a community directly such as changing the distribution of its members or indirectly, such as altering resources of the environment. Either way, populations will seek to re-assemble post-disturbance, and any fitness advantages could make certain populations more successful. Some strategies could include increased growth rate or the production of secondary metabolites that bestow an advantage to the producer. Despite many efforts seeking to characterize microbial assembly in response to disturbance, few studies have described both the microbial response and secondary metabolic response to disturbance in cultivated mesocosms. Here, we repeatedly disturbed a cultivated wastewater community via ten transfers into soft agar matrix over the course of one month. We sought to determine which taxa were most successful during repeated disturbance; and to characterize the changes in the metabolome in response to repeated disturbance, with the prediction that organisms that were both prolific and capable of producing secondary metabolites would prevail. We observed that disturbance selected for different community assembly patterns but overall antimicrobial resistant members; we also observed that changes in the microbial community structure were linked to changes in the metabolome structure. We observed the presence of certain taxa co-occuring with the presence of certain metabolites. These data suggest that changes in the microbiome and metabolome are linked as communities seek to re-assemble in response to disturbance, and that variable outcomes can result from identical starting conditions. These data also describe potential ecological mechanisms used by certain populations to proliferate post-disturbance, further characterizing chemical ecology of microbial communities.

Introduction

One collective experience of microbial communities is disturbance. A disturbance is an event that either directly alters a community or changes the environment, indirectly altering the community. Disturbances can be continuous events (press) or discrete events (pulse) (Bender et al., 1984). For microbial communities, disturbances can describe the change in environmental conditions or nutrients (Antonopoulos et al., 2009) or a mixing of spatial arrangements and distributions of organisms within a community (Shade, Read, et al., 2012). In both, habitat structure can affect the resiliency of communities to disturbance by partitioning the effects of disturbance. Resiliency to disturbance starts at the individual level, which determines success at the population level and eventually the community level (reviewed (Shade, Peter, et al., 2012)). Disturbance can lead to a variety of outcomes, some of which will depend on the rate of disturbance (e.g., intermediate disturbance hypothesis) or the traits of community members (reviewed (Shade, Peter, et al., 2012)). Though microbial ecology has typically studied disturbance to determine the mechanism of response and predict the outcome (Shade, 2018; Shade, Peter, et al., 2012), the variability of responses to disturbance should be viewed as an outcome itself (Fraterrigo & Rusak, 2008). This variability is influenced by the members of the starting community (e.g., historical contingency) and the responses to disturbance can result in 'alternative stable states' (Blaustein & Margalit, 1996; Chase, 2003; Samuels & Drake, 1997), where compositional outcomes in response to increased disturbance vary under identical conditions (Belyea & Lancaster, 1999); though the alternative is also true where increased disturbance can continually select for one compositional outcome (Jiang & Patel, 2008). For a microbial community, disturbance results in re-assembly; and for researchers, an opportunity to characterize drivers of microbial community assembly.

A central goal in microbial ecology is to characterize the drivers of community assembly (selection, drift, dispersal, and diversification) (Nemergut et al., 2013; Vellend, 2010) across different systems and in some cases, in response to disturbance (Ferrenberg et al., 2013; Nemergut et al., 2013). For example, dispersal resulting from disturbance can invoke priority effects, where the order of arrival will determine success; it can also lead to founder's effects, where the success of certain populations is dependent on the success of the previous generation (Fukami, 2015). Further, selection resulting from disturbance can be abiotic, where environmental filtering selects for certain groups based on changes in nutrient availability or habitat structure (Andersen et al., 2013; Hargreaves et al., 2015; Yan et al., 2019). Selection resulting from disturbance can also be biotic, where members of the community are forced to re-assemble and interact (Shade, Read, et al., 2012), and in doing so could be forced to either compete for nutrients and space, or seek out different niches.

Many of these interactions are mediated by secondary metabolites and are used for both communication and competition (Raguso et al., 2015). Studying microbial interactions of whole communities relies primarily on interaction networks (Faust et al., 2012; Steele et al., 2011), and metabolic networks based on primary metabolism (Lawson et al., 2017) and secondary metabolism (Ali et al., 2020). Interactions also can drive a community's response to disturbance (Bissett et al., 2013; Hunt & Ward, 2015). As the tools to characterize the metabolome become more ubiquitous in microbiome research, studies have started to characterize the response of the metabolome to disturbance (Antunes et al., 2011; Bowers et al., 2020; Hossain et al., 2019; McCall et al., 2018), though fewer studies focus on the response of the metabolome to repeated

disturbance outside of host-associated systems, and as of this study, none have looked at cultivated communities and repeated disturbance. Here, we continually disturbed a cultivated wastewater community via ten transfers into soft agar matrix over the course of one month. Each transfer represented a successive selection event that builds on previous outcomes (i.e., founder's effects).We sought to determine which taxa were most successful during repeated disturbance; and to characterize the changes in the metabolome in response to repeated disturbance. We hypothesized that the repeated selection events within a structured habitat would also favor taxa that could quickly adapt. One potential mechanism would be through the production of secondary metabolites. Understanding the metabolite changes in response to disturbance can help determine the mechanisms behind community assembly patterns which long-term lead to the discovery of new chemical strategies.

Materials and Methods

Sampling and Experimental Set Up

Wastewater was collected from the aerator basin at the Norman Water Reclamation Facility, Norman, OK on 4 January 2020, stored on ice, and transported immediately back to the laboratory (< 30 minutes). The sample (50 mL volume) was homogenized vigorously on a vortex mixer for ~2 minutes to break up flocks and produce a slurry. A 5.0 mL sample was serially diluted (1:10) in sterile R2B liquid medium (Reasoner & Geldreich, 1985). Based on optimization serial diluting and plating experiments, the 1:1000 dilution was used for subsequent analyses because it generated countable colonies when plated. From the prepared sample, 183 µL was transferred to bead tubes in triplicate to characterize the starting inoculum prior to setting up the soft agar plates. These soft agar plates were prepared by using sterile 100 mL petri dishes filled with 100 mL sterile R2A agar (1.5% agar) as a base. The overlay consisted of 49.375 mL of soft R2A (0.7%) inoculated with 635 µL of culture for a final volume of 50 mL.

In addition to large petri plates, standard petri dishes were used to make R2A (1.5% agar) soft agar plates for antibiotic resistance screen. The soft agar overlay was prepared with 3.946 mL R2A (0.7% agar), 50 µL inoculum, and 4 µL of antibiotic or control (sterile water). Antibiotics used were ampicillin (100 µg/mL), tetracycline (10 µg/mL), ciprofloxacin (10 µg/mL). These antibiotics were chosen based on differential mechanisms of action and resistance. Ampicillin targets cell wall synthesis (Campoli-richards & Brogden, 1987), tetracycline targets protein synthesis (Chopra & Roberts, 2001), and ciprofloxacin targets DNA replication (LeBel, 1988).

For the initial plate set up, called T0 or Transfer zero, 652 µL of the prepared 10^{-3} dilution of wastewater was added to the soft agar tubes, vortexed, and poured onto petri dishes (n=6). Plates were cultivated at 30°C for three days. After cultivation, three plates were used to represent the subsequent biological replicates (called lineage A, B or C) while the remaining three were used for colony picking.

After cultivation, one plate from each lineage (A, B, C) was broken down for propagation. To do this, the soft agar layer was sterilely scraped into chemically clean glassware. A volume of 100 mL of sterile R2B liquid medium was added to the soft agar to aid in homogenization. A steel wood boring bit was flame sterilized and a drill was used to homogenize each sample. From each sample (A B, and C), triplicate 750 µL aliquots were taken for 16S rRNA gene sequencing. A volume of 300 µL of the homogenized sample was diluted 1:1000. From each diluted sample, duplicate 500 µL aliquots were transferred to 500 µL 40% glycerol for storage at -80°C, and duplicate 625 µL aliquots were taken to be used as the inoculum for the subsequent overlay, as described above for the initial overlay. This generated two large petri plates for each A, B, or C lineage. The remaining plate served as a backup if needed. Lastly, a volume of 250 µL was taken from each diluted sample, pooled to a final volume of 750 µL, and further serially diluted 1:10 in sterile R2B medium. Each dilution was plated in a soft agar overlay in triplicate containing each antibiotic or control, described above. Plates were incubated at 30°C and colonies were counted at 24h, 48h, 72h, and 96h. Colony forming units were calculated for each time point and antibiotic or control. Lastly, once CFUs were calculated, colonies were picked for isolation by using a sterile toothpick to transfer each colony to a deep 96 well plate containing 800 µL of sterile R2B medium and incubated for 3 days at 30°C. After

incubation 500 µL of sterile 40% glycerol was added to each well and isolates were stored at -80°C.

Once 16S rRNA gene sequencing samples and the aliquot for propagation and the antibiotic screen were collected, the remaining homogenized sample was extracted for metabolite analysis. An overnight organic extraction was performed by adding 100 mL of 100% ethyl acetate to the sample. See Supplemental Figure 1 for a visualization of overall experiment set-up.

Controls for cultivation included a large, uninoculated soft agar plate per transfer to check for contamination of the medium. In addition to kit reagent controls, sequencing controls consisted of one large, uninoculated soft agar plate collected and processed as described above. Controls for metabolomics consisted of glassware extraction blanks, one large, uninoculated soft agar plate for medium blank, methanol control and ethyl acetate control. Controls for the antibiotic screen consisted of uninoculated, soft agar overlays for each antibiotic and negative control (sterile water) per transfer.

DNA Extraction and Sequencing

Before DNA extraction, each sample was thawed on ice and homogenized for 30 seconds at maximum speed using a bead beat for 45 seconds using a hand-held reciprocating saw. DNA was extracted according to manufacturer specifications using the Zymo Quick-DNA Fungal/Bacterial Miniprep kit (Cat# D6005, Zymo Research Corp., Irvine, CA). Samples were stored at -20°C until all samples could be extracted at once. For community characterization, a conserved region of the SSU rRNA gene of most bacteria, archaea, and eukarya was amplified

using primers 515F-Y and 926R (Parada et al., 2015) via the following PCR protocol: initial denaturation at 94˚C for 2 min, followed by 30 cycles of denaturation at 94˚C for 45 s, annealing at 50˚C for 45 s, and extension at 68˚C for 90 s, with a final extension at 68˚C for 5 min. These primers produced two amplicons, a ~400 bp fragment for bacteria and archaea, and a 600 bp fragment for eukaryotes. The forward primer 515F-Y (5'-GTA AAA CGA CGG CCA G CCG TGY CAG CMG CCG CGG TAA-3') contains the M13 tag (underlined) fused to the 5' end of the forward primer (Kraus et al., 2018). The reverse primer 926R (5'-CCG YCA ATT YMT TTR AGT TT-3') was unmodified from Parada et. al 2015. Each PCR contained 5 PRIME HOT master mix (1X; 5 PRIME Inc., Gaithersburg, MD), 0.2 µM of each primer, and 3.0 µL of extracted DNA at a final volume of 50 µL. The amplified fragments were purified using Sera-Mag magnetic beads (GE) with the AmPureXP (Beckman Coulter) protocol at a final concentration of 1.8x v/v. After purification, 3 µL of each PCR product, 1x 5 PRIME HOT master mix (Quantabio, Massachusetts, USA), 0.2µM of the 926R primer, and 0.2 µM of a specific 12 bp oligonucleotide was used in a separate barcoding PCR (6 cycles) in 50 µL reactions to attach a unique barcode to amplicons of each library. The same thermocycler protocol was used as above but only run for 6 cycles. The now barcoded amplicons were purified using Sera-Mag (GE) beads with the AmPureXP (Beckman Coulter) protocol to a final volume of 40 µL, quantified using the QuBit HS DS DNA assay kit (Thermo Fisher Scientific Inc., Waltham, MA), and pooled in equimolar amounts. The pooled, barcoded amplicon libraries were then concentrated to a final volume of 100 µL (16.6 ng/ µL) with an Amicon-Ultra filter (Millipore, Burlington, MA, USA) following manufacturer's protocol. The combined amplicon libraries were denatured according to Miseq library preparations protocol (Illumina, San Diego,

CA, USA). The sample was loaded at a concentration of 10 pM and sequenced using 2x250 paired-end strategy on the Miseq (Illumina San Diego, CA, USA) platform for 251 cycles.

Organic Extractions for LCMS/MS

After overnight ethyl acetate extractions, the organic layer was transferred to round-bottom flasks and evaporated on a Heidolph Hei-Vap Value (Heidolph) rotovap. Samples were resuspended in 100% methanol and transferred to pre-weighed scintillation vials. Samples were dried under nitrogen gas and weighed for final yield (mg). For this experiment, all beakers were brand-new from the same lot, rinsed with 100 % methanol prior to use, and one beaker was incubated overnight with 103 mL of ethyl acetate to generate a glassware blank to attribute any features in the metabolomics data to glassware alone. Once samples were dry, they were stored at -80°C until experiment was complete.

After all metabolite samples were dried and stored at -80°C, each sample was resuspended to a final concentration of 11 mg/mL in 100% isopropanol containing 0.1 µM of sulfadimethoxine as an internal standard. A volume of 100 µL of each sample was transferred to a LCMS/MS 96-well plate and stored at -80°C until it was run on the instrument. In addition to the internal control, a pooled sample control containing 2 µL of every sample and a 100% isopropanol with internal standard were also prepared. Lastly, 100% methanol blanks were run between each experimental sample.

HPLC-ESI Mass Spectrometry

HPLC separation of the samples (5µL injection volume at 11 mg/mL) was performed on

an Agilent 1290 HPLC system with a Waters ACQUITY UPLC BEH column (ODS-18; 2.1 x 100 mm; 1.7 μm particle size). Mobile phase A was comprised of 0.1% formic acid in Optima LC/MS water (Fisher), while mobile phase B was comprised of 0.1% formic acid in 100% HPLC-grade acetonitrile. A binary gradient at 0.4 mL/min flow rate was used under the following steps: 90% solvent A and 10% solvent B from 0 to 1 min, linear gradient to 20% solvent B from 1-2 min, linear gradient to 80% solvent B from 2-16 min, linear gradient to 100% solvent B from 16-18 min, and 100% solvent B from 20 – 21 min, which was maintained for an additional 2 min before the next sample injection. The resulting eluent from the HPLC column was run on an Agilent 6545 accurate mass Q-TOF mass spectrometer with an electrospray ionization source operating in the positive mode. Nitrogen was used as a nebulizing gas (40 lbs/in^2) and as a drying gas (325 °C; 10 L/min flow rate)(Bartley et al., 2013). Fragmentor voltage was 180 V, skimmer voltage was 45 V, and capillary voltage was 4,000 V. Untargeted MS/MS was used with mass-dependent collision energy ramp from 20-35 V using UHP Nitrogen. Data was collected with Mass Hunter Acquisition software (B.08.00) and analyzed with Mass Hunter Quantitative Analysis software (B.06.00).

ASV Detection

Barcodes from raw SSU rRNA amplicon sequences were trimmed and sequences were dereplicated using cutadapt v 3.0 (M. Martin, 2011). Demultiplexed reads were quality filtered, forward and reverse reads were merged (~372 bp) and amplicon sequence variants (ASVs) were inferred as a part of the *dada2* pipeline (Callahan et al., 2017). Taxonomy was assigned using the SILVA database v32 (Quast et al., 2013; Yilmaz et al., 2014). Contaminants were removed via the r package *decontam* v. 1.10 using the frequency- and prevalence-based filtering with a

threshold of 0.5 (Davis et al., 2017). The final, experimental dataset consisted of ~1.2 million reads resulting in 484 ASVs with an average of 12368 reads per samples (min = 2860, max = 34441).

Tree Building and Community Analysis

Sequences were aligned using the r package *DECIPHER* v 2.0 (Wright, 2015, 2016) and the r package *phangorn* v2.5.5 (Schliep, 2011) was used to first construct a neighbor-joining tree to use as the starting point for a GTR+G+I maximum likelihood tree. Community analysis was done using environments created from the r package *phyloseq* v1.26.1 (McMurdie & Holmes, 2013).

Shannon's Index was used to describe alpha diversity. Bray-Curtis distances were used to calculate beta diversity and ordinated using Principal Coordinate Analysis (PCoA) and drawn with 95% confidence ellipses. Statistical comparisons between lineages and transfers were done using PERMANOVA with the adonis function in the *vegan* package (~ Lineage*Transfer) (Oksanen et al., 2019). To fully understand the significance described in the PERMANOVA, a PERMDISP with the betadisper function in *vegan* was used to describe any within-sample variance (i.e., across replicates) that could explain any significant differences detected in the PERMANOVA (Oksanen et al., 2019). Hypothesis testing via ANOVA with permutations (n=999) was used with PERMDISP to determine any significant differences in variation within samples.

To characterize microbial community dynamics over time and between lineages, the r package *corncob* (B. D. Martin et al., 2020) was used to characterize significant changes in abundance and variability of dominant taxa using count regression using the beta binomial distribution and corrected for false discovery rates. To investigate changes in relative abundance for the significant and dominate genera, a model was constructed using the formula ~Lineage + Transfer + Lineage:Transfer and compared to a null (~1) using a parametric bootstrapped (n=1000) Likelihood Ratio Test to determine if transfer and lineage affected taxon abundance.

To characterize the resistance of cultivated communities, colony forming units (CFUs) were calculated using the following formula (CFU/mL = (number of colonies/ (volume plated (mL)) * (dilution)). A nonlinear regression was used to describe the changes in CFUs over time via PRISM v9 (GraphPad Software, San Diego, California USA). Next, antibiotic plate CFUs were normalized to the untreated plate CFUs to generate a ratio where the untreated plates equals 1. A Wilcoxon test was used to compare the mean ratio of each antibiotic plate to a hypothetical median of 1 that represents the non-treated plate in PRISM v9 (GraphPad Software, San Diego, California USA).

Molecular Networking and Spectral Library Search in GNPS

Raw data obtained from the LCMS/MS instrument was converted to .mzXML files using msConvert within ProteoWizard v3 (Chambers et al., 2012). The mass spectrometry data were first processed with MZmine2 v2.51 (Pluskal et al., 2010). Peaks were identified (MS1 noise cutoff: 1.0E2, MS2 noise cutoff: 1.0E1). Chromatograms were built under the following criteria

using the ADAP chromatogram builder (Myers et al., 2017): MS1 scans; Minimum Group Number: 5; Group Intensity: 2.0E2; Minimum High Intensity: 1.0E3; *m/z* tolerance: 0.02 *m/z* or 0.0 ppm. Deconvolution was achieved through the following parameters via the local minimum algorithm: Chromatographic Threshold: 5.0%; search minimum in RT range: 0.4 min; Minimum Relative Height: 5.0%; Minimum Absolute Height: 2.0E3; Minimum Ratio Peak Top/Edge: 1; Duration: 0.01-3.50 min; Algorithm: Median; *m/z* MS2: 0.01Da; RT Range: 0.2 min. Isotopes were grouped under the following criteria: *m/z* total: 0.02; RT total: 0.10 min; Maximum charge: 3; Representative Isotope: Max Intensity. The feature table was constructed with using the following alignment parameters: *m/z* tolerance: 0.02 *m/z* or 200 ppm; Weight for *m/z*: 75; RT tolerance: 0.1 min; Weight for RT: 25 and filtered using the following criteria: Min peaks in per row : 3; Keep only peaks with MS2: yes. Poor peaks were removed from the peak list (long tails or plateaus) and any feature found in a blank (methanol, glassware, or isopropanol) was removed from the feature table. Results were exported to GNPS (Wang et al., 2017), for feature based molecular networking (FBMN) analysis (Nothias et al., 2020).

In GNPS the data were analyzed using FBMN workflow (Nothias et al., 2020). The data were filtered by removing all MS/MS fragment ions within +/- 17 Da of the precursor *m/z*. MS/MS spectra were filtered by choosing the top 6 fragment ions in the +/- 50 Da window. The precursor ion mass tolerance was set to 0.02 Da and the MS/MS fragment ion tolerance to 0.02 Da. A molecular network was created with edges filtered via cosine score > 0.7 and > 4 matched peaks. Also, edges between nodes were kept only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100. The analogue search mode searched against MS/MS spectra with a maximum difference of

100 Da in the precursor ion value. The library spectra were filtered as the input data was filtered, as discussed above, cosine score above 0.7 and at least 4 matched peaks.

Data Deposition and Job Accessibility

Sequence data has been deposited at NCBI's Sequence Read Archive (SRA) database under accession number PRJNA737136.

The mass spectrometry data were deposited on MassIVE (MSV000086507). The molecular networking job can be publicly accessed at https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=7910db941bf74d4592f982044399dc49.

Metabolomics Data Analysis and Hypothesis Testing

Principal Coordinate Analysis using Bray-Curtis distances calculated from total ion current (TIC)-normalized feature table was used to visualize the composition of metabolites using the r package *phyloseq* v1.26.1 (McMurdie and Holmes 2013). The molecular networks were visualized using Cytoscape software v3.8.1 (Shannon et al., 2003). To test for differences, a PERMANOVA and PERMDISP was used (as described above) to characterize differences between metabolite composition among samples. A chi-square test was used to determine if the proportions of shared and unique metabolites between lineages was significant. Features were normalized to total ion current (TIC). Richness (observed) and Shannon's Index was also used to describe feature alpha diversity. NPclassifier (https://npclassifier.ucsd.edu/) was used to group analog annotated features in molecular superclass groups with relevant natural product nomenclature (Kim et al., 2020) and the abundance of all features in each superclass were

summed to generate the total abundance. A heatmap was used visualized difference in total abundance in PRISM v9 (GraphPad Software, San Diego, California USA).

Microbiome Metabolome Comparisons

A Procrustes rotational analysis using r packages *ade4* (Dray & Dufour, 2007) and *vegan* (Oksanen et al., 2019) was used to determine if the metabolome and the microbiome beta diversity changed similarly. To compare ASV and metabolite co-occurrences, mmvec (Bolyen et al., 2019; Morton et al., 2019) via qiime2 plug-in (Bolyen et al., 2019) was used to generate co-occurrence probabilities via neural networks following default parameters. The mmvec algorithm produces conditional probabilities (not correlations), because of this, negative values represent lower probability rather than an anti-correlation.

Results

Disturbance Resulted in Diverging Community Structures

Three replicate cultivation linages resulted in three distinct community assembly patterns, where differences in community structure were driven by both transfer and lineage. Early transfers from all lineages initially clustered together, indicating similarity in the starting inoculum, but linaeges A and C eventually became more different from each other with later transfers (Figure 1A). PERMANOVA indicated that much of the variation can be explained by transfer ($p < 0.001$, $r^2 = 0.42$), lineage ($p< 0.001$, $r^2 = 0.27$), or interactions between lineage and transfers (lineage*transfers, $p < 0.001$, $r^2 = 0.21$). Considering the effect dispersion could have on differences observed between communities, a PERMDISP was used to see if any significant effect was related to within-group dispersion, rather than between group differences. There was no significant dispersion within-groups when considering transfers ($p = 0.656$) indicating that changes in community structure among samples followed predictable trends in response to disturbance. In contrast, the dispersion within lineage groups did significantly differ ($p= 0.028$) and was not predictable, which is highlighted by the different compositional changes that occurred in each lineage over time (Figure 1B). To address the effect lineage had on beta diversity variability, each lineage was treated as a separate community and beta-diversity was recalculated using Bray-Curtis distance (Figure 1C). Plotting the primary axis of the Bray-Curtis distance matrix against alpha diversity, we observed that beta diversity for lineage A converged rapidly with later transfers remaining very similar. In contrast, the beta diversity of lineage B changed along the gradient that follows transfers, highlighting the gradual changes observed in Figure 1B. Lastly, the changes in beta diversity for lineage C also followed the transfer gradient

but mid-late transfers began to cluster together, as observed in the Figure 1B. In all cases (lineages A, B, and C), alpha diversity decreased with increasing transfers (see Figure 1D).

Disturbance Resulted in Different Community Compositions

Each lineage followed a different community assembly pattern in response to continued disturbance (Figure 1B) that altered alpha diversity of each community over time (Figure 1C, Figure 1D). Overall, *Chryseobacterium* sp. was most dominant in early communities, while *Klebsiella* sp. became more dominant in later transfers. However, there were differences in how relative abundance changed over time. This resulted in two different outcomes, one where *Klebsiella* sp. eventually became dominant by T10 (lineages A and B), and one were *Klebsiella* sp. and *Chryseobacterium* sp. coexisted (lineage C) (Figure 1B). A differential analysis was used to determine which genera increased in relative abundance significantly with transfers by controlling for the effects both transfers and lineage could have on the abundance and dispersion. Multiple genera were associated with increases over transfers (Supplemental Figure 2). Here, we present five taxa that are members of the dominant genera (Figure 1B) and were significantly associated with changes in abundance over time and between transfers (Likelihood Ratio Test: $p < 0.05$): *Acinetobacter* sp., *Pseudomonas* sp., *Chryseobacterium* sp., an Enterobacteriaceae species, and *Klebsiella* sp. (Figure 2).

Specifically, the relative abundance of *Klebsiella* sp. increased rapidly in lineage A remaining stable in higher abundance, gradually in lineage B, and slightly in lineage C eventually becoming stable. In contrast, the relative abundance of *Chryseobacterium* sp. and the Enterobacteriaceae species decreased rapidly in lineage A, and decreased gradually in lineage B. Further, *Chryseobacterium* sp. represented a constant ~45-50% of the composition in lineage C

(similar to *Klebsiella* sp.). Lastly, the relative abundance of *Pseudomonas* sp. increased in lineage A only and decreased in lineages B and C; and the relative abundance of *Acinetobacter* sp. increased only in lineage C (Figure 2).

Populations That Could Successfully Colonize Were Also Resistant to Antimicrobials

One strategy that can make microorganisms more competitive is the production of antimicrobials and resistance to antimicrobials. To characterize the resistance profile of the resulting microbial communities, we pooled a subsample from lineages A, B and C, and plated it on agar plates containing different antibiotics. Despite not selecting for antimicrobial resistance during cultivation, we observed that the frequency of resistant organisms increased with successive transfers. First, a nonlinear regression showed that CFUs for each condition increased over time by comparing the slope to a null slope via a Likelihood Ratio Test ($p < 0.0001$) (Figure 3). As expected, we observed that the number of CFUs from plates containing antibiotics were significantly less than CFUs from plates with no antibiotics (Wilcoxon: $W = -66$, $p < 0.001$) (Figure 3), but the proportion of resistance CFUs increased with each transfer in all treatments except for ciprofloxacin. Lastly, we observed the highest levels of resistance among cultivated communities were to ampicillin and tetracycline compared to ciprofloxacin. Overall, we observed that repeated disturbance selected for antimicrobial resistant organisms.

Differences in the Metabolome Were Driven by Both Transfer and Lineage

Similar to the microbiome, the structure of the metabolome was different between lineages and with each transfer (Figure 4A), though richness was similar between lineages (Supplemental Figure 4). A Principal Coordinate Analysis showed that samples clustered with their own lineage

after transfer 0. Within those clusters, there was separation based on transfer (Figure 2A). PERMANOVA indicated that these compositional differences were driven by lineage ($p = 0.001$, $r^2 = 0.18$), transfer ($p = 0.001$, $r^2 = 0.45$), and interactions between the two ($p = 0.001$, $r^2 = 0.33$). Despite significant differences in the compositional structure between transfers, samples did exhibit significant with-in group dispersion ($p = 0.001$), which could explain observed significance in the PERMANOVA, and indicates that the composition of the metabolome changed differently over time between lineages. Samples did not have significant dispersion when considering lineage ($p= 0.184$). This is not surprising as many more metabolites were shared between lineages, compared to metabolites unique to each lineage (chi-sq = 6.7, $p=0.04$) (Figure 4B), although the proportion of shared metabolites decreased over time (Supplemental Figure 5). The difference in metabolome structure and the decrease in shared metabolites could indicate that despite having many shared taxa, metabolic strategies and responses can differ.

After features were categorized into relevant natural product superclasses, some groups were more ubiquitous across lineage and transfers than others (Figure 4C). Pseudoalkaloids were the most abundant annotated superclass and observed in each lineage across transfers. Some superclasses were only abundant in later transfers for each lineage, like macrolides (a class of antibiotics). Lastly, some superclasses were more abundant in certain lineages, like monoterpenoids in lineage C.

Metabolomics Can Help Describe Potential Mechanisms of Microbial Community Response to Disturbance

We expected the structure of the microbial community to change in response to disturbance as members sought to re-assemble, but we also wanted to know if the metabolome changed in a

similar way. A Procrustes analysis showed that the structure of the metabolome was significantly correlated to the structure of the microbiome ($r^2 = 0.45$, $p = 0.001$) (Figure 5). Further, co-occurrence analysis via mmvec characterized differential patterns of co-occurrence between ASVs and metabolites (Supplemental Figures 6-8).

For lineage A, the most abundant genus co-occured with certain natural product super classes like nucleosides and pseudoalkaloids (log conditional probability > 0) (Supplemental Figure 6). For lineage B, members of Enterobacteriaceae, *Klebsiella* spp. and some *Chryseobacterium* spp. were associated with many natural product super classes, like nucleosides, cyclic polyketides, or monoterpenoids (log conditional probability > 0) (Supplemental Figure 7). Lastly, some *Chryseobacterium* spp. and *Klebsiella* spp. were observed to co-occur with monoterpenoids, nucleosides, pseudoalkaloids in lineage C (log conditional probability > 0) (Supplemental Figure 8). These co-occurrence patterns coupled with the significant Procrustes analysis (Figure 5) show that the microbial community structure and metabolome structure changed similarly, and those certain taxa can be linked to the presence of certain superclasses of natural products. If certain taxa are observed to significantly change in response to disturbance and the metabolome changes in a similar way, it follows that those organisms are driving the secondary metabolic responses observed in the community.

Discussion

We sought to explore changes in microbial and chemical diversity by combining approaches that used habitat structure to elicit the production of secondary metabolites, and perturbations to characterize community assembly patterns. Overall, we observed that not all outcomes of community composition were predictable in response to disturbance; we also observed that changes in the microbial community structure were linked to changes in the metabolome. Specifically, we were able to (1) characterize three patterns of microbial community assembly, (2) identify the most successful taxa and characterize their antimicrobial resistance, and (3) describe changes in the metabolome of each community in response to continued disturbance. These data highlight the relationship between a microbial community and their chemical environment through the production of secondary metabolites as populations undergo successive selection events (i.e. disturbance).

Three patterns of community assembly emerged through successive transfers, highlighting that not all outcomes of community assembly were predictable. Through a combination of founder's effects and selection for taxa which can proliferate more successfully, we observed taxa with more generalist strategies to be more be more successful (i.e., win). As we generated founder's effects via repeated transfers, diversity decreased. These founder's effects constrained the populations that underwent selection resulting in three distinct microbial community assembly patterns. In two outcomes, *Klebsiella* sp. became the dominant taxa, potentially through competitive exclusion. First, *Klebsiella* sp. become abundant very quickly in lineage A, and more gradually in lineage B. A third outcome resulted in the coexistence of *Klebsiella* sp. and *Chryseobacterium* sp. However, determining the mechanisms behind the community

assembly patterns is notoriously complex, and the different ways populations could respond to disturbance introduces variability in outcomes of community assembly. In the case of lineage C, the structured environment's ability to contain otherwise rapid growth while also partitioning resources and potential antimicrobial secondary metabolites may have allowed for coexistence and variability which we observed the in the community assembly patterns (Geyrhofer & Brenner, 2020; Amarasekare, 2003). Lastly, by taking a subsample of the community, mixing it, and allowing it to grow in a structured environment, we allowed populations to build upon the reproductive success from previous transfers (e.g. community-level historical contingency). Though this was not explicitly tested, we anticipate that the increase in relative abundance is also related to the size of the inoculum and its ability to capture the total diversity of the community (see Supplemental Figure 3 for sequence count abundance), similar to a species-area relationship (Green & Ostling, 2003). In this case, it is challenging to separate increases in abundance from a population's resiliency in response to disturbance and the effect of dispersal (i.e. founder's effects) on the transfers. Overall, specific dynamics of community assembly were unpredictable, but the general trend of the same organisms being successful colonizers were shared among lineages.

We sought to explore the potential mechanism through which an organism is a successful colonizer, particularly in response to repeated disturbance. Our data showed the ability for taxa in smaller populations in wastewater to quickly become abundant and dominant during cultivation. This suggests two things (1) that both *Klebsiella* sp. and *Chryseobacterium* sp. can quickly establish growth after each transfer, and (2) are more resilient to disturbance, in this case intermittent mixing. In each case, these attributes make them good pioneer taxa. One mechanism

to explain the high relative abundance of *Klebsiella* sp. is competitive exclusion. Competitive exclusion describes the idea that two organisms cannot overlap in niche space unless something relieves that selection pressure by partitioning resources (Hardin, 1960). This can be achieved over time through natural selection or by phenotypic plasticity of certain populations (Rainey & Travisano, 1998). In the case of lineages A and B, we observed *Klebsiella* sp. to significantly increase over time and become abundant by the final transfers, making them consistently successful during community assembly in structured habitats. This is not surprising as they are adept nosocomial, antimicrobial resistant, opportunistic pathogens (Podschun & Ullmann, 1998; Wyres & Holt, 2018). Hallmarks of many opportunistic pathogens are their generalist tendencies and their phenotypic plasticity (Brown et al., 2012), and organisms with these capabilities are often selected for in response to disturbances (Atlas et al., 1991). We suggest that competitive exclusion driven by potential phenotypic plasticity of *Klebsiella* sp. and its antimicrobial resistance allowed for its dominant relative abundance (van Dorp et al., 2019; Paczosa & Mecsas, 2016). As such, *Klebsiella* sp. is well suited to outcompete others for resources, particularly if those mechanisms (e.g., growth rate) can adapt quickly, compared to costly and timely production of antimicrobial compounds.

Second, we observed the frequency of resistant organisms to increase after each transfer, where disturbance selected for antimicrobial resistant organisms. The mechanisms of microbial warfare include both weapons and armor. Weapons could be bactericidal compounds and antagonist behavior like increased nutrient uptake or acidification of the medium. Alternatively, armor would be mechanisms of antimicrobial resistance (e.g. beta-lactamases). Both offensive or defensive strategies describe forms of interference competition or exploitative competition, respectively (Ghoul & Mitri, 2016). As a community undergoes repeated disturbances, those

organisms that produce an advantage, perhaps through competitive interactions, will be the most successful. In this case, both *Klebsiella* spp. and *Chryseobacterium* spp. are known to be highly resistant to antibiotics (Hayward et al., 2020; Verner-Jeffreys et al., 2017: Podschun & Ullmann, 1998) and resistance to antimicrobial compounds could represent a fitness advantage as populations seek to assemble post-disturbance (i.e. transfer).

Interestingly, the increase in resistance was observed only for ampicillin and tetracycline, but not ciprofloxacin. Both ampicillin and tetracycline are derived from microbial sources (Campoli-richards & Brogden, 1987; Chopra & Roberts, 2001; respectively), while ciprofloxacin is a synthetic drug (LeBel, 1988). One explanation for this could be that resistance in populations has been selected for through interactions mediated by the natural analogs of these compounds. Further, wastewater is known to harbor microbial communities with extensive antimicrobial resistance (Bouki et al., 2013) as well as populations of antimicrobial producers (Bensultana et al., 2010), suggesting that the populations that were able to proliferate in our cultivated mesocosms, could have already undergone selection for resistance which would provide an advantage if a competing organism were to use these compounds as a combative strategy. We suggest that since ciprofloxacin would not be produced as a competitive advantage, it would not apply an equal selective force on a population compared to ampicillin or tetracycline. The mechanisms of resistance to each of these antimicrobials can also explain the prevalence of resistance. The prevalence of resistance to ampicillin was not surprising. Our cultivation conditions selected for members of family Enterobacteriaceae which express beta-lactamase genes, making them resistance to beta-lactam drugs, like ampicillin. Not only are these members of this family known to express these genes, but the genes themselves have been shown to disseminate through horizontal gene transfer and be expressed in many, diverse populations

(reviewed in von Wintersdorff et al., 2016). Resistance to tetracycline differs in mechanisms of resistance (e.g. efflux proteins or ribosomal protection), but is mobilized on genetic elements (e.g. plasmids), similar to beta-lactamases (Grossman, 2016; Roberts, 1996). Resistance to tetracycline is also observed in similar organisms resistant to beta-lactam drugs (Grossman, 2016) and wastewater communities (Adesoji et al., 2015; Auerbach et al., 2007). Considering the widespread use of antibiotics (Ventola, 2015) and the resistance observed in wastewater communities, the selection for resistance observed here could play a role in fitness advantage during colonization and in response to disturbance.

Another mechanism for taxa to be successful colonizers is the production of secondary metabolites that give a fitness advantage to the producer. We observed some metabolites to co-occur with certain taxa. For example, *Chryseobacterium* sp. are known producers of terpenoids (Jia et al., 2019; Rudolf et al., 2021), molecules that have antibacterial activity (e.g. inhibition of biofilm formation) (Mahizan et al., 2019). We observed that communities with monoterpenoids, were also communities with high abundances of *Chryseobacterium* sp (Figure 4C). We can hypothesize that the production of terpenoids may decrease the growth of *Klebsiella* sp., a known biofilm producer (Dumaru et al., 2019; Stewart et al., 1997), and result in coexistence between the two organisms in a structured environment. This hypothesis could be tested using isolates and bioactivity assays, something this study did not set out to do, but gives rationale to do so. Also, *Klebsiella* sp. and the natural product superclass of metabolites, nucleosides, were observed in each lineage and observed to co-occur (Supplemental Figures 6-8). Nucleosides are proposed to have antimicrobial activity and the specific metabolite identified here, methylthioadenosine (Supplemental Figure 9), has been observed in *Klebsiella* sp. methionine

salvage pathways (Sekowska et al., 2019; Parveen & Cornell, 2012; Riscoe et al., 1988). In both examples, the secondary metabolites could indicate mechanisms that increased fitness.

Disturbance or perturbation is a common phenomenon encountered by microbial communities and can be a maintainer of ecosystem diversity (e.g. intermediate disturbance hypothesis) (Buckling et al., 2000; Santillan et al., 2019; Connell, 1978). In this case, we exposed communities to a fixed rate of disturbance using transfers and found that diversity initially decreased, though this decrease will also be initially confounded with the added selection pressure when going from a wastewater system to a cultivated mesocosm, a disturbance itself. Here, we observed different, distinct patterns of compositional changes over time, where changes in the microbiome and metabolome were correlated (Figure 5). This indicates that while each lineage began to stabilize, either through competitive exclusion (lineage A and B) or co-existence (lineage C), the metabolic response to disturbance is linked to the changes in the diversity of the microbiome. The patterns in both the microbial data and the metabolome data highlight the complexity of community assembly and metabolic responses, both to changes in the environment and the interactions among the members of a community. We observed certain taxa to 'win' and can hypothesize about the potential mechanisms driving community dynamics, showing that sequencing data fails to capture the complexity of a system. As microbial ecology merges with chemical ecology, using tools that capture total microbial diversity and chemical diversity describe the untold stories from which hypotheses can be generated and tested. This study adds to the current efforts to describe the chemical response to microbial community assembly.

Figures

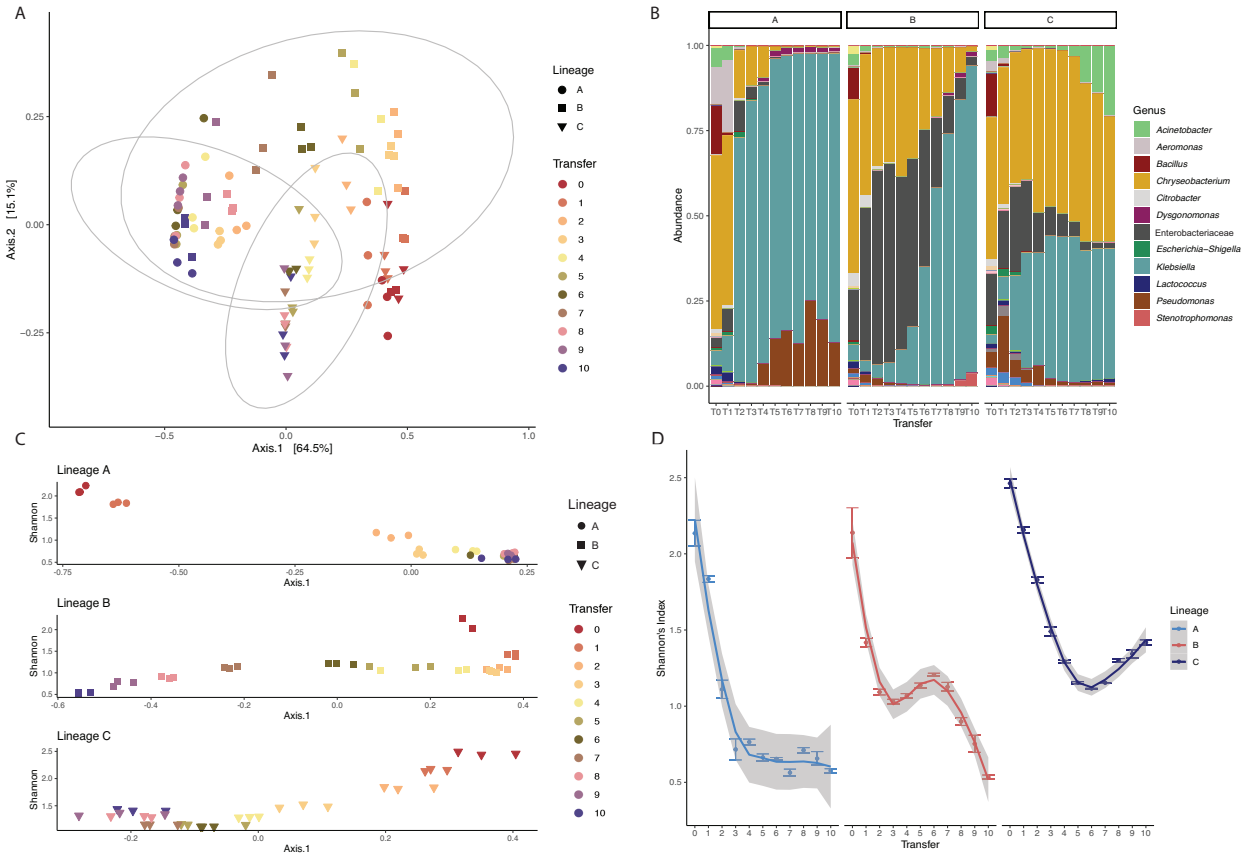


Figure 1: Microbial community dynamics. (A) Principal Coordinate Analysis using Bray-Curtis distance. Shape and 95% confidence ellipses refer to lineage, color refers to transfer. (B) Relative abundance of genera over time for each lineage. (C) Relationship between alpha and beta diversity for each lineage. Color refers to lineage. (D) Alpha diversity of the microbial community measured in Shannon's index (mean with standard deviation). Data fit with a loess regression line and grey area indicates 0.95 confidence intervals.

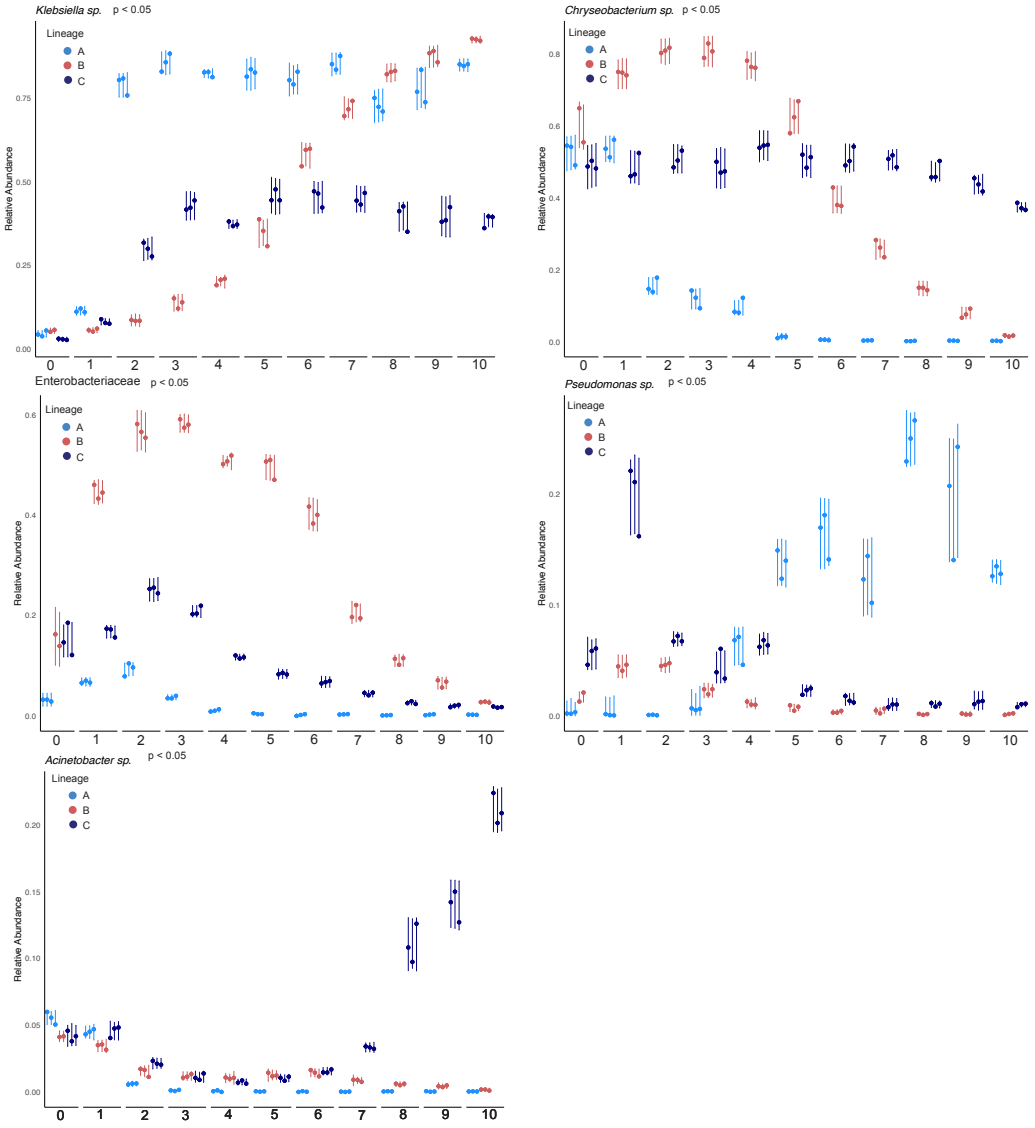


Figure 2: Differential abundance of taxa significantly associated with changes over time (see Supplemental Figure 1). Color corresponds to lineage. 95% prediction intervals are represented with bars.

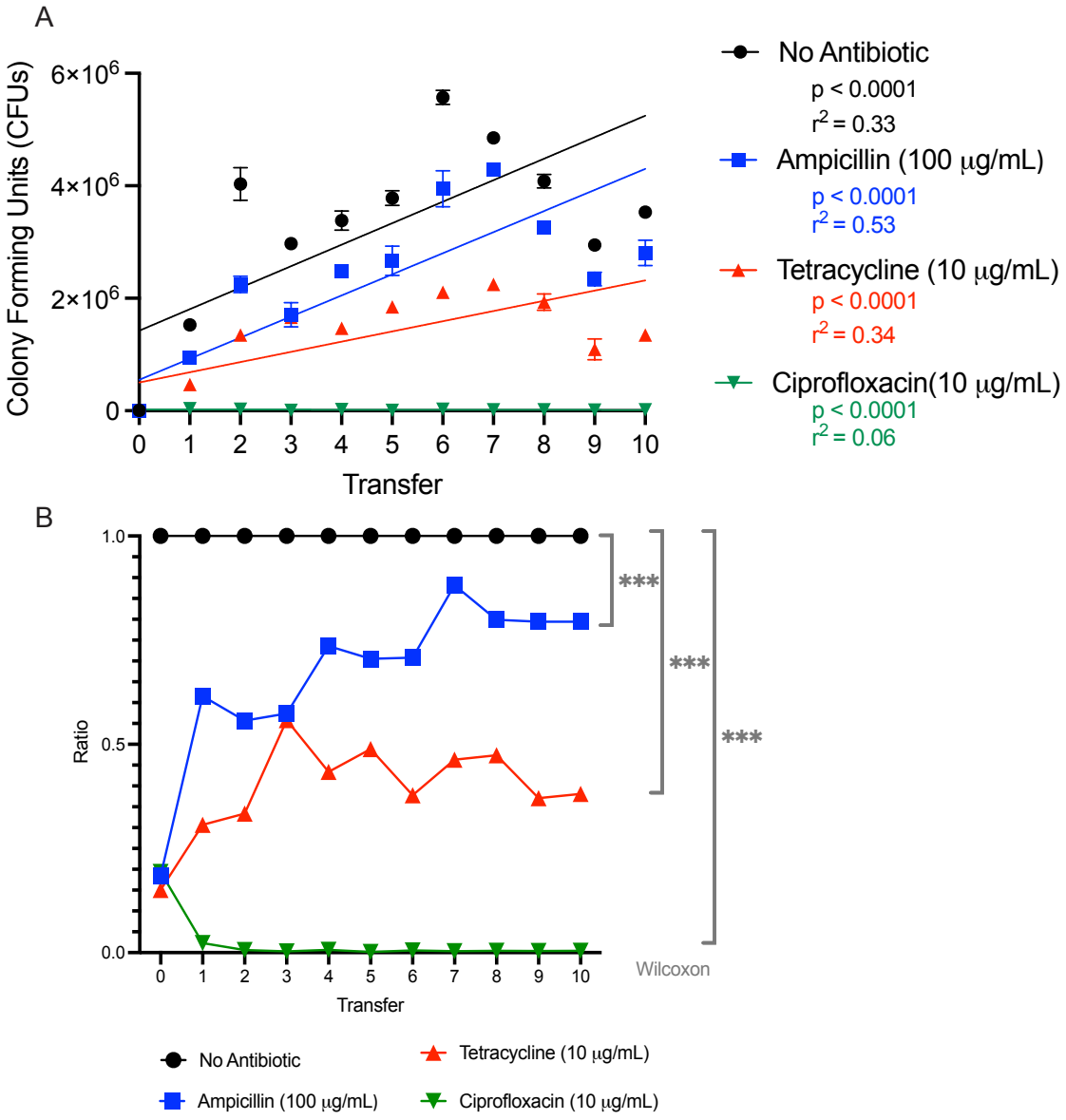


Figure 3: Changes in colony forming units (CFU/mL) over time using (A) nonlinear regressions showing changes in CFUs over time and (B) ratio of nontreated:treated plates, (Wilcoxon: *** $p < 0.001$). Plates containing Ciprofloxacin had fewer CFUs than other treatments resulting in a flatter line on the y-axis.

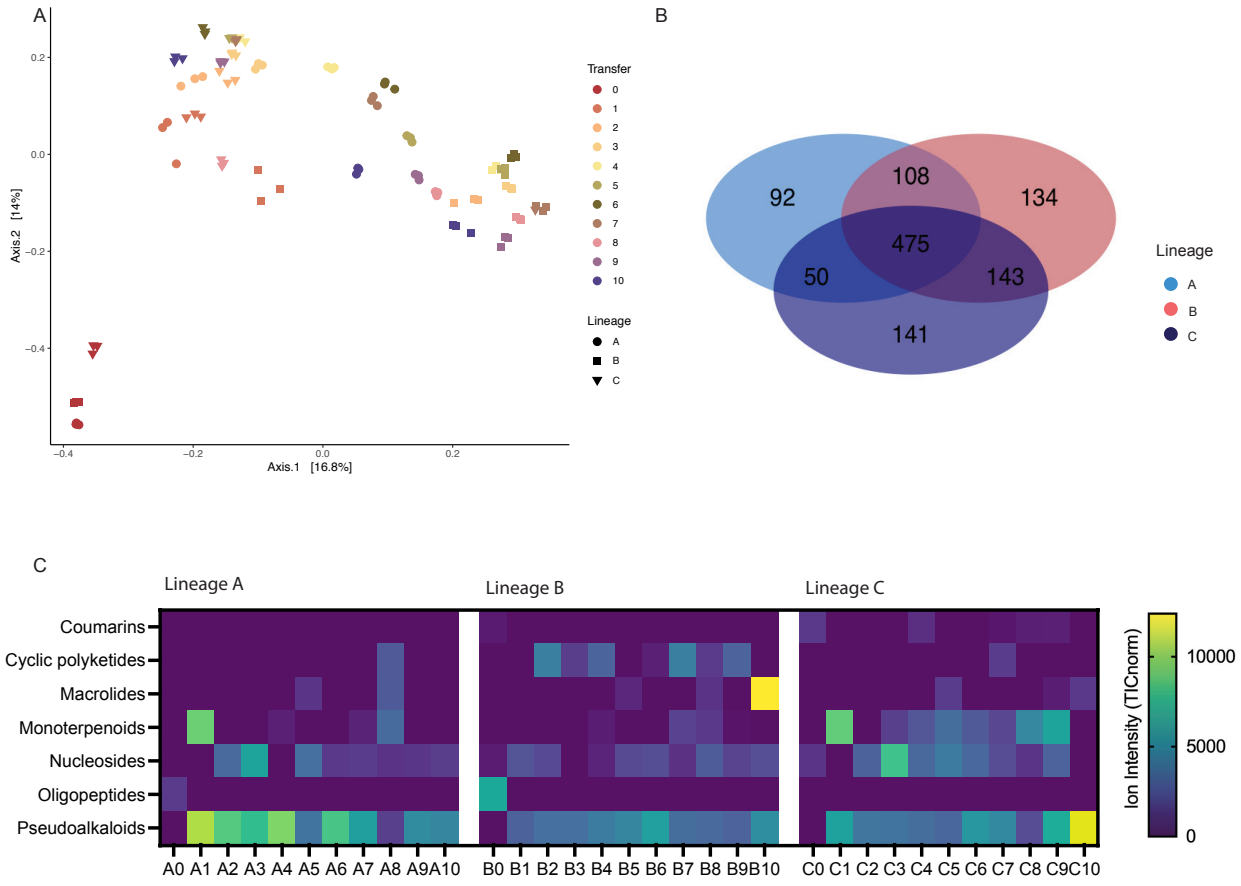


Figure 4: Metabolome dynamics. (A) Principal Coordinate Analysis of features using Bray-Curtis distance. Shape refers to lineage, color refers to transfer. (B) Features shared or unique among lineages (Venn Diagram). (C) Heatmap describing the abundance of different metabolite superclasses for each lineage. Values represent summed TIC normalized ion intensities for each superclass.

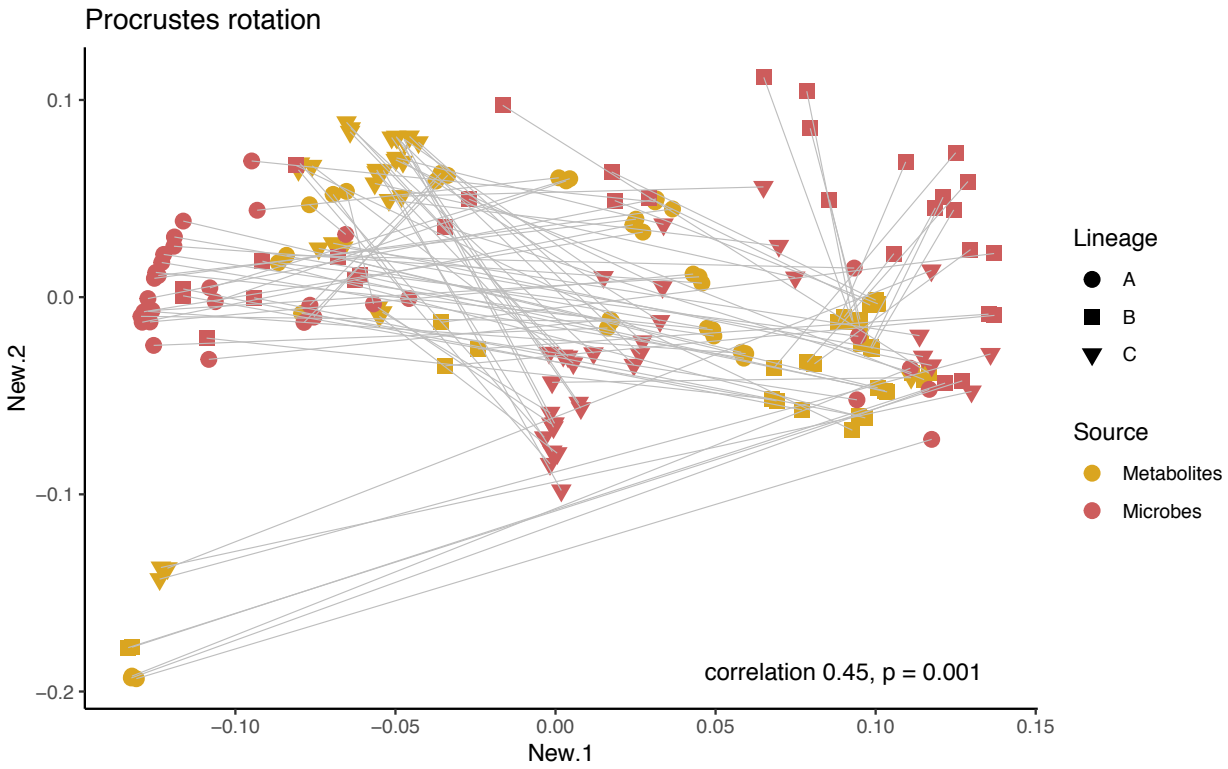
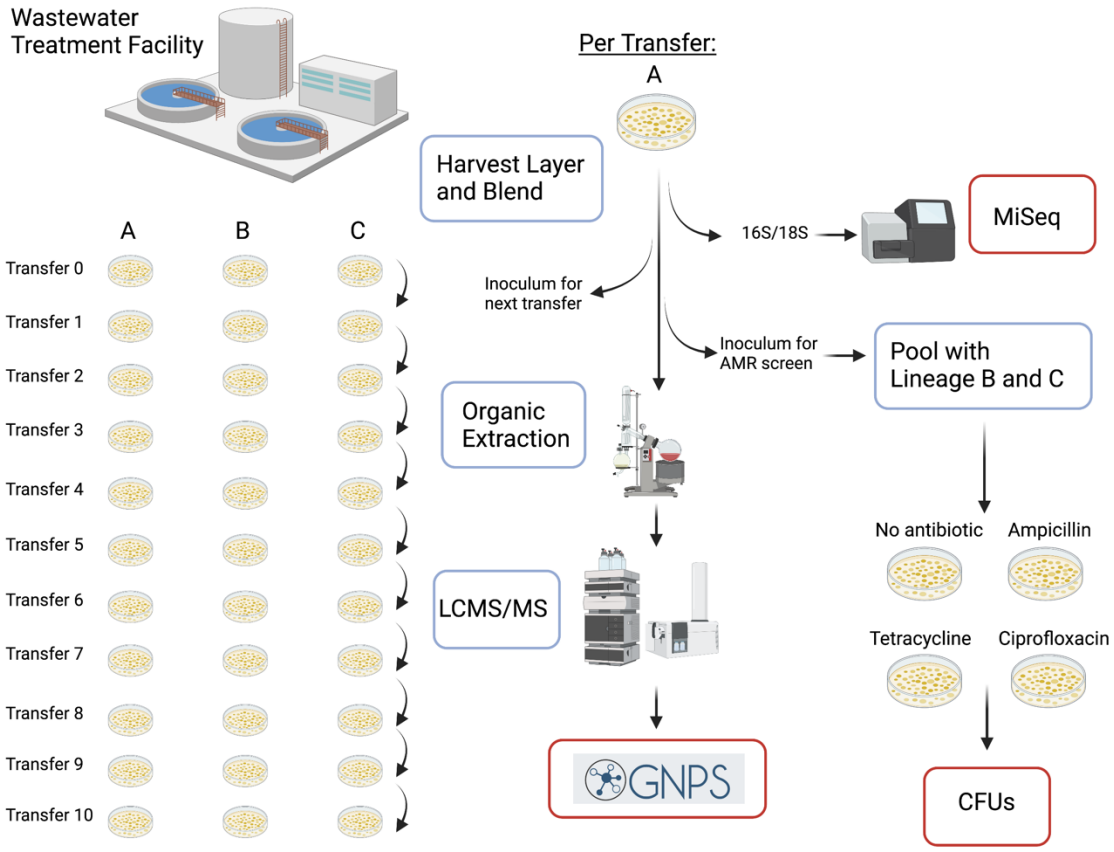
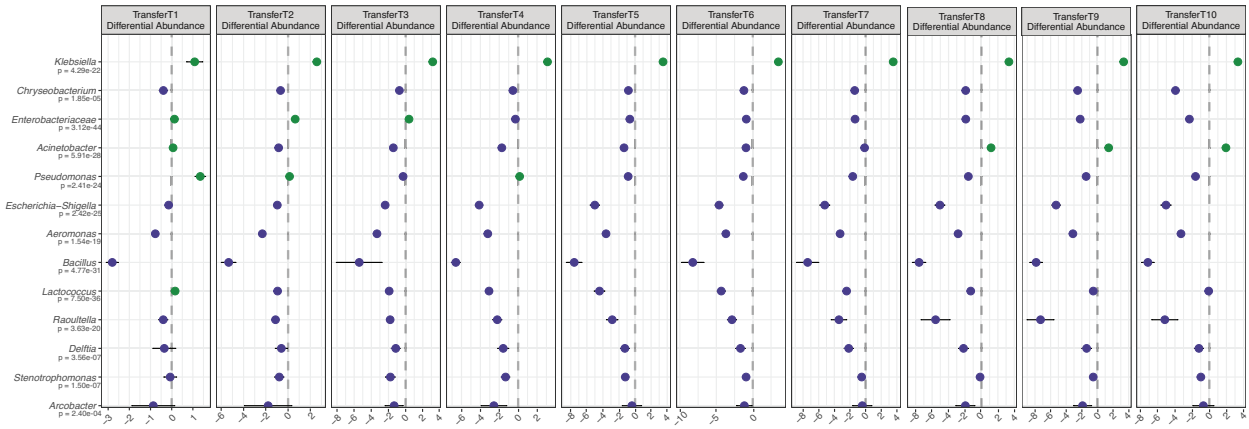


Figure 5. Procrustes rotational analysis comparing the changes in beta diversity between the microbiome (red) and the metabolome (gold). Shape refers to lineage. Corresponding microbiome-metabolome samples connected with grey line. Significant correlation: coefficient = 0.45, p=0.001.

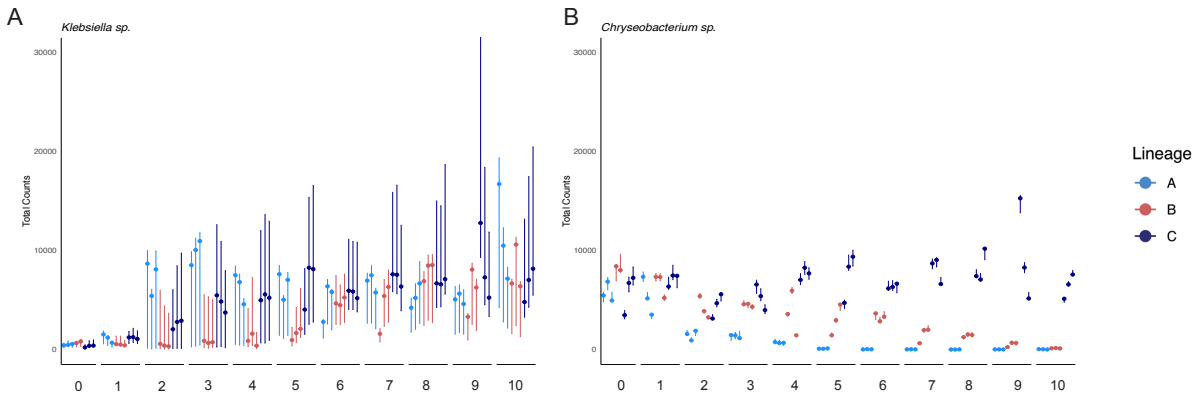
Supplemental Figures



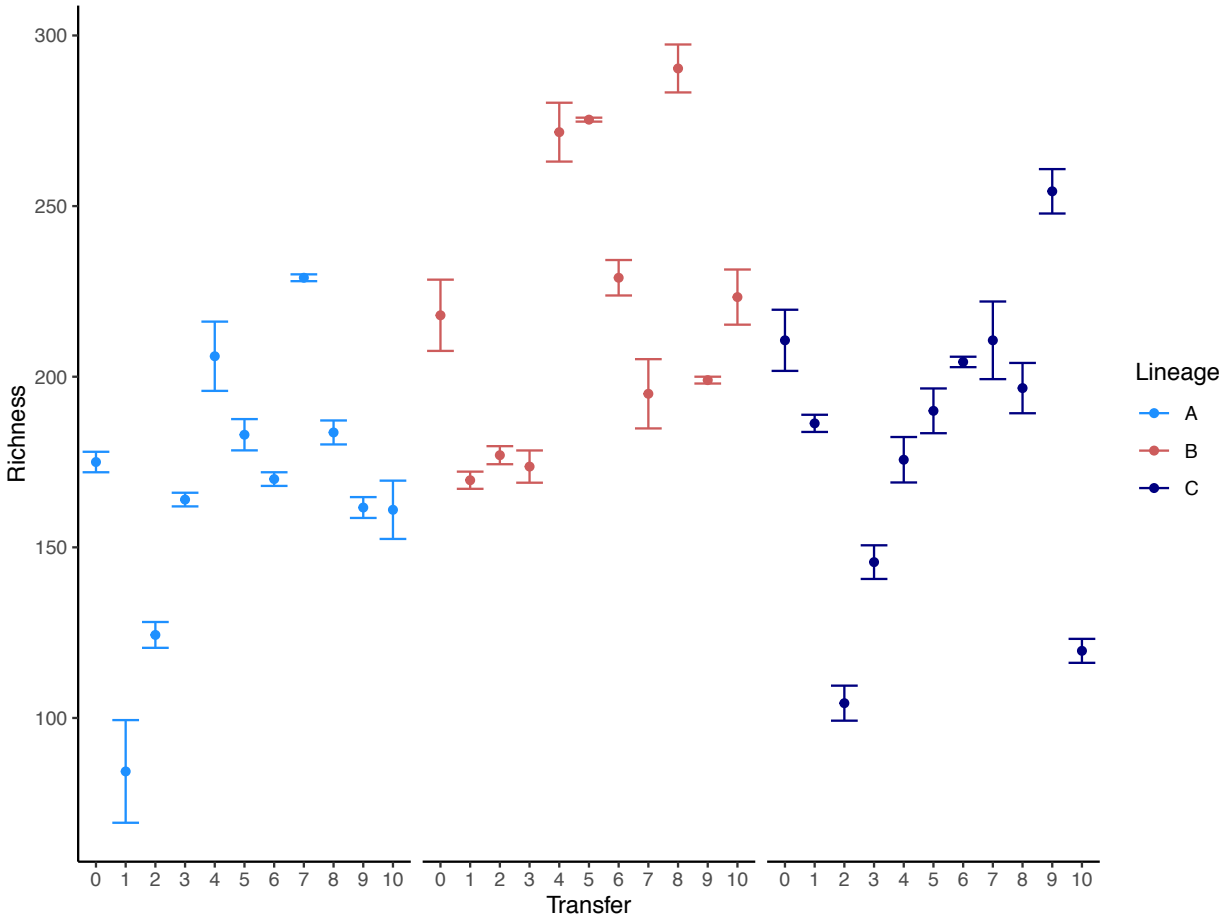
Supplemental Figure 1: Experimental design schematic showing transfer and example breakdown using one biological replicate (lineage A). Blue boxes are major steps, red boxes are types of data generated. Created with Biorender.com.



Supplemental Figure 2: ASVs significantly associated with changes in relative abundance between transfers. FDR-corrected p-values are presented for each ASV. Green indicates a positive or increase in relative abundance, while purple indicates a negative or decrease in relative abundance. Note, because this is a binomial regression, all values are relative to the initial transfer 0.

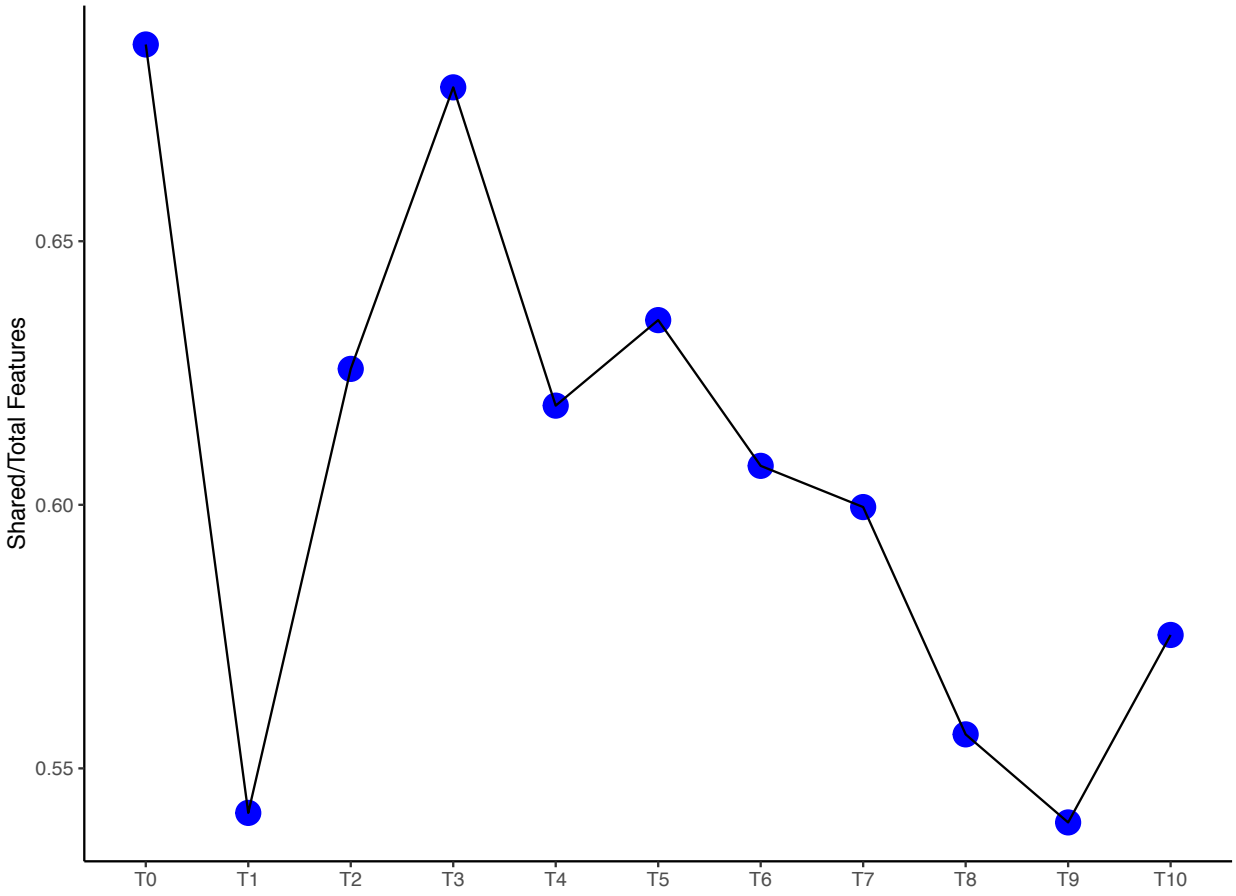


Supplemental Figure 3: Changes in sequence counts over time for (A) *Klebsiella* sp. and (B) *Chryseobacterium* sp. Color refers to lineage.

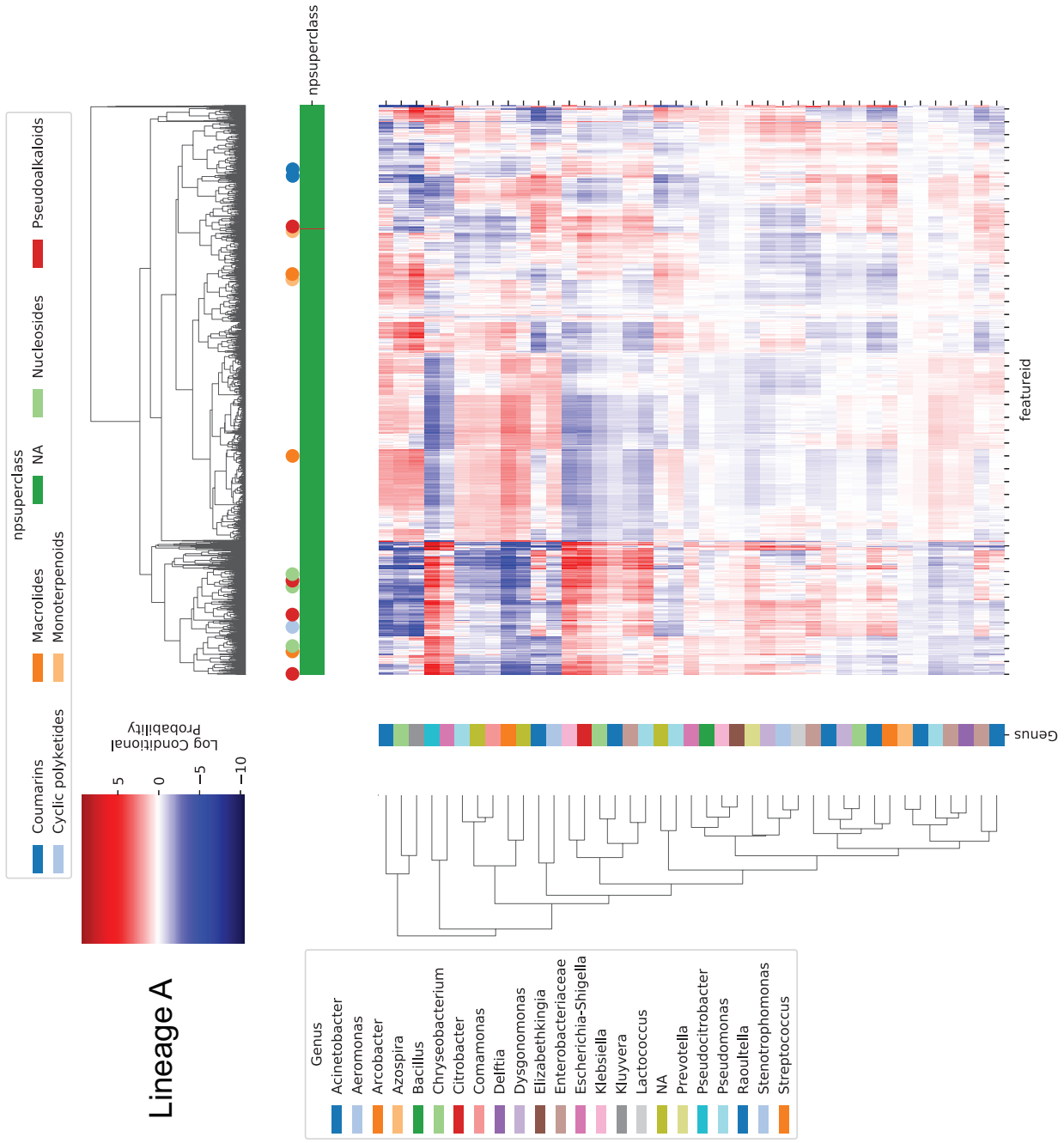


Supplemental Figure 4. Richness of metabolites measured in observed features over transfers.

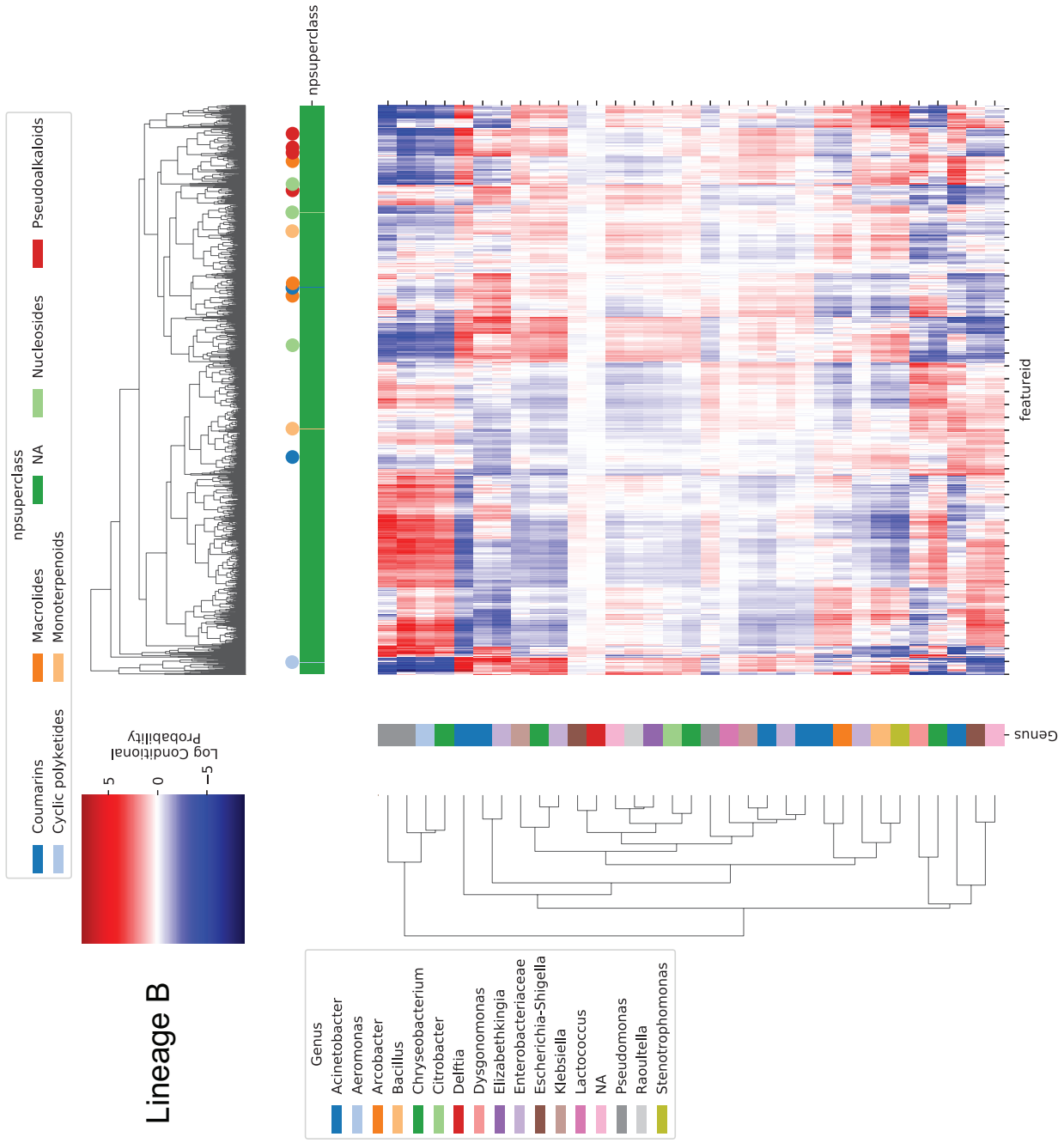
Points represent mean with standard deviation described with bars. Color refers to lineage.



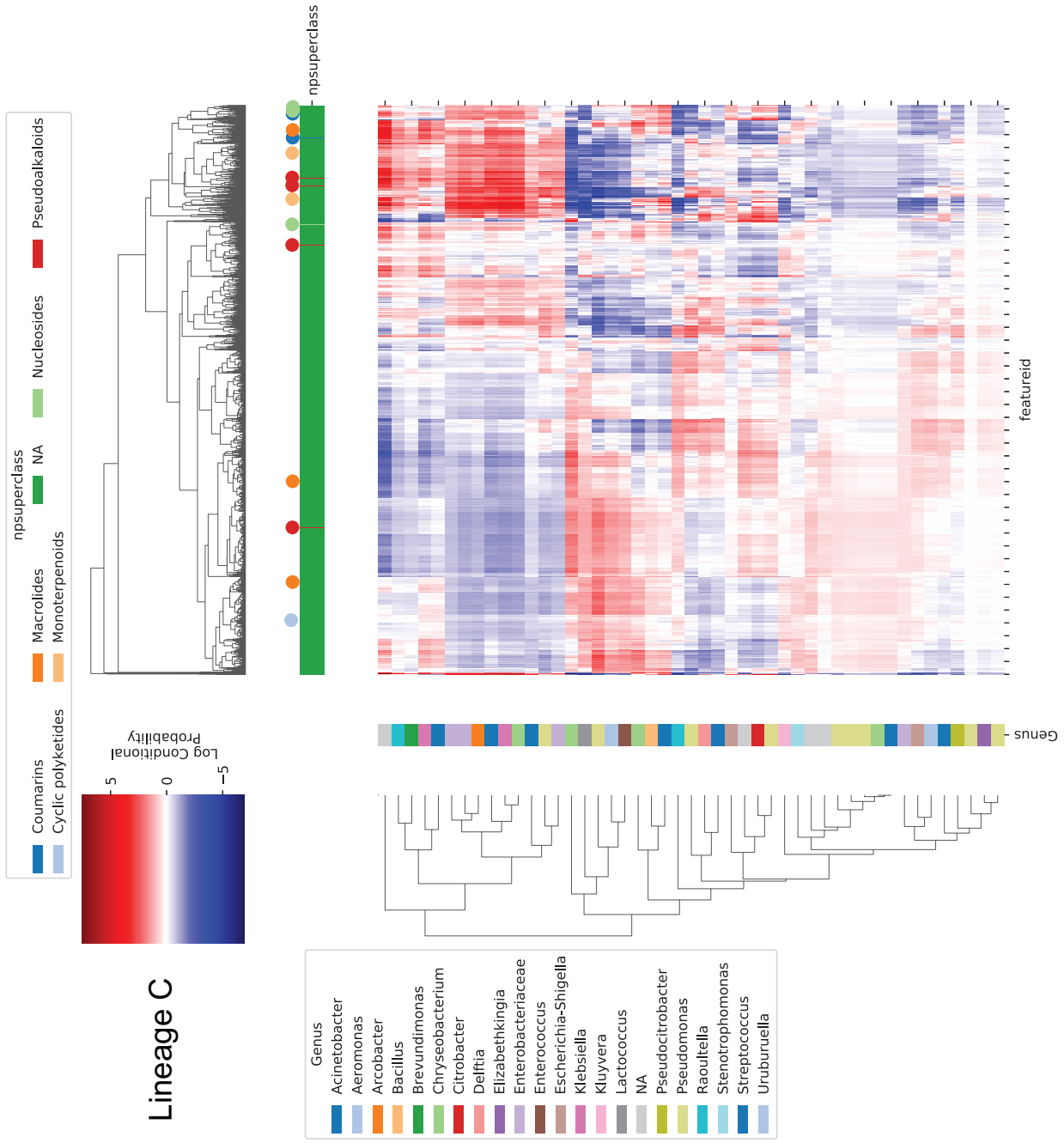
Supplemental Figure 5: Change in proportion of shared metabolites over time. Shared metabolites between each lineage (A:B:C) for each transfer divided by the total number of detected metabolites for that transfer.



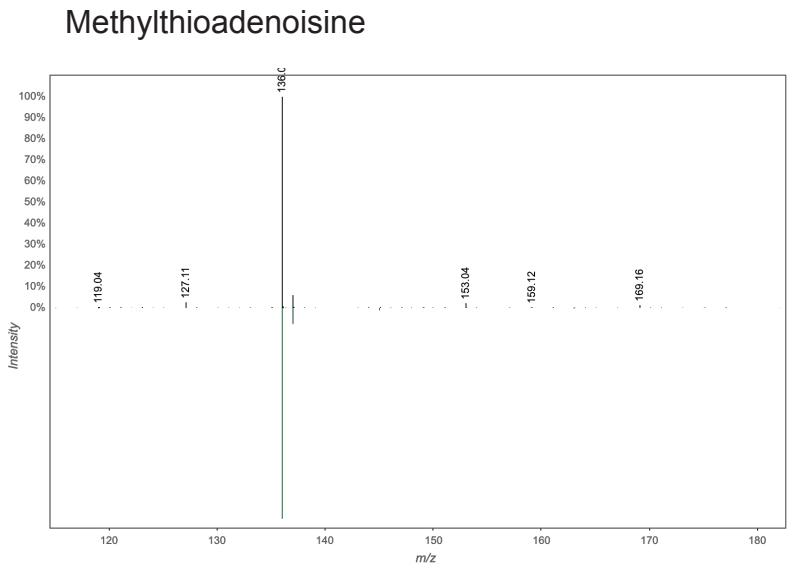
Supplemental Figure 6: mmvec heatmap of lineage A describing conditional probabilities of metabolite and microorganism co-occurrence. Dendrogram on the y axis clusters ASVs, colored by genera. Dendrogram on the x axis clusters metabolite features, colored by NPClassifier superclass; colored spheres are callouts. Red corresponds to high co-occurrence probabilities, while blue correspond to lower co-occurrence (not anti-correlation).



Supplemental Figure 7: mmvec heatmap of lineage B describing conditional probabilities of metabolite and microorganism co-occurrence. Dendrogram on the y axis clusters ASVs, colored by genera. Dendrogram on the x axis clusters metabolite features, colored by NPClassifier superclass; colored spheres are callouts. Red corresponds to high co-occurrence probabilities, while blue correspond to lower co-occurrence (not anti-correlation).



Supplemental Figure 8: mmvec heatmap of lineage C describing conditional probabilities of metabolite and microorganism co-occurrence. Dendrogram on the y axis clusters ASVs, colored by genera. Dendrogram on the x axis clusters metabolite features, colored by NPClassifier superclass; colored spheres are callouts. Red corresponds to high co-occurrence probabilities, while blue correspond to lower co-occurrence (not anti-correlation).



Supplemental Figure 9: Mirror plot of nucleoside feature identification. Black is experimental MS2 spectrum, green is analog library MS2 spectrum.

References

Adesoji, A. T., Ogunjobi, A. A., Olatoye, I. O., & Douglas, D. R. (2015). Prevalence of tetracycline resistance genes among multi-drug resistant bacteria from selected water distribution systems in southwestern Nigeria. *Annals of Clinical Microbiology and Antimicrobials*, *14*(1), 1–8. https://doi.org/10.1186/s12941-015-0093-1

Ali, J. G., Casteel, C. L., Mauck, K. E., & Trase, O. (2020). Chemical Ecology of Multitrophic Microbial Interactions: Plants, Insects, Microbes and the Metabolites that Connect Them. *Journal of Chemical Ecology*, *46*(8), 645–648. https://doi.org/10.1007/s10886-020-01209-y

Amarasekare, P. (2003). Competitive coexistence in spatially structured environments: A synthesis. *Ecology Letters*, *6*(12), 1109–1122. https://doi.org/10.1046/j.1461-0248.2003.00530.x

Andersen, R., Chapman, S. J., & Artz, R. R. E. (2013). Microbial communities in natural and disturbed peatlands: A review. *Soil Biology and Biochemistry*, *57*, 979–994. https://doi.org/10.1016/j.soilbio.2012.10.003

Antonopoulos, D. A., Huse, S. M., Morrison, H. G., Schmidt, T. M., Sogin, M. L., & Young, V. B. (2009). Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. *Infection and Immunity*, *77*(6), 2367–2375. https://doi.org/10.1128/IAI.01520-08

Antunes, L. C. M., Han, J., Ferreira, R. B. R., Lolić, P., Borchers, C. H., & Finlay, B. B. (2011). Effect of antibiotic treatment on the intestinal metabolome. *Antimicrobial Agents and Chemotherapy*, *55*(4), 1494–1503. https://doi.org/10.1128/AAC.01664-10

Atlas, R. M., Horowitz, A., & Krichevsky, M. (1991). Response of Microbial Populations to Environmental Disturbance. *Microbial Ecology*, *22*, 249–256.

Auerbach, E. A., Seyfried, E. E., & McMahon, K. D. (2007). Tetracycline resistance genes in activated sludge wastewater treatment plants. *Water Research*, *41*(5), 1143–1151. https://doi.org/10.1016/j.watres.2006.11.045

Bartley, L. E., Peck, M. L., Kim, S. R., Ebert, B., Manisseri, C., Chiniquy, D. M., Sykes, R., Gao, L., Rautengarten, C., Vega-Sánchez, M. E., Benke, P. I., Canlas, P. E., Cao, P., Brewer, S., Lin, F., Smith, W. L., Zhang, X., Keasling, J. D., Jentoff, R. E., … Ronald, P. C. (2013). Overexpression of a BAHD acyltransferase, OsAt10, alters rice cell wall hydroxycinnamic acid content and saccharification. *Plant Physiology*, *161*(4), 1615–1633. https://doi.org/10.1104/pp.112.208694

Belyea, L. R., & Lancaster, J. (1999). Assembly Rules within a Contingent Ecology. *OIKOS*, *86*(3), 402–416.

Bender, E. A., Case, T. J., & Gilpin, M. E. (1984). Perturbation Experiments in Community Ecology: Theory and Practice. *Ecology*, *65*(1).

Bensultana, A., Ouhdouch, Y., Hassani, L., Mezrioui, N. E., & Rafouk, L. (2010). Isolation and characterization of wastewater sand filter actinomycetes. *World Journal of Microbiology and Biotechnology*, *26*(3), 481–487. https://doi.org/10.1007/s11274-009-0194-0

Bissett, A., Brown, M. v., Siciliano, S. D., & Thrall, P. H. (2013). Microbial community responses to anthropogenically induced environmental change: Towards a systems approach. *Ecology Letters*, *16*(SUPPL.1), 128–139. https://doi.org/10.1111/ele.12109

Blaustein, L., & Margalit, J. (1996). Priority Effects in Temporary Pools: Nature and Outcome of Mosquito Larva-Toad Tadpole Interactions Depend on Order of Entrance. *The Journal of Animal Ecology*, *65*(1), 77. https://doi.org/10.2307/5701

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bouki, C., Venieri, D., & Diamadopoulos, E. (2013). Detection and fate of antibiotic resistant bacteria in wastewater treatment plants: A review. *Ecotoxicology and Environmental Safety*, *91*, 1–9. https://doi.org/10.1016/j.ecoenv.2013.01.016

Bowers, S. J., Vargas, F., González, A., He, S., Jiang, P., Dorrestein, P. C., Knight, R., Wright, K. P., Lowry, C. A., Fleshner, M., Vitaterna, M. H., & Turek, F. W. (2020). Repeated sleep disruption in mice leads to persistent shifts in the fecal microbiome and metabolome. *PLoS ONE*, *15*(2), 1–25. https://doi.org/10.1371/journal.pone.0229001

Brown, S. P., Cornforth, D. M., & Mideo, N. (2012). Evolution of virulence in opportunistic pathogens: Generalism, plasticity, and control. *Trends in Microbiology*, *20*(7), 336–342. https://doi.org/10.1016/j.tim.2012.04.005

Buckling, A., Kassen, R., Bell, G., & Rainey, P. B. (2000). Disturbance and diversity in experimental microcosms. *Nature*, *408*(6815), 961–964. https://doi.org/10.1038/35050080

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, *11*(12), 2639–2643. https://doi.org/10.1038/ismej.2017.119

Campoli-richards, D. M., & Brogden, R. N. (1987). Sulbactam / Ampicillin Properties , and Therapeutic Use. *Drugs*, *33*, 577–609.

Chambers, M. C., MacLean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M. Y., Paulse, C., Creasy, D., … Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918–920. https://doi.org/10.1038/nbt.2377

Chase, J. M. (2003). Community assembly: When should history matter? *Oecologia*, *136*(4), 489–498. https://doi.org/10.1007/s00442-003-1311-7

Chopra, I., & Roberts, M. (2001). Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews*, *65*(2), 232–260. https://doi.org/10.1002/anie.197507211

Connell, J. H. (1978). Diversity in tropical rain forests and coral reefs. *Science*, *199*(4335), 1302–1310. https://doi.org/10.1126/science.199.4335.1302

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2017). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *BioRxiv*, 1–14. https://doi.org/10.1101/221499

Dray, S., & Dufour, A. B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20. https://doi.org/10.18637/jss.v022.i04

Dumaru, R., Baral, R., & Shrestha, L. B. (2019). Study of biofilm formation and antibiotic resistance pattern of gram-negative Bacilli among the clinical isolates at BPKIHS, Dharan. *BMC Research Notes*, *12*(1), 1–6. https://doi.org/10.1186/s13104-019-4084-8

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., & Huttenhower, C. (2012). Microbial co-occurrence relationships in the Human Microbiome. *PLoS Computational Biology*, *8*(7). https://doi.org/10.1371/journal.pcbi.1002606

Ferrenberg, S., O'neill, S. P., Knelman, J. E., Todd, B., Duggan, S., Bradley, D., Robinson, T., Schmidt, S. K., Townsend, A. R., Williams, M. W., Cleveland, C. C., Melbourne, B. A., Jiang, L., & Nemergut, D. R. (2013). Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *ISME Journal*, *7*(6), 1102–1111. https://doi.org/10.1038/ismej.2013.11

Fraterrigo, J. M., & Rusak, J. A. (2008). Disturbance-driven changes in the variability of ecological patterns and processes. *Ecology Letters*, *11*(7), 756–770. https://doi.org/10.1111/j.1461-0248.2008.01191.x

Fukami, T. (2015). Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics*, *46*, 1–23.

Geyrhofer, L., & Brenner, N. (2020). Coexistence and cooperation in structured habitats. *BMC Ecology*, *20*(1), 1–15. https://doi.org/10.1186/s12898-020-00281-y

Ghoul, M., & Mitri, S. (2016). The Ecology and Evolution of Microbial Competition. *Trends in Microbiology*, *24*(10), 833–845. https://doi.org/10.1016/j.tim.2016.06.011

Green, J. L., & Ostling, A. (2003). Endemics-Area Relationships : The Influence of Species Dominance and Spatial Aggregation. *Ecology*, *84*(11), 3090–3097.

Grossman, T. H. (2016). Tetracycline antibiotics and resistance. *Cold Spring Harbor Perspectives in Medicine*, *6*(4), 1–24. https://doi.org/10.1101/cshperspect.a025387

Hardin, G. (1960). The Competitive Exclusion Principle. *Science*, *131*(3409), 1292–1297.

Hargreaves, S. K., Williams, R. J., & Hofmockel, K. S. (2015). Environmental filtering of microbial communities in agricultural soil shifts with crop growth. *PLoS ONE*, *10*(7), 1–14. https://doi.org/10.1371/journal.pone.0134345

Hayward, C., Ross, K. E., Brown, M. H., & Whiley, H. (2020). Water as a source of antimicrobial resistance and healthcare-associated infections. *Pathogens*, *9*(8), 1–21. https://doi.org/10.3390/pathogens9080667

Hossain, E., Khanam, S., Wu, C., Lostracco-Johnson, S., Thomas, D., Katemauswa, M., Gosmanov, C., Li, D., Woelfel-Monsivais, C., Sankaranarayanan, K., & McCall, L. I. (2019). 3D mapping of host-parasite-microbiome interactions reveals metabolic determinants of tissue tropism and disease tolerance in Chagas disease. *BioRxiv*, *July*, 1–15. https://doi.org/10.1101/727917

Hunt, D. E., & Ward, C. S. (2015). A network-based approach to disturbance transmission through microbial interactions. *Frontiers in Microbiology*, *6*(OCT), 1–8. https://doi.org/10.3389/fmicb.2015.01182

Jia, Q., Chen, X., Köllner, T. G., Rinkel, J., Fu, J., Labbé, J., Xiong, W., Dickschat, J. S., Gershenzon, J., & Chen, F. (2019). Terpene Synthase Genes Originated from Bacteria through Horizontal Gene Transfer Contribute to Terpenoid Diversity in Fungi. *Scientific Reports*, *9*(1), 1–11. https://doi.org/10.1038/s41598-019-45532-1

Jiang, L., & Patel, S. N. (2008). Community assembly in the presence of disturbance: A microcosm experiment. *Ecology*, *89*(7), 1931–1940. https://doi.org/10.1890/07-1263.1

Kim, H. W., Wang, M., Leber, C. A., & Nothias, L. (2020). NPClassifier : deep neural structural classification tool for natural products. *Bioarxiv*. https://chemrxiv.org/articles/preprint/NPClassifier_A_Deep_Neural_Network-Based_Structural_Classification_Tool_for_Natural_Products/12885494/1

Kraus, E. A., Beeler, S. R., Mors, R. A., Floyd, J. G., Stamps, B. W., Nunn, H. S., Stevenson, B. S., Johnson, H. A., Shapiro, R. S., Loyd, S. J., Spear, J. R., & Corsetti, F. A. (2018). Microscale biosignatures and abiotic mineral authigenesis in Little Hot Creek, California. *Frontiers in Microbiology*, *9*(MAY), 1–13. https://doi.org/10.3389/fmicb.2018.00997

Lawson, C. E., Wu, S., Bhattacharjee, A. S., Hamilton, J. J., McMahon, K. D., Goel, R., & Noguera, D. R. (2017). Metabolic network analysis reveals microbial community interactions in anammox granules. *Nature Communications*, *8*(May), 1–12. https://doi.org/10.1038/ncomms15416

LeBel, M. (1988). Ciprofloxacin: Chemistry, Mechanism of Action, Resistance, Antimicrobial Spectrum, Pharmacokinetics, Clinical Trials, and Adverse Reactions. *Pharmacotherapy*, *8*(1), 3–30.

Mahizan, N. A., Yang, S. K., Moo, C. L., Song, A. A. L., Chong, C. M., Chong, C. W., Abushelaibi, A., Erin Lim, S. H., & Lai, K. S. (2019). Terpene derivatives as a potential agent against antimicrobial resistance (AMR) pathogens. *Molecules*, *24*(14), 1–21. https://doi.org/10.3390/molecules24142631

Martin, B. D., Witten, D., & Willis, A. D. (2020). Modeling microbial abundances and dysbiosis with beta-binomial regression. *Annals of Applied Statistics*, *14*(1), 94–115.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal*, *17*(1).

McCall, L. I., Tripathi, A., Vargas, F., Knight, R., Dorrestein, P. C., & Siqueira-Neto, J. L. (2018). Experimental Chagas disease-induced perturbations of the fecal microbiome and metabolome. *PLoS Neglected Tropical Diseases*, *12*(3), 1–15. https://doi.org/10.1371/journal.pntd.0006344

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0061217

Morton, J. T., Aksenov, A. A., Nothias, L. F., Foulds, J. R., Quinn, R. A., Badri, M. H., Swenson, T. L., van Goethem, M. W., Northen, T. R., Vazquez-Baeza, Y., Wang, M., Bokulich, N. A., Watters, A., Song, S. J., Bonneau, R., Dorrestein, P. C., & Knight, R. (2019). Learning representations of microbe–metabolite interactions. *Nature Methods*, *16*(12), 1306–1314. https://doi.org/10.1038/s41592-019-0616-3

Myers, O. D., Sumner, S. J., Li, S., Barne, S., & Du, X. (2017). One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks. *Analytical Chemistry*, *89*(17), 8696–8703.

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and Processes of Microbial Community Assembly. *Microbiology and Molecular Biology Reviews*, *77*(3), 342–356. https://doi.org/10.1128/mmbr.00051-12

Nothias, L. F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., Protsyuk, I., Ernst, M., Tsugawa, H., Fleischauer, M., Aicheler, F., Aksenov, A. A., Alka, O., Allard, P. M., Barsch, A., Cachet, X., Caraballo-Rodriguez, A. M., da Silva, R. R., Dang, T., …

Dorrestein, P. C. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nature Methods*, *17*(9), 905–908. https://doi.org/10.1038/s41592-020-0933-6

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., & Peter Solymos, M. Henry H. Stevens, E. S. and H. W. (2019). *vegan: Community Ecology Package* (2.5-6). R.

Paczosa, M. K., & Mecsas, J. (2016). Klebsiella pneumoniae: Going on the Offense with a Strong Defense Michelle. *Microbiology and Molecular Biology Reviews*, *80*(3), 629–661. https://doi.org/10.1128/MMBR.00078-15.Address

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2015). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414.

Parveen, N., & Cornell, K. A. (2012). Critical Enzyme for Bacterial Metabolism. *Molecular Microbiology*, *79*(1), 7–20. https://doi.org/10.1111/j.1365-2958.2010.07455.x.Methylthioadenosine/S-adenosylhomocysteine

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

Podschun, R., & Ullmann, U. (1998). Klebsiella spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clinical Microbiology Reviews*, *11*(4), 589–603. https://doi.org/10.1128/cmr.11.4.589

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing

and web-based tools. *Nucleic Acids Research*, *41*(D1), 590–596. https://doi.org/10.1093/nar/gks1219

Raguso, R., Agrawal, A., Douglas, A., Jander, G., Kessler, A., Poveda, K., & Thaler, J. (2015). The raison d ' e ˆ tre of chemical ecology. *Ecology*, *96*(3), 617–630.

Rainey, P. B., & Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*, *394*(6688), 69–72. https://doi.org/10.1038/27900

Reasoner, D. J., & Geldreich, E. E. (1985). A new medium for the enumeration and subculture of bacteria from potable water. *Applied and Environmental Microbiology*, *49*(1), 1–7. https://doi.org/10.1128/aem.49.1.1-7.1985

Riscoe, M. K., Ferro, A. J., & Fitchen, J. H. (1988). Analogs of 5-methylthioribose, a novel class of antiprotozoal agents. *Antimicrobial Agents and Chemotherapy*, *32*(12), 1904–1906. https://doi.org/10.1128/AAC.32.12.1904

Roberts, M. C. (1996). Tetracycline resistance determinants: Mechanisms of action, regulation of expression, genetic mobility, and distribution. *FEMS Microbiology Reviews*, *19*(1), 1–24. https://doi.org/10.1016/0168-6445(96)00021-6

Rudolf, J. D., Alsup, T. A., Xu, B., & Li, Z. (2021). Bacterial terpenome. *Natural Product Reports*. https://doi.org/10.1039/D0NP00066C

Samuels, C. L., & Drake, J. A. (1997). Divergent perspectives on community convergence. *Trends in Ecology and Evolution*, *12*(11), 427–432. https://doi.org/10.1016/S0169-5347(97)01182-8

Santillan, E., Seshan, H., Constancias, F., Drautz-Moses, D. I., & Wuertz, S. (2019). Frequency of disturbance alters diversity, function, and underlying assembly mechanisms of complex

bacterial communities. *Npj Biofilms and Microbiomes*, *5*(1). https://doi.org/10.1038/s41522-019-0079-4

Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, *27*(4), 592–593. https://doi.org/10.1093/bioinformatics/btq706

Sekowska, A., Ashida, H., & Danchin, A. (2019). Revisiting the methionine salvage pathway and its paralogues. *Microbial Biotechnology*, *12*(1), 77–97. https://doi.org/10.1111/1751-7915.13324

Shade, A. (2018). Understanding Microbiome Stability in a Changing World. *MSystems*, *3*(2), 1–4. https://doi.org/10.1128/msystems.00157-17

Shade, A., Peter, H., Allison, S. D., Baho, D. L., Berga, M., Bürgmann, H., Huber, D. H., Langenheder, S., Lennon, J. T., Martiny, J. B. H., Matulich, K. L., Schmidt, T. M., & Handelsman, J. (2012). Fundamentals of microbial community resistance and resilience. *Frontiers in Microbiology*, *3*(DEC), 1–19. https://doi.org/10.3389/fmicb.2012.00417

Shade, A., Read, J. S., Youngblut, N. D., Fierer, N., Knight, R., Kratz, T. K., Lottig, N. R., Roden, E. E., Stanley, E. H., Stombaugh, J., Whitaker, R. J., Wu, C. H., & McMahon, K. D. (2012). Lake microbial communities are resilient after a whole-ecosystem disturbance. *ISME Journal*, *6*(12), 2153–2167. https://doi.org/10.1038/ismej.2012.56

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models. *Genome Research*, *13*, 2498–2504. https://doi.org/10.1101/gr.1239303.metabolite

Steele, J. A., Countway, P. D., Xia, L., Vigil, P. D., Beman, J. M., Kim, D. Y., Chow, C. E. T., Sachdeva, R., Jones, A. C., Schwalbach, M. S., Rose, J. M., Hewson, I., Patel, A., Sun, F., Caron, D. A., & Fuhrman, J. A. (2011). Marine bacterial, archaeal and protistan association

networks reveal ecological linkages. *ISME Journal*, *5*(9), 1414–1425. https://doi.org/10.1038/ismej.2011.24

Stewart, P. S., Camper, A. K., Handran, S. D., Huang, C. T., & Warnecke, M. (1997). Spatial distribution and coexistence of Klebsiella pneumoniae and Pseudomonas aeruginosa in biofilms. *Microbial Ecology*, *33*(1), 2–10. https://doi.org/10.1007/s002489900002

van Dorp, L., Wang, Q., Shaw, L. P., Acman, M., Brynildsrud, O. B., Eldholm, V., Wang, R., Gao, H., Yin, Y., Chen, H., Ding, C., Farrer, R. A., Didelot, X., Balloux, F., & Wang, H. (2019). Rapid phenotypic evolution in multidrug-resistant Klebsiella pneumoniae hospital outbreak strains. *Microbial Genomics*, *5*(4), 1–11. https://doi.org/10.1099/mgen.0.000263

Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A., & Knight, R. (2013). EMPeror: A tool for visualizing high-throughput microbial community data. *GigaScience*, *2*(1), 2–5. https://doi.org/10.1186/2047-217X-2-16

Vellend, M. (2010). Conceptual synthesis in community ecology. *Quarterly Review of Biology*, *85*(2), 183–206. https://doi.org/10.1086/652373

Ventola, C. L. (2015). The Antibiotic Resistance Crisis Part 1: Causes and Threats. *P&T*, *40*(4), 277–283.

Verner-Jeffreys, D. W., Brazier, T., Perez, R. Y., Ryder, D., Card, R. M., Welch, T. J., Hoare, R., Ngo, T., McLaren, N., Ellis, R., Bartie, K. L., Feist, S. W., William, W. M., Adams, A., & Thompson, K. D. (2017). Detection of the florfenicol resistance gene floR in Chryseobacterium isolates from rainbow trout. Exception to the general rule? *FEMS Microbiology Ecology*, *93*(4), 1–11. https://doi.org/10.1093/femsec/fix015

von Wintersdorff, C. J. H., Penders, J., Van Niekerk, J. M., Mills, N. D., Majumder, S., Van Alphen, L. B., Savelkoul, P. H. M., & Wolffs, P. F. G. (2016). Dissemination of

antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Frontiers in Microbiology*, *7*(FEB), 1–10. https://doi.org/10.3389/fmicb.2016.00173

Wang, M., Carver, J. J., Phelan, V. v., Sanchez, L. M. ;, Garg, N. ;, Peng, Y. ;, Nguyen, D. D. ;, Watrous, J. ;, Kapono, C. A. ;, Luzzatto-Knaan, T. ;, Porto, C. ;, Bouslimani, A. ;, Melnik, A. V. ;, Meehan, M. J. ;, Liu, W.-T. ;, Crüsemann, M. ;, Boudreau, P. D. ;, Esquenazi, E. ;, Sandoval-Calderon, M. ;, … Bandeira, N. (2017). Sharing and community curation of mass spectrometry data with GNPS. *Nature Biotechnology*, *34*(8), 828–837. https://doi.org/10.1038/nbt.3597.Sharing

Wright, E. S. (2015). DECIPHER: Harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*, *16*(1), 1–14. https://doi.org/10.1186/s12859-015-0749-z

Wright, E. S. (2016). Using DECIPHER v2.0 to analyze big biological sequence data in R. *R Journal*, *8*(1), 352–359. https://doi.org/10.32614/rj-2016-025

Wyres, K. L., & Holt, K. E. (2018). Klebsiella pneumoniae as a key trafficker of drug resistance genes from environmental to clinically important bacteria. *Current Opinion in Microbiology*, *45*, 131–139. https://doi.org/10.1016/j.mib.2018.04.004

Yan, Y., Klinkhamer, P. G. L., van Veen, J. A., & Kuramae, E. E. (2019). Environmental filtering: A case of bacterial community assembly in soil. *Soil Biology and Biochemistry*, *136*(July), 107531. https://doi.org/10.1016/j.soilbio.2019.107531

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., & Glöckner, F. O. (2014). The SILVA and "all-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research*, *42*(D1), 643–648. https://doi.org/10.1093/nar/gkt1209