

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Using Genomes to Infer Interactions Shaping Community Structure in the Soil Microbiome

Permalink

<https://escholarship.org/uc/item/5k291852>

Author

Nicolas, Alexa M

Publication Date

2022

Peer reviewed|Thesis/dissertation

Using Genomes to Infer Interactions Shaping Community Structure
in the Soil Microbiome

By

Alexa M. Nicolas

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mary K. Firestone, Co-Chair

Professor Michiko E. Taga, Co-Chair

Professor Jillian F. Banfield

Professor Britt Koskella

Fall 2022

Abstract

Using Genomes to Infer Interactions Shaping Community Structure in the Soil Microbiome

by

Alexa M. Nicolas

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Mary K. Firestone, Co-Chair

Professor Michiko E. Taga, Co-Chair

Microbial residents of soil interact with one another to form communities that serve critical roles in carbon sequestration, agriculture, and nutrient cycling. Soil microbes are abundant – a gram of soil can house billions of cells – and diverse – cells represent thousands of species of archaea and bacteria. However, as new technologies have enabled cataloging a greater portion of which organisms persist and coexist in soil, mapping the network of microbial interactions is the next hurdle for appreciating the critical role of the soil microbiome in terrestrial biogeochemical cycling. From microbial genomes we can predict how organisms live and from culturing microorganisms in laboratories we can see how they behave, but illuminating how microbes act *in situ* remains the canonical paradox: in order to observe soil microbes we must perturb them. In this dissertation, multiple genome resolved ‘omic approaches are combined to study how microbes may interact in soil. Each chapter focuses on a distinct type of interaction and the cascading effects that interaction has on soil microbial communities at-large.

The first chapter reviews the varied microbial interactions that exist in soil: from contact-dependent molecular transfers to interactions at a distance that depend on molecules moving through soil. Here, I introduce the framework that a subset of interactions can be inferred in an organism’s genome. Reading genomes provides the basis for appreciating an organism’s nutritional requirements which can be used to predict how and with which organisms act in community.

In the second chapter, I turn to often overlooked ultrasmall microbes in soil from the Candidate Phyla Radiation (CPR) bacteria and DPANN archaea (an acronym of the first included phyla: Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota and Nanohaloarchaeota) lineages. We affirm that, in part, these organisms are overlooked because they comprise the soil rare biosphere and that we were able to capture these genomes by optimizing previous methods used to establish viral-enriched metagenomes. We found that, as in other systems, these organisms must acquire much of their essential building blocks from neighboring cells, suggesting they require living in close association with other organisms perhaps as epibionts or parasites. We also

determined that soil CPR appear to harbor a likely local adaptation to the oxic environment of top soils – the capacity for some form of aerobic respiration.

Moving into the third chapter, I turn attention to understanding the role of viruses in the soil microbiome. Viruses, like other soil microbes, are ubiquitous and recently have become better cataloged, but a comprehensive understanding of the ecological role of these specific and diverse predators has not yet emerged. I examined a highly dynamic moment in soil where carbon is mineralized at a disproportionately high rate compared to the rest of the year – the first rainfall in a seasonally dry annual California grassland. This system has long been fruitful for relating microbes directly to the turnover of soil carbon. Thus, we probed the role of viruses in soil and how viruses may contribute to soil carbon turnover via the characteristic resuscitation and mortality of specific microbial lineages immediately following wet-up. We combined metagenomics with stable isotope probing (SIP) using H_2^{18}O and found that dry soil is a sparse yet diverse reservoir of putative virions, of which only a subset thrives following wet-up. These active viruses appeared to closely follow host population dynamics.

In the final chapter of this dissertation, I focus on a model nutrient and its rippling effects on the microbiome, shifting focus from a specific type of soil microbe. Turning to corrinoids – cobalt-containing enzymatic cofactors used in a variety of carbon and nitrogen metabolisms by most organisms, typified by Vitamin B12 – we analyzed how varying types of this non-assimilated nutrient could differentially shift composition and function of soil microbiomes both in soil and in soil-derived enrichment cultures. We observed in soil that the addition of distinct corrinoid types led to differences in community composition, but that this effect was more muted in soil-enrichment cultures. Using metagenomics and metatranscriptomics on soil enrichment cultures, we found that some organisms showed specific differences in the number of genes differentially expressed for certain corrinoid types over others, while other organisms showed a non-specific response to distinct corrinoid types, but these expression changes only corresponded to subsequent changes in an organism's abundance in a few instances.

Overall, the results presented in this dissertation use a variety of sequencing methods and lenses to tease apart the types of interactions that cohere the soil microbiome. These interactions illustrate both the highly physically connected nature of the soil microbiome and the simultaneous diffuseness of these connections. CPR bacteria and DPANN archaea often live physically associated with other cells. Similarly, viruses are only alive inside host cells, but may also persist in soil awaiting a physical association with a host – perhaps facilitated by rain. Finally, soils maintain a stock of corrinoids, but changes in the soil corrinoid profile may preferentially target growth of certain bacteria. Through unraveling how microbes interact we will advance our understanding of critical soil processes.

*There can be no power greater
Anywhere beneath the sun
Yet what force on earth is weaker
Than the feeble strength of one?*
— Ralph Chaplin (1915)

Under any circumstances sociability is the greatest advantage in the struggle for life
— Pyotr Kropotkin, *Mutual Aid: A Factor of Evolution* (1902)

Table of Contents

Acknowledgements	iii
1. Interactions in the soil microbiome	1
<i>1.1 Contact-dependent interactions: long-term physical associations, lethal transient meetings</i>	<i>3</i>
<i>1.2 Interactions at a distance: molecular mediators of cellular communication</i>	<i>6</i>
2. Soil Candidate Phyla Radiation bacteria encode components of aerobic metabolism and co-occur with nanoarchaea in the rare biosphere of rhizosphere grassland communities	11
<i>2.1 Main Text</i>	<i>12</i>
<i>2.2 Figures</i>	<i>16</i>
3. Isotope-enrichment reveals active viruses follow microbial host dynamics during rewetting of a California grassland soil	21
<i>3.1 Introduction</i>	<i>22</i>
<i>3.2 Results</i>	<i>24</i>
<i>3.3 Discussion</i>	<i>29</i>
<i>3.4 Materials and Methods</i>	<i>35</i>
<i>3.5 Figures</i>	<i>42</i>
4. Corrinoids influence community assembly in a soil-derived enrichment	53
<i>4.1 Introduction</i>	<i>54</i>
<i>4.2 Results</i>	<i>55</i>
<i>4.3 Discussion</i>	<i>61</i>
<i>4.4 Methods</i>	<i>64</i>
<i>4.5 Figures</i>	<i>68</i>
Concluding remarks	80
References	82
Appendix / Glossary of Terms	106

Acknowledgements

During my doctoral studies I had the intellectual experience I've always dreamed of – discussions that extended out of classrooms and followed us down the hill back to our labs, successful lab meetings measured in thoughtful questions and engaging conversations, coffee catchups ending in sharing new literature from unlikely fields. For all this, I am forever indebted to the community of scholars in the Plant and Microbial Biology (PMB) Department. PMB is a community of brilliant, inquisitive, passionate coworkers that mostly don't center the human body in biology. I feel incredibly lucky for the lifelong group of peers I've found here.

In particular, thinking alongside my two(!) PhD advisors has been an absolute privilege. I am so grateful to Mary Firestone and Michi Taga for teaching me how to ask questions, develop hypotheses, and, eventually, see projects through to completion! Under Mary's guidance I learned a crucial piece of advice in academia that has never led me astray – if someone never says, "I don't know," that's the real problem. Mary is a trailblazer and situated me in a large network of exciting and truly interdisciplinary collaborations in the world of soils. I am so proud to be Mary's final conferred doctorate.

In Michi Taga at times I found a scientific confidant, who showed me that scientific expertise is more than the content of your training; it is truly in developing an analytical mind to ask the right questions. Michi created a lab culture where I learned to shirk the embarrassment of sharing my work and eventually felt empowered to present my ideas and think through issues of equity. I am so grateful to Michi for line edits on fellowship applications hours before deadlines, but more importantly for modeling so many best practices – writing grants together and working on peer reviews; teaching me pedagogy both as a first-year student in my favorite course, and again co-teaching microbial genomics remotely during the more stringent lockdown portion of the pandemic.

I have also had the impressive opportunity to work with world-renowned collaborators. I cherish every moment I've had to examine data and writing with Jill Banfield. I hope to one day share Jill's ability to look at a figure and immediately understand the nuance of a system or quickly find the substance in a dense paragraph of writing. I am thankful for all the afternoons in Jill's office, remote meetings, chats on hikes, and evenings at Serpens Ridge. Similarly, I am so grateful to Jennifer Pett-Ridge and all the moments of mulling over fantastical scientific questions we've shared, her generosity of access to resources and expertise a scientist can only dream of, and of course her delicious confections and meals. A huge thank you to Steve Blazewicz and all the iterative thinking we did together. When Steve and I hit our stride in marathon meetings, science was a fun and creative pursuit. Thank you also to my committee member, Britt Koskella who dispensed critical career advice and contributed crucial ecology expertise.

I worked with incredible scientific researchers, postdoctoral scholars, and graduate students. I am appreciative of the science and friendship I shared with Rachel Hestrin. I had two labs full of experts – Mengting Maggie Yuan, Evan Starr, Ella Sieradzki, Katerina Estera-Molina, Rachel Neurath, Jordan Hoff, Anne Kakouridis, Kenny Mok, Zoila Alvarez-Aponte, Becky Procknow, Zachary Hallberg, Amanda Shelton, Olga Sokolovskaya, Kristopher Kennedy. Gordon Pherribo's friendship in graduate school and beyond has been so important to me. Thank you to my collaborators at Lawrence Berkeley National Labs and Lawrence Livermore National Labs especially Peter Weber, Hans Carlson, and Erin Nuccio. I had an incredibly enriching time with the Environmental Microbial Genomics Group at Ecole Central de Lyon. Thank you to Mia Sungeun Lee, Graeme Nicol, and Christina Hazard for collaborations in France and beyond.

In thinking back to the moments that led me to biology, I always remember scoffing at my freshman high school biology teacher who suggested I might become a biologist. In senior year, I made my way to Dr. Allison Wheeler's AP Biology class and when she found the poetry in lecturing on X-inactivation, I was converted. I would be remiss not to thank the deeply underfunded New York City public school system, and reluctantly, The Bronx High School of Science, especially.

The University of California for me meant the first time I became a union member. Thank you to my union, UAW 2865, for, at its best, being a community of academic workers committed to leaving academia a better place than how we found it, and for the lifelong friendship I hope to cultivate with Yueran Zhang. I will always dream of a University of California for all.

Doctorates are dizzying journeys, and I am so grateful to Margiana Petersen-Rockney and Alexander Jaffe for tracing this path alongside me. Without Alex's friendship and scientific collaboration everything would have been different.

I have an extraordinary transcontinental support network of friends and family who imbue everything I do with meaning. Because best friends are the best friends, to my BFATBF, I have loved you since I was 12 and NYC will always be home because of all of you. I don't know how I lucked out with such a strong and deep crew of California best friends, but because of you Oakland almost felt like mine.

The best present in my life is my sister. Jenny, I am so grateful for always growing together. Thank you for being a keeper of my secrets even when you tell on me.

My parents are my greatest fortune. Thelma and Albert, thank you for letting me be me, unapologetically. With absolute pride my parents gave me every possible support in the world I would need to pursue my dreams. Every milestone I cross makes me feel so grateful to be your kid.

Finally, I found my commitment to science in soil. To the soils of the world, to the peoples who've inhabited California, farmed these soils, and managed these ecologies long before this state was mapped.

1. Interactions in the soil microbiome

Soils hold more carbon than the earth's atmosphere and terrestrial vegetation combined; they are second only to the ocean in their carbon sequestration potential (T W Crowther et al., 2016; Scharlemann, et al., 2014; Sokol et al., 2022). Crucial to understanding soils' carbon-holding capacity are the soil microbial communities that constitute a complex network of metabolisms that drive biogeochemical cycles and shape soil organic carbon stocks (Sokol et al., 2022). But, beyond the critical role the soil microbiome plays in carbon sequestration, microbes impact agriculture (Averill et al., 2022; Banerjee & van der Heijden, 2022; Herrera Paredes & Lebeis, 2016; Jiao et al., 2020; Stopnisek & Shade, 2021). As earth's climate continues to change, and droughts are predicted to grow longer and more severe (Cook et al., 2018; Diffenbaugh, et al., 2015), research shows microbes may serve a pivotal role as associations of microbes with plant roots can protect against drought (Coleman-Derr & Tringe, 2014; Schimel, 2018; Xu et al., 2018). Microbial communities also harness and direct nutrients to plants (Braga et al., 2016; Kakouridis et al., 2022; Kuypers et al., 2018). On bulk, the soil microbiome is vital for nutrient cycling, carbon sequestration, agriculture, and other ecosystem services, but the term "soil microbiome" is merely shorthand for a web of interactions among its microbial residents (Dal Co et al., 2020; Embree et al., 2015; Ponomarova & Patil, 2015; Seth & Taga, 2014).

With thousands of bacterial species and billions of cells per gram, soil is believed to harbor the most diverse microbial communities on earth, and is perhaps the final frontier of environmental microbiology (Fierer, 2017; Fierer & Lennon, 2011; Shang et al., 2017; Z. Zhang et al., 2017). The sheer scale of microbial diversity in soil means that methods used to characterize other environmental systems require greater inputs to capture the same or even a smaller proportion of the total community composition (Alteio et al., 2020; Lombard et al., 2011; Lynch & Neufeld, 2015; Nicolas et al., 2021). For instance, using 'omic techniques requires higher sequencing depth and subsequently greater computational power to capture more complete genomes of a similarly or smaller representative set of organisms due to the long tail in the rank abundance curve of soil microorganisms – the soil microbiome is not only highly diverse, another feature of its complexity is that the community is not dominated by a few organisms (Delgado-Baquerizo et al., 2018; Diamond et al., 2019). Further, essentially all methods to study soil microbiomes require separating the physical inanimate matter that we think of constituting soils – minerals, clay – from the biological material of soil – cells, nucleotides, proteins (Delmont et al., 2011; Göller et al., 2020; Lombard et al., 2011; Sokol et al., 2022). This separation or suspension of biology from the complex matrices of soil has dictated the lenses we can understand soil microbiology – isolations, functional assays, and more recently in situ 'omic studies (Bartelme et al., 2020; Fierer, 2017; Steen et al., 2019; Wilpiseski et al., 2019). These techniques equip researchers with the power to identify how an individual organism lives and functions, the capacity of a genome to harbor a certain function, rather than how organisms living together interact with one another.

The study of the soil microbiome actually aggregates many soil microbial communities that span soil pores, minerals, roots, fungal hyphae, and decaying plant material (Wilpiseski et al., 2019). When we isolate microbes from soil, simultaneously extricating them from their neighbors and domesticating them in a pure culture laboratory environment, we also detach them from the varied structures they inhabit. While on bulk soil microbial diversity is staggering, soil neighborhoods

are estimated to contain fewer than 100 species (Raynaud & Nunan, 2014). The soil habitat provides unique challenges for microbes – microniches of low nutrient availability, potential for oscillating redox conditions depending on the balance between inundation with water and gas exchange, physical segregation between populations that may rely on each other (Bartelme et al., 2020; Pett-Ridge et al., 2013; Sokol et al., 2022; Wilpiseski et al., 2019). The question remains: how do scientists understand the interactions between organisms that yield communities.

Microbial interactions underlie the complex structure of microbiomes

The unique obstacles and features of soil life for residing microbes also serves to make the studying of interactions between microbes challenging. The interactions between soil microbes span physical scales of proximity from nanometers to tens of microns (Raynaud & Nunan, 2014; Wilpiseski et al., 2019). These scales can be split into two broad categories: direct physical interactions between resident microbes and interactions at a distance, mediated by small molecules (Boopathi et al., 2021; Tecon et al., 2018). At each scale of interaction, there exists, in anthropomorphic terms, beneficial, neutral, and negatively connotated associations between organisms (Campbell et al., 2018; Pande et al., 2016; Rakoff-Nahoum et al., 2016; Romdhane et al., 2022; Verbruggen et al., 2012). Direct physical associations between microbial cells of different populations have distinct impacts on the flow of nutrients in and ecology of soil. For instance, predation affects where and how nutrients move between soil trophic levels (Pomeroy et al., 2007; Sistla et al., 2013), epi- or endosymbioses between organisms maintain organisms of reciprocal ecological roles (Castelle & Banfield, 2018; Musat et al., 2016), and contact-dependent secretion systems are thought to operate through mechanisms of kin selection (Boopathi et al., 2021; Wellington & Greenberg, 2019). With respect to more-distanced microbial interactions, molecular communication also takes on many forms: one organism producing a nutrient another organism needs or consuming a byproduct extracellularized by another organism (Seth & Taga, 2014); signaling between organisms through lethal means – toxins, antibiotics – or through non-virulent messenger molecules (Kothari et al., 2019; Miao et al., 2018; Roux et al., 2019).

Given the discussed challenges of studying microbial interactions in soil, my dissertation concerns itself with interactions encoded in genomes. My work reflects the perspective that certain interactions are “hardwired into genomes” and through genomes we can predict through their nutritional requirements or immune systems how organisms interact with one another (Zengler & Zaramela, 2018). These modes of interaction contrast with interlocking metabolism which may be highly dynamic through time and depend on understanding locally accessible substrates for organisms (Zengler & Zaramela, 2018). Through studying nutritional needs encoded and viral insertions in genomes, organisms may eventually be connected to one another through their interlocking requirements and traces of the past viral infection. In particular, in soil, microbial genomes on average are larger than microbial genomes in other systems i.e., gut, marine (Brewer et al., 2016; Chuckran et al., 2022; Giovannoni et al., 2014). This is thought to be because soil resources are diverse but sparse and maintaining a greater genetic inventory may enhance survival across conditions (Bartelme et al., 2020; Nielsen et al., 2021). However, soil microbes still exhibit requirements for nutrients individuals themselves cannot produce (D’Souza et al., 2018). This chapter reviews microbial interactions through the lens of contact-dependent interactions and interactions at a distance, paying particular attention to the subset of these interactions that can be observed directly in genomes. The body of the dissertation addresses: obligate parasites or

episymbionts such as bacteria of the candidate phyla radiation (CPR) discussed in chapter 2, predation via viruses examined in chapter 3, and nutritional interactions via corrinoids as a model metabolite investigated in chapter 4.

1.1 Contact-dependent interactions: long-term physical associations, lethal transient meetings

Predatory-prey interactions move nutrients through the soil microbiome

Predators can serve as top-down controls on the soil microbiome (Hungate et al., 2021; Thakur & Geisen, 2019; Zimmerman et al., 2020). They span scales, trophic levels, and taxonomic domains. Soil predators range in size from microarthropods (e.g. mites, springtails) to nematodes, protists, predatory bacteria (e.g. *Bdellovibrio*), and – at the smallest scale – soil viruses. The types of predation are diverse and may yield different effects on the composition and function of the soil microbiome. Due to this diversity, predators can be broadly categorized into: (1) those that feed selectively vs. non-selectively, due to their distinct effects on microbial community structure, and (2) those that engulf their prey compared to those that pierce or lyse their prey, due to their distinct effects on biogeochemistry (Chow et al., 2014; Zhao et al., 2019).

Nematodes and protists are two dominant bacterial predators that vary in the selectivity of their feeding habits (Gao et al., 2019; Martins et al., 2022). Bacterivorous nematodes tend to be generalists, non-selectively consuming bacteria via filter feeding (Rønn et al., 2012), whereas bacterivorous protists discriminate between prey based on factors like body size and palatability (Thakur & Geisen, 2019). For example, protists generally favor consuming Gram-negative bacteria, which are considered easier to digest, thus promoting the abundance of the thicker cell-walled Gram-positive bacteria (Rønn et al., 2012; Thakur & Geisen, 2019). The consumptive effects of protists and nematodes influence rhizosphere biogeochemistry, particularly nutrient mineralization. Protists and nematodes possess a higher carbon to nitrogen ratio than the prey they consume, they excrete excess nitrogen as a waste product (as well as other excess nutrients) (Osler & Sommerkorn, 2007), which are then used by plants and other microbes – i.e. in the ‘microbial loop’ (Clarholm, 1985). Soil fauna such as mites, springtails, and pseudoscorpions are also prominent predators of bacteria and fungi, and these mesofauna can modify microbial community structure (Crowther et al., 2015) and impact microbial efficiency (Lucas et al., 2020), and the rate of litter degradation and soil carbon accrual (Shu et al., 2021; Soong & Nielsen, 2016).

When larger predators engulf and ingest their prey, organic matter from the prey is moved into higher trophic levels (Pomeroy et al., 2007). In contrast to these larger predators, viruses and predatory bacteria are too small to engulf their prey, but can lyse them instead (Hungate et al., 2021; K. E. Williamson et al., 2017). Soil viruses act by transforming cellular biomass of the prey into extracellularized organic compounds, in the process of producing virions (nucleotides and proteins) (Kuzakov & Mason-Jones, 2018). Moreover, unlike larger predators, viruses can persist in their preys’ cells (Brum et al., 2016; Correa et al., 2021; Touchon et al., 2016). Viruses of bacteria (i.e. bacteriophages or phages) and archaeal viruses can recombine into host cells thereby integrating themselves into host genomes as prophages (process known as lysogeny) (Hampton et al., 2020; Luk et al., 2014; Pina et al., 2011). Upon certain environmental cues, prophages can excise themselves from host genomes, replicate, and may lyse host cells (Brum et al., 2016;

Turnbull et al., 2016). Another, important feature that links viral genomes to their hosts is clustered regularly interspaced short palindromic repeats (CRISPR) systems (Ofir & Sorek, 2018). CRISPR systems are a form of adaptive immunity encoded in bacteria and archaea (Burstein et al., 2016; Pinilla-Redondo et al., 2022). In these CRISPR loci, genomes harbor short stretches of viral DNA that host populations have previously encountered thereby maintaining a record of previous viral infections (Burstein et al., 2016; Hampton et al., 2020). Though many microbial lineages are devoid of CRISPR systems, through CRISPR loci and lysogenic infections, viruses are intrinsically connected to their hosts and maintain a coupled genomic record of interaction *in situ* (Andersson & Banfield, 2008; Burstein et al., 2016; Meaden et al., 2022). Simultaneously, some viruses encode their own CRISPR systems which maintains a record of its hosts' genome (Seed et al., 2013). In soil, CRISPR loci have been used to connect phages to their methylotrophic bacterial hosts and, in combination with $^{13}\text{CH}_4$ stable isotope probing, implicate phages in terrestrial methane cycling (Lee et al., 2021).

The soil virome has attracted considerable attention recently, because soil viruses are highly diverse and abundant, outnumbering microbial cells by as much as 100:1 (Sokol et al., 2022). Understanding interactions between viruses and their hosts may have profound impacts on how we understand the soil microbiome. In marine systems, viruses play a prominent role in biogeochemical cycling, as they kill up to ~40% of bacteria daily and sustain up to 55% of bacterial production by continuously liberating this carbon – which is known as ‘the viral shunt’ (Kuzyakov & Mason-Jones, 2018; Suttle, 2007; Suttle & Culley, 2007). Viruses are posited to play a similar role in soils – a specter of microbial death, shaping a large portion of organic matter turnover in soil (Kuzyakov & Mason-Jones, 2018; Sokol et al., 2022). To date, however, though more, and larger soil viral datasets have been published, researchers have not yet unified these massive cross-soil datasets to theorize a ‘soil viral shunt’. Research has started to elucidate bulk metrics of soil viral diversity, and to identify genomic capacities of viruses for organic compound degradation – such as through metabolic genes (Bi et al., 2020; Emerson et al., 2018; K. E. Williamson et al., 2017). Although linking virus and host on a fine-grain taxonomy has proved difficult *in silico*, a few recent time-resolved studies are beginning to relate viral and host dynamics (Emerson et al., 2018; Starr et al., 2019), but studies are needed to elucidate how boom and bust cycles of viruses may contribute to microbial succession and further, nutrient cycling. In one study that assembled time-resolved metatranscriptomes, RNA viral populations appeared to be highly divergent based on the presence of root litter, similar to their fungal hosts (Starr et al., 2019). These mycoviruses may thus play a key role in shaping carbon cycling in the root detritusphere as they prey on their hosts.

Predatory bacteria have received less attention than other soil predators, but recent evidence suggests they are widespread, and that they influence food web structure (Ye et al., 2020) and carbon and nutrient flows (Hungate et al., 2021). These bacterial predators employ different strategies: some are social and hunt in packs (‘wolf-pack predation,’ i.e. *Myxococcales*); some insert fang-like structures into their host to extract cytoplasm (e.g. *Vampirovibrio* and *Vampirococcus*), and others lyse their host immediately, or after taking up temporary residence in their host cytoplasm and producing daughter cells (e.g. *Bdellovibrionales*) (Hungate et al., 2021). In a meta-analysis of 15 different sites, obligate predatory bacteria were shown to grow 36% faster and assimilated carbon at rates 211% higher than nonpredatory bacteria (Hungate et al., 2021). This provides evidence in the soil microbiome for the ecological theory that higher productivity

increases predator control of lower trophic levels, and influences carbon flow through microbial food webs.

Candidate Phyla Radiation bacteria genomes reveal requisite interactions

Beyond predation, many biotic interactions remain unknown in their nature, such as those of candidate phyla radiation (CPR) bacteria, which are ultrasmall bacteria that may live as parasites or episymbionts of bacteria (Castelle et al., 2018; Castelle & Banfield, 2018). Their genomes reveal a requirement for interaction based on incomplete pathways for nucleotide biosynthesis, amino acid and lipid biosynthesis, and other critical cellular building blocks and suggest a larger role in the interaction networks that comprise the soil microbiome (Castelle et al., 2018). It is speculated that, especially in nutrient depleted environments, CPR may play an essential role in recycling nutrients and microbial biomass from dead community members (Castelle et al., 2018; Probst et al., 2018).

These unique organisms have proved difficult to isolate and were only first revealed to exist through sequencing-based techniques. Subsequent cultivation and electron microscopy of a TM7 phylotype revealed that this organism lives attached to a host cell of *Actinomyces* spp. *XH001* (X. He et al., 2015). The researchers speculated that TM7 may live a facultative parasitic lifestyle since under starvation conditions it appears TM7 may kill its Actinobacterial host. These contact-dependent interactions, perhaps characteristic of CPR, were highlighted in a groundwater system where electron microscopy, similarly, revealed a chain of small cells (each less than 10 nm in diameter) directly attached through pili-like projections to a much larger ($> 1 \mu\text{m}$) host cell (C. He et al., 2021). The small genome sizes portend smaller cells, and the loss of essential metabolic platforms implies the requirement of close association of cells provisioning cellular building blocks for the existence of these distinct bacteria phyla (C. He et al., 2021; Jaffe et al., 2020). Thus, CPR bacteria are an illustration of hypothesized interactions encoded in genomes, followed by the nature of the interaction visualized. In soil specifically, how CPR live has not been investigated with similarly in-depth analyses to other systems (i.e. host-associated systems (Jaffe et al., 2021; Utter et al., 2020)), but as explored in Chapter 2, we see the same trends in loss of complete metabolic pathways for CPR in soil as reported across environments. In soil, however, we see lineages other than Saccharibacteria encode components of aerobic metabolism and suggest adaptation to a microaerophilic environment may be more widespread in soil CPR (Nicolas et al., 2021; Starr et al., 2018).

Contact-dependent secretion systems transfer molecular effectors directly into neighboring cells

A critical component of microbial interactions is how individual cells alter the environment around them through the export of various signals, nutrients, and toxins. Bacteria maintain seven major secretion systems (Green & Mecsas, 2016). Of these systems that translocate proteins across membranes, three in gram negative bacteria can reach into the membrane of a neighboring cell: type III secretion systems (T3SS) enable ‘injection’ of effectors into eukaryotic cells; type IV secretion systems (T4SS) and type VI (T6SS) secretion systems transfer effectors into bacteria or eukaryotes (Green & Mecsas, 2016). These secretion systems allow bacterial effector proteins to enter neighboring cells to, potentially, reprogram cells. For these direct contact dependent interactions the degree of desiccation in a soil influences the probability of cells coming into

contact – higher water potentials are predicted to decrease cell-cell contacts (Tecon et al., 2018). In a metagenomic study analyzing the cross-domain interactions between residents of the soil microbiome, researchers measured ^{13}C Carbon isotopic incorporation in microbes from the roots of *Avena fatua* grown in $^{13}\text{CO}_2$ and assembled bacterial genomes that encoded all types of secretion systems (Starr et al., 2021). In particular, a *Burkholderiales* genome with detected isotopic enrichment encoding T4SS and T6SS, also maintained a T3SS alongside a cluster of genes with homology to plant effector proteins, suggesting that this organism may act as a plant pathogen (Starr et al., 2021). However, rather than T3SS seeming to serve strictly in pathogenesis, there is evidence of promoted mutualisms. Specifically, for differing strains of *Pseudomonas fluorescens* T3SS contribute to mutualisms with mycorrhizal fungi that promote plant-fungal symbioses (Nazir et al., 2017).

Contact-dependent secretory systems such as T6SS may have profound ecological implications. In a plant assay, a soil proteobacteria, *Agrobacterium tumefaciens*, used its unique nuclease effector via a T6SS to kill other co-inoculated bacteria *Pseudomonas aeruginosa* and sibling *Agrobacterium tumefaciens* cells, which was interpreted as beneficial for niche colonization (Ma et al., 2014). Broadly studies of T6SS in plant models are thought to exemplify microbial competition through what is described as inter-bacterial warfare for access to nutrients (Gallegos-Monterrosa & Coulthurst, 2021). Of course, bacteria also encode mechanisms to defend against lethal effectors, implying that living in a mixed microbial environment where exposure to toxins is a feature of microbiomes means genomes must encode defense systems to anticipate toxin exposure. First, bacteria recognize self-versus-non-self via identification proteins encoded in genomes (Zepeda-Rivera et al., 2018) and self is thus not absolutely defined as genetic relatedness, but rather presence of a gene or specific allele (Garcia et al., 2016). Second, organisms, such as those in the human gut, can acquire gene clusters via mobile elements that neutralize T6SS-delivered toxic effectors, i.e., acquired interbacterial defense gene clusters (Ross et al., 2019). While these mechanisms to neutralize toxins from contact-dependent interactions themselves are encoded in genomes, these are not necessarily cognate systems. Thus, we can presume which organisms initiate these interactions from genomes, but not tease apart their interacting partner in a community context.

1.2 Interactions at a distance: molecular mediators of cellular communication

Contact-dependent secretion systems mediate transfer of small molecules directly into neighboring cells, but given that around 12 μm represents the average calculated inter-cell distance between soil microbes (Raynaud & Nunan, 2014), perhaps it is more common for what is exported out of cells to reach cells at a distance – for cellular alteration of a local environment to modulate peripheral cells. Water content is an important consideration in these interactions at a distance – lower water content means water soluble molecules are limited in their ability to diffuse across distances (Stark & Firestone, 1995). For gas phase molecules, decreasing water content can mean faster movement through soil pores (Stark & Firestone, 1995). Changing water potential, too, can impact what cells release: a large increase in water potential, as a shock, can increase solutes released by cells (Halverson et al., 2000). The distances extracellularized molecules diffuse are also influenced by the scale of soil aggregates. Most microbes live between and within aggregates and microaggregates measuring in the range of 250 μm (Wilpiseski et al., 2019). A variety of molecules constitute the molecular messages of soil interactions: signaling molecules such as those

of quorum sensing systems, distinct forms of nitrogen that are transformed in subsequent steps to comprise the soil nitrogen cycle, and nutrients such as vitamins produced by a subset of microbes and used broadly.

Cells monitor and make decisions based on population density using quorum sensing systems (Ng & Bassler, 2009). These systems encode cell-to-cell communication wherein bacteria produce, extracellularize, and detect from the environment molecular signals, termed autoinducers, as a proxy for population density (Ng & Bassler, 2009). Studies of plant-microbe interactions have shown that quorum sensing systems seem to influence mechanisms of plant growth promotion, primarily through the canonical quorum sensing signal N-acyl homoserine lactones (AHLs) relying on LuxR/LuxI homologs in gram-negative bacteria (Chamkhi et al., 2020). While the mechanisms of these interactions still call for teasing apart in soil, there is evidence that quorum sensing supports both within population communication of cell density and between population communication (Garcia et al., 2016; Scott & Hasty, 2016; Silpe & Bassler, 2019). In the rhizosphere of the common bean plant, *Phaseolus vulgaris*, the quorum sensing regulator CinR is important when *Rhizobium etli* are co-inoculated with the soil isolate *Rhizobium fabae* (Miao et al., 2018). The co-inoculation of these rhizobia led to higher wildtype *R. etli* adherence to root hairs compared to co-inoculation with a *cinR* knockout strain of *R. etli* 51]. When *R. etli* was grown with *R. fabae* AHLs there was observed downstream upregulation of *R. etli*'s quorum sensing systems 51]. Underscoring the importance of quorum sensing for plant growth, quorum sensing signals are reported in higher abundance in the rhizosphere compared to measured bulk soil levels (DeAngelis et al., 2008). Quorum sensing-based cellular communication is not strictly between plants and bacteria, or bacteria and other bacteria: more recently soil viruses were also found to alter induction of prophage in the presence of bacterial AHLs (Ghosh et al., 2009; Liang et al., 2020) and others have shown that phages encode proteins that detect host autoinducers (Silpe & Bassler, 2019).

Quorum sensing system-based interactions across populations intersect with other cellular processes to regulate and connect to soil metabolic transformations. Specifically, there exists a growing link between understandings of quorum sensing interactions and nitrogen cycling metabolisms. As individual microbes cannot perform the complete cycle of nitrogen – from nitrogen fixation to ammonia oxidation i.e. nitrification to the reduction of nitrate and nitrite to nitrous oxide i.e. denitrification – many microbes perform the isolated functions through metabolic handoffs in modules that constitute the holistic process (Diamond et al., 2019; Hug & Co, 2018; Kuypers et al., 2018; Matheus Carnevali et al., 2021). In the rhizosphere, where many nitrogen-fixing bacteria reside in symbiosis with plants there is an observed correlation between increased cell densities, AHLs, and enzymes important in nitrogen cycling (DeAngelis et al., 2008). Further, the production of extracellular enzymes that can degrade bioavailable forms of nitrogen – chitinases and proteases – was found to be disrupted in soil isolates when quorum sensing is disrupted (DeAngelis et al., 2008). In another study, a nitrite oxidizing bacteria, *Nitrobacter winogradskyi*, revealed the connection between nitrification and quorum sensing, finding that the expression of *nirK* and *nirK* gene clusters was controlled by quorum signaling and that interruption of quorum sensing through the targeted degradation of AHLs significantly decreased nitrification (Mellbye et al., 2016).

One microbe's byproduct becomes another's substrate

In the soil nitrogen cycle the products of one microbial metabolism become the fuel for another microbe's life. As products become substrates, these interactions may span distinct portions of the soil microstructure, with microbes capable of distinct nitrogen metabolisms residing in different size and chemistry microaggregates (Wilpiseski et al., 2019). The distinct reactions of the nitrogen cycle also predominate in different soil hotspots and soil depths (Diamond et al., 2019; Kuypers et al., 2018; Matheus Carnevali et al., 2021). Rhizobia in the rhizosphere fix nitrogen and often inhabit root nodules and symbiose with plants as they provision bioavailable ammonia (Kuypers et al., 2018). Ammonium leaks into the surrounding environment and in the oxidation of ammonia to nitrate cooperative interactions can form such as that between the ammonia-oxidizing archaea, *Nitrosopumilus* spp., and the nitrite-oxidizing bacteria, *Nitrospina* spp. (H. Koch et al., 2019; Kuypers et al., 2018). While these modular processes of nitrogen cycling are encoded in genomes, which microbes are involved in and how these metabolic handoffs occur may be spread across soil structure especially due to the solubility of the nitrogen species. Thus, these metabolic interactions may be highly dynamic through space and time (Zengler & Zaramela, 2018).

Of literature on metabolic interactions between microbes in a microbiome context there exists much more research in human microbiome studies than in soil. A critical component of this type of research is understanding how and which microbes differentially uptake nutrients freely available in their environment. These nutrients may be produced by other microbes (Baran et al., 2015), a nearby plant root (Zhalnina et al., 2018), fungal hyphae (Nuccio et al., 2013, 2022), or another nearby organism. In a study of exometabolites released by the common oat plant, *Avena barbata*, bacterial isolates from rhizosphere samples exhibited different responses to different root exudates such as amino acids, organic acids, sugars, nucleotides, fatty acids, and amines, which suggested a mechanism for community assembly around plant roots during plant growth (Zhalnina et al., 2018). These distinct substrate preferences were also shown to impact niche partitioning in the context of soil biological crusts (Baran et al., 2015). Of 372 abundant microbial metabolites, only two were demonstrated to be metabolized by all six studied biocrust isolates (Baran et al., 2015). The potential for cross-feeding was also illustrated in this study: 80 metabolites were observed to be taken up by at least one isolate and released by a different isolate (Baran et al., 2015). However, determining *in situ* substrate preferences and availability from genomes alone to determine interactions is not yet possible.

Producer microbes complement encoded nutritional deficiencies

Organisms' nutritional requirements, on the other hand, are constant and fixed in genomes (Zengler & Zaramela, 2018). In the study of a soil fungal-bacterial co-culture, the fungal partner, *Serendipita indica*, exists as an auxotroph for the B vitamin, thiamin, provided by *Bacillus subtilis* (Jiang et al., 2018). However, the complementation of *S. indica*'s thiamin auxotrophy required spatiotemporal segregation of the interacting partners when grown in a petri dish (Jiang et al., 2018). Another model across environmental systems for understanding nutritional interactions is through iron-scavenging siderophores (Braga et al., 2016; Butaite et al., 2017; Traxler et al., 2012). A study of soil *Pseudomonas* observed that pyoverdine, a type of siderophore that binds environmental iron, is not produced by all *Pseudomonas* strains, but that non-producers co-occur with pyoverdine producers and encode multiple pyoverdine receptors that enable uptake of pyoverdine produced by other community members (Butaite et al., 2017). These types of analyses

combine genomics with lab culture studies and advance the mapping of the complex network of microbial interactions in soil.

Outside of laboratory co-cultures, consortia, and supernatant studies, which organisms supplement an organism's nutritional requirements and the nature of the interaction that may complement an auxotroph are largely unknown. For example, further examined in the fourth chapter of this dissertation, is the broad use of Vitamin B12 analogs, corrinoids, of soil bacteria. Corrinoids are cobalt-containing modified tetrapyrroles that are used as cofactors in a variety of carbon and nitrogen metabolisms by most organisms (Roth et al., 1996). Corrinoid metabolisms were found in 86% of bacterial genomes from a database of over 11,000 genomes on IMG, but the complete *de novo* corrinoid biosynthesis pathway was only present in 37% of bacterial genomes (Shelton et al., 2019), suggesting corrinoids are a widely shared metabolite in microbiomes. A study of Vitamin B12 in soil recently revealed that the bacterial-wide phenomenon of corrinoid production maintained in only a subset of genomes holds true for soil, predicting that less than 10% of soil microbes encode complete *de novo* biosynthesis of B12 while the majority of soil microbes contain B12 transporters and dependent enzymes (Lu et al., 2020). Chemical analysis of corrinoids in soil uncovered that B12 is the dominant form, with a small portion of soil's corrinoid profile persisting as 5-methylbenzimidazolyl cobamide (Hallberg et al., 2022). As early as 1951, scientists showed that addition of vitamin B12 to culture media supported isolation and growth of more microbes from soil (Lochhead & Thexton, 1951). In an 18-member microbial consortia containing a cyanobacterial autotroph, researchers combined metagenomics with isolate growth studies to hypothesize a network of interactions centered around cofactor syntrophy (Romine, Rodionov, Maezato, Osterman, et al., 2017). In particular, they proposed a network of interactions based on the exchange of corrinoids and precursors between two identified corrinoid producers, six auxotrophs for corrinoids that could not use produced precursors, and seven organisms that could salvage corrinoid intermediates (Romine, Rodionov, Maezato, Osterman, et al., 2017). This type of detailed analysis of interactions centered on fixed nutritional requirements has not yet been performed for soil microbiomes.

Conclusion

Extending models of interaction from other systems or laboratory studies to soil microbiomes are challenged by the physical complexity and sheer diversity of the soil microbiome. Both contact-dependent and contact-independent interactions are deeply influenced by soil structure, water potential, and how soluble and diffusible exchanged molecules are. The interactions that comprise soil microbial communities are varied – from predation by predators larger and smaller than their prey to episympioses and contact-dependent secretion systems; to cells sensing the size of their population, metabolic handoffs, cross-feeding and syntrophies, to facultative and obligate nutritional dependencies. The interactions intertwining soil microorganisms are intertwined – cells communicate their population density and simultaneously enable transformations of nitrogen. These interactions can be encoded in genomes either as an organisms' lifestyle – is a predator, encodes a secretion system, relies on another organism for major cellular building blocks, can perform a module in nitrogen cycling – or in a pairwise manner, directly linking a set of organisms. We see these pairwise interactions encoded in genomes, perhaps most strongly, for viruses where hosts include snippets of their viral predators in CRISPR loci or when viral genomes integrate into host genomes. However, these pairwise interactions are also encoded by cognate systems: quorum

sensing autoinducers and their corresponding receptors; secreted toxins and their neutralizers; siderophores and their receptors. When many organisms encode a specific signal, and many other organisms contain the ability to receive that signal the specific interactions that may occur become muddled. This is pertinent in the case of corrinoid nutritional requirements where a subset of organisms that produce corrinoids may broadly fill the need of many organisms that rely on corrinoids. Understanding organisms' distinct requirements encoded in their genomes portends who and how they may interact with others. Ultimately, microbial communities may do more than the sum of their parts, but it is their parts we read.

2. Soil Candidate Phyla Radiation bacteria encode components of aerobic metabolism and co-occur with nanoarchaea in the rare biosphere of rhizosphere grassland communities

Alexa M. Nicolas¹, Alexander L. Jaffe¹, Erin E. Nuccio², Michiko E. Taga¹, Mary K. Firestone^{3,4}, Jillian F. Banfield^{3,5-7*}

¹ Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA

² Nuclear and Chemical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA, USA

³ Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA

⁴ Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA

⁵ Department of Earth and Planetary Science, University of California, Berkeley, Berkeley, CA

⁶ Chan Zuckerberg Biohub, San Francisco, CA

⁷ Innovative Genomics Institute, University of California, Berkeley, Berkeley, CA

* Corresponding author email: jbanfield@berkeley.edu, 510-643-2155, 2151 Berkeley Way, Berkeley, CA 94720

Published in *mSystems*, 2021.¹

Candidate Phyla Radiation (CPR) bacteria and nanoarchaea populate most ecosystems, but are rarely detected in soil. We concentrated particles less than 0.2 μm from grassland soil, enabling targeted metagenomic analysis of these organisms, which are almost totally unexplored in largely oxic environments like soil. We recovered a diversity of CPR bacteria and some archaea sequences, but no sequences from other cellular organisms. The sampled sequences include Doudnabacteria (SM2F11) and Pacearchaeota, organisms rarely reported in soil, as well as Saccharibacteria, Parcubacteria and Microgenomates. CPR and DPANN (archaea of the phyla Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota and Nanohaloarchaeota) enrichments of 100- to 1000-fold were achieved compared to bulk soil, in which we estimate each of these organisms comprise about 1 to 100 cells per gram of soil. Like most CPR and DPANN sequenced to date, we predict these microorganisms live symbiotic, anaerobic lifestyles. However, Saccharibacteria, Parcubacteria, and Doudnabacteria genomes sampled here also encode ubiquinol oxidase operons that may have been acquired from other bacteria, likely during adaptation to aerobic soil environments. We conclude that CPR bacteria and DPANN archaea are part of the soil rare biosphere and encode unique metabolic platforms potentially evolved to live symbiotically under relatively oxic conditions.

Importance

Here we investigated overlooked microbes in soil, CPR bacteria and DPANN archaea, by size-fractionating small particles from soil, an approach typically used for recovery of viral metagenomes. Concentration of these small cells (<0.2 μm) allowed us to identify these organisms as part of the soil rare biosphere and to sample genomes that were absent from non-size fractionated metagenomes. We find that some of these predicted symbionts, which have been largely studied in anaerobic systems, have aerobic capacity acquired via lateral transfer that may enable adaptation to oxic soil environments. We estimate that there are approximately 1 to 100 cells of each of these lineages per gram of soil, highlighting that the approach provides a window onto the soil rare biosphere and its associated genetic potential.

¹All main figures for this manuscript can be found below in section 2.2. All supplementary files (including tables and methods) can be found online with the published manuscript (<https://doi.org/10.1128/mSystems.01205-20>).

2.1 Main Text

Interactions among soil microorganisms impact biogeochemical cycling and overall ecosystem function. Recent metagenomic analysis of soil microbial communities revealed that many steps of key reaction pathways central to transformations in soil are partitioned among coexisting organisms (Diamond et al., 2019). Other interactions are mediated by molecules such as vitamins and antimicrobial compounds (Abreu & Taga, 2016; Abrudan et al., 2015; Jiang et al., 2018). Further, there is the potential for a variety of symbiotic interactions, including those that involve obligate reliance on coexisting organisms for even the most basic requirements (X. He et al., 2015; Starr et al., 2018). Candidate Phyla Radiation (CPR) bacteria and DPANN archaea (an acronym of the names of the first included phyla: Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota and Nanohaloarchaeota) are detected across ecosystems and are often predicted to be obligate, anaerobic (epi)symbionts that depend on other organisms for basic cellular building blocks (Castelle et al., 2018; C. He et al., 2021; X. He et al., 2015). However, CPR bacteria and DPANN archaea have rarely been studied in relatively oxic environments or identified in soil (Delgado-Baquerizo et al., 2018; Diamond et al., 2019; Portillo et al., 2013). Genome-resolved metagenomic analyses circumvent the limitations of isolation-based methods that fail for organisms unable to grow alone and bacteria that evade detection by primers used in 16S rRNA gene surveys (Brown et al., 2015), yet there are few reports of CPR metagenome-assembled genomes (MAGs) from soil (Alteio et al., 2020; Kroeger et al., 2018; Lavy et al., 2019; Matheus Carnevali et al., 2021) almost certainly because of the rarity of these bacteria.

Prior studies of groundwater have taken advantage of the observation that CPR bacteria have ultra-small cells that pass through 0.2 μm filters to enable genome recovery for these organisms (Brown et al., 2015). Studies of other systems reveal that size fractionation of particles prior to sequencing impacts the composition and function of metagenomes (S. J. Williamson et al., 2012). However, to our knowledge, studies of 0.2 μm filtrates from soil have focused on viromes, and not assessed their microbial contents (Bi et al., 2020; Han et al., 2017; Segobola et al., 2018). Here, we took advantage of the expected very small sizes of CPR bacteria and DPANN archaeal cells (Luef et al., 2015) to concentrate them from soil. Thus, we could test the hypothesis that these anaerobic organisms are understudied parts of the soil rare biosphere, where they may have evolved pathways to persist in relatively oxic environments. We sequenced concentrated-soil-effluent that had passed through a 0.2 μm filter used to remove larger cells and recovered a diversity of bacterial and archaeal sequences.

We sampled rhizosphere-associated soil from the top 10 cm of an annual grassland from the Hopland Research and Extension Center in February 2018. We added a potassium citrate-based buffer and collected the effluent from a subset of the soil samples, which we passed through a 0.2 μm -filter, concentrated, and treated with DNase to remove extracellular DNA that could have derived from larger lysed cells (**Fig 1a, Supplementary Methods**) (Trubl et al., 2016; K. E. Williamson et al., 2013). To evaluate enrichment, bulk DNA was extracted from the same soil samples for whole community shotgun DNA sequencing, generating what are here referred to as “bulk metagenomes.” Approximately 20 Gbp of sequence was obtained from each of six concentrates and two bulk samples. In addition to recovering viral sequences and mobile elements from these small-particle-concentrate metagenomes, we reconstructed sequences from CPR and nanoarchaea genomes. From these data, we resolved 26 draft genomes that were > 70% complete

(estimated using a CPR-specific single copy gene set (Brown et al., 2015)), with < 10% contamination derived from either CPR or DPANN. No CPR or DPANN genomes were recovered from the bulk metagenomes.

Sequences from cells < 0.2 μm were almost exclusively from 15 lineages of CPR bacteria and one DPANN archaea phylum (**Fig 2; Supplementary Fig. 1**). Importantly, CPR and DPANN sequences were completely absent in bulk metagenome samples, and were only detectable at very low, if any, coverage via read mapping to assembled sequences from the small size fraction metagenomes. (**Fig. 1c**). Further, from the 74 bacterial 16S rRNA sequences recovered from the concentrate metagenomes, all of which were assigned to CPR lineages, we predict more than half (42 16S rRNA gene sequences) would not have been detected using standard amplicon sequencing primers. Notable was the phylum-level diversity of CPR lineages in the concentrate metagenomes. Previously, a genome of TM7 (Saccharibacteria) was reported from the same soil but sampled at less than 1x coverage from bulk soil (Starr et al., 2018), and Microgenomates and Parcubacteria have been genomically sampled at low abundance (Kroeger et al., 2018; Sharrar et al., 2020). To our knowledge, this is one of very few reports of Pacearchaeota and Doudnabacteria (Lavy et al., 2019; Matheus Carnevali et al., 2021) in soil and the first report of a novel clade of Saccharibacteria.

Comparing sequence coverage from concentrates to that from the bulk metagenomes, we calculate that filtration enriched the relative abundance of genomes by 100-1000x (**Fig 1c**). We approximate, given the relative abundance of the most and least abundant CPR genomes in each of the bulk and concentrate metagenome, that CPR cells may comprise on the order of 1-100 cells per gram of soil. Given estimates of 10^9 microbial cells per gram of soil, this would equate to, at maximum, $\sim 10^{-5}$ percent of microbial cells in a gram of soil (**Supplementary Methods**).

Given the unique challenges of the soil environment for microbes, we next assessed whether these soil CPR and DPANN exhibited similar traits to their counterparts in other environments. Recent studies show that CPR bacteria generally appear to have the capacity for glycolysis and fermentation (Jaffe et al., 2020), but often lack complete pathways to synthesize nucleotides *de novo* and have many gaps in metabolism that suggest an obligate symbiotic lifestyle (Castelle et al., 2018). We find that most of the genomes from this sampling effort encode the three central glycolysis enzymes reportedly found in nearly all CPR bacteria: triose phosphate isomerase (TIM), glyceraldehyde 3-phosphate (GAPDH), and phosphoglycerate kinase (PGK) (Jaffe et al., 2020). The genomes also contain few if any TCA cycle genes and lack the vast majority of genes of the electron transport chain, and for synthesis of lipids and nucleotides (**Fig 1d**), suggesting they live anaerobic lifestyles and depend on resources from other organisms (Castelle et al., 2018). However, we identified an operon encoding a multi-subunit cytochrome bo₃ ubiquinol terminal oxidase in genomes of three Doudnabacteria, eight Saccharibacteria, and one Parcubacteria as well as in non-binned CPR phyla sequences from the concentrate metagenomes. We then performed a synteny analysis (**Fig 2c**) to compare these loci to a related one from the first Saccharibacteria genome described from soil, *T. rhizospherense* (Starr et al., 2018). The comparison shows identical gene order for the *cyo* operon as in the highly studied *E. coli* K12 operon (Abramson et al., 2000) and in the *T. rhizospherense* genome (Starr et al., 2018), although some CPR loci were incomplete due to assembly fragmentation. Several CPR loci also included an ORF annotated as an oxidoreductase or a conserved hypothetical protein. While the genomes recovered do not contain

quinone biosynthesis genes, the combination of this ubiquinol oxidase and the associated oxidoreductase (**Fig 2c**), which often cooccur in genomes encoding an NADH dehydrogenase and, to varying-completeness, F-type ATPase (**Fig 1d**) suggests the possibility of some aerobic respiratory capacity. Perhaps, some form of aerobic respiration may be common in soil associated Saccharibacteria specifically, and perhaps soil CPR more broadly. We, thus, hypothesize that this operon may confer an adaptive advantage for CPR to live in aerophilic environments such as surface soil.

Next, we generated a maximum-likelihood tree of subunit 1 (CyoB) of the *cyo* operon to test whether the operon exhibited a pattern of vertical inheritance in our CPR genomes (**supplementary methods; Fig 2b**). This analysis suggests that this gene cluster has been laterally transferred from other bacteria such as Proteobacteria or Chlamydiae into these CPR at least once, with perhaps different origins for gene clusters from Parcubacteria and Doudnabacteria compared to those from Saccharibacteria. Further, based on CyoB phylogeny, *T. rhizosphaerense* appears more closely related to Saccharibacteria sequences from this study than when compared to RpS3 phylogeny, which may further underscore local adaptation to soil.

Here, we conducted a targeted study of CPR bacteria and nanoarchaea in a soil ecosystem to expand our understanding of rare soil-dwelling microbes. Using typical sequencing allocations for soil metagenomics, we were only able to recover genomes for these understudied community members through size-fractionating buffered soil. Our results indicate that CPR bacteria and DPANN archaea are relatively rare in soil as they can be difficult to recover with typical metagenomic sequencing allotments.

While the precise ecological roles of these organisms remains unclear, their predicted requirement for interaction with nearby community members to satisfy their metabolic needs, and their previously reported close physical association with other cells (C. He et al., 2021; X. He et al., 2015) suggest that they may play still undescribed roles in soil microbial interaction networks.

The ability to selectively filter soil solutions to recover CPR and DPANN genomes suggests that either these organisms attach to larger microbial cells and the association can be physically disrupted or that they are, at times, not attached to other cells. The approach enabled us to sample genetic inventories of rare soil-adapted microbes and uncover numerous genes and pathways, some of which likely evolved to handle symbiotic lifestyles under relatively oxic conditions. Specifically, we expanded the known diversity of genes and pathways in soil-adapted CPR bacteria and found that these inventories could explain the presence of these organisms, widely understood to be anaerobic, in soil. More generally, our approach provides a route to expand the known diversity of genes and pathways in the soil biosphere.

Acknowledgements: This research was supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Awards DE-SC0020163 and DE-SC0016247 to MKF and the NIH grant DP2AI117984 to MET. Work conducted at Lawrence Livermore National Laboratory was supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Award SCW1678, LLNL Lab Directed Research and Development Award 18-ERD-041, and under the auspices of the U.S. DOE under Contract DE-AC52-07NA27344. We acknowledge funding support to JFB from the Chan Zuckerberg Biohub and the Innovative Genomics Institute at UC Berkeley.

We thank Cindy J. Castelle for support inferring archaea phylogenetic relationships and Katerina Estera-Molina for her expertise and management of the Firestone Lab field work site at the Hopland Research and Extension Center (HREC). We acknowledge that HREC sits on traditional, unceded land of the Pomo Indians.

Data Availability

Curated genomes described in this study are available from ggKbase: <https://ggkbase.berkeley.edu/soilcpr> (please note that it is necessary to register for an account by provision of an email address before download) and are available under NCBI BioProject accession no. PRJNA744897. NCBI accession numbers for metagenome assembled genomes are provided in Supplementary Table 2b.

2.2 Figures

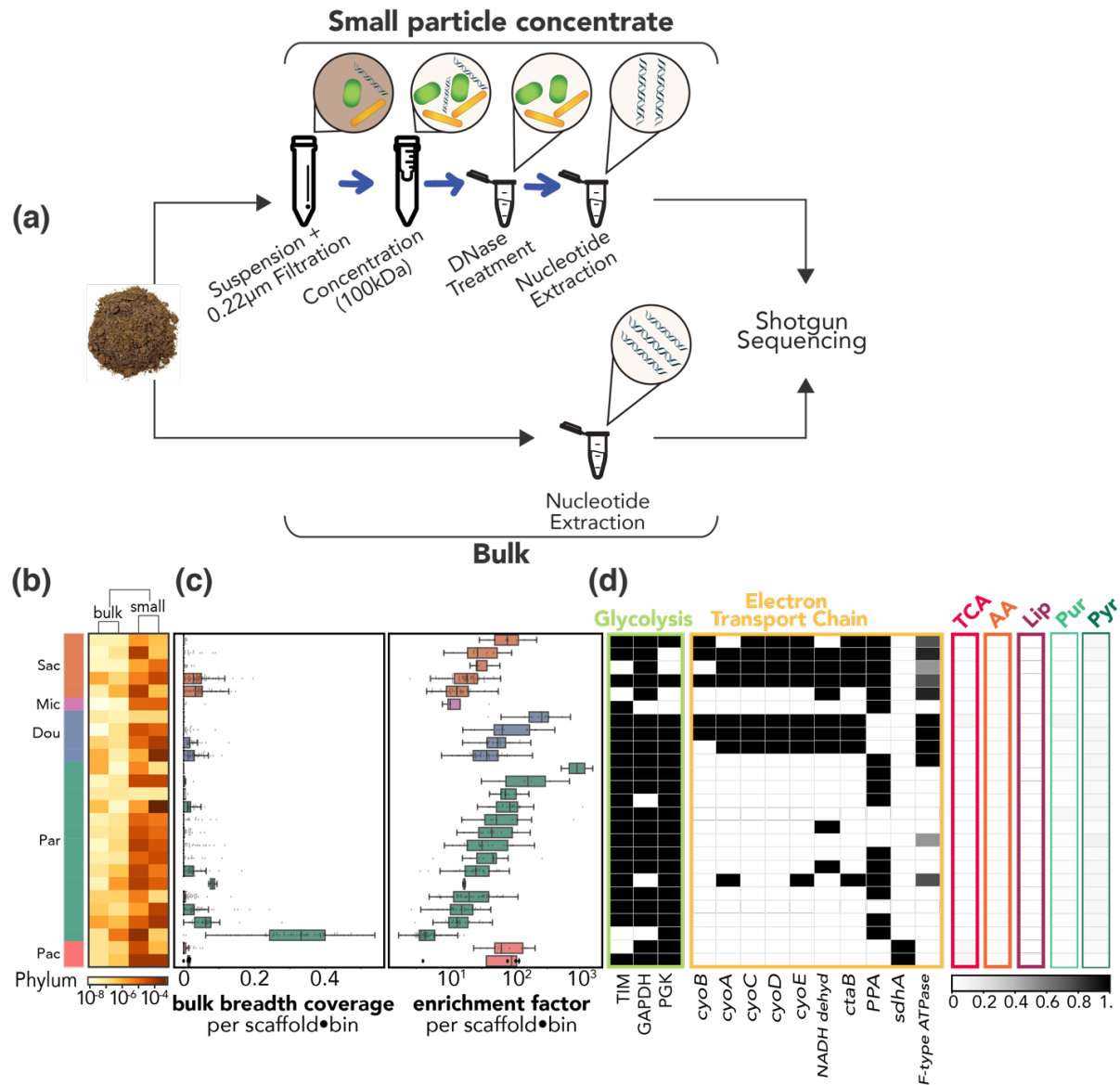
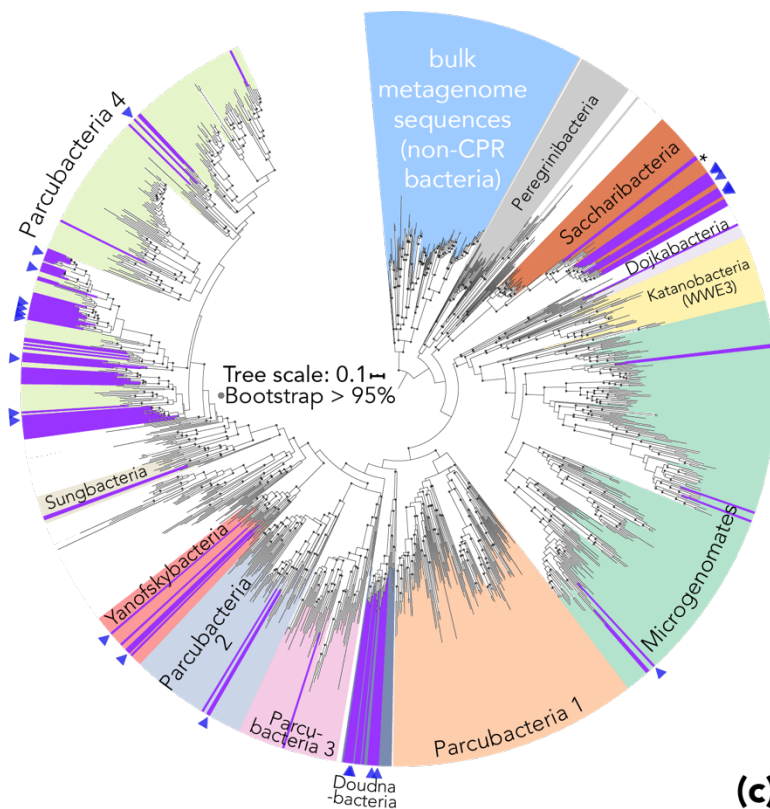
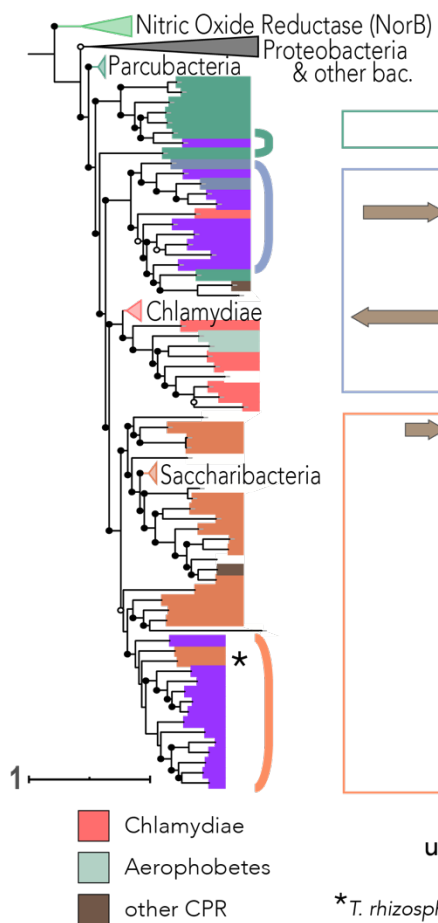


Figure 1 | Enrichment and metabolic profiles of CPR in soil concentrate metagenomes. (a) Method for concentration of small particles from soil for metagenomic sequencing (*top*) compared to sample preparation methods for bulk soil metagenomes (*bottom*). (b) Heatmap showing relative abundance of 26 organisms by phylum (Sac: Saccharibacteria, Mic: Microgenomates, Dou: Doudnabacteria, Par: Parcubacteria, Pac: Pacearchaeota) across bulk metagenomes and concentrate metagenomes. (c) Coverage based metrics showing recovery and enrichment in all concentrates combined, relative to bulk fractions combined as boxplots. *Left*: Breadth of coverage of scaffolds comprising each genome (bin) in the bulk fraction. *Right*: Enrichment factor (i.e., relative abundance of a scaffold from the concentrate metagenome over a scaffold's bulk metagenome relative abundance) for each genome. (d) Metabolic analysis of each genome including: presence of each of three glycolysis genes that are highly conserved among CPR; genes involved in electron transport chain (NADH dehydrogenase, *ctaB*: heme o synthase [EC:2.5.1.141], *PPA*: inorganic pyrophosphatase [EC:3.6.1.1], *sdhA*: succinate dehydrogenase / fumarate reductase, flavoprotein subunit [EC:1.3.5.1 1.3.5.4], including percentage completeness of F-type ATPase; and percentage completeness (grey scale) of the TCA cycle (tricarboxylic acid cycle) and pathways for AA (amino acid biosynthesis), Lip (lipid biosynthesis), Pur (purine biosynthesis), Pyr (pyrimidine biosynthesis).

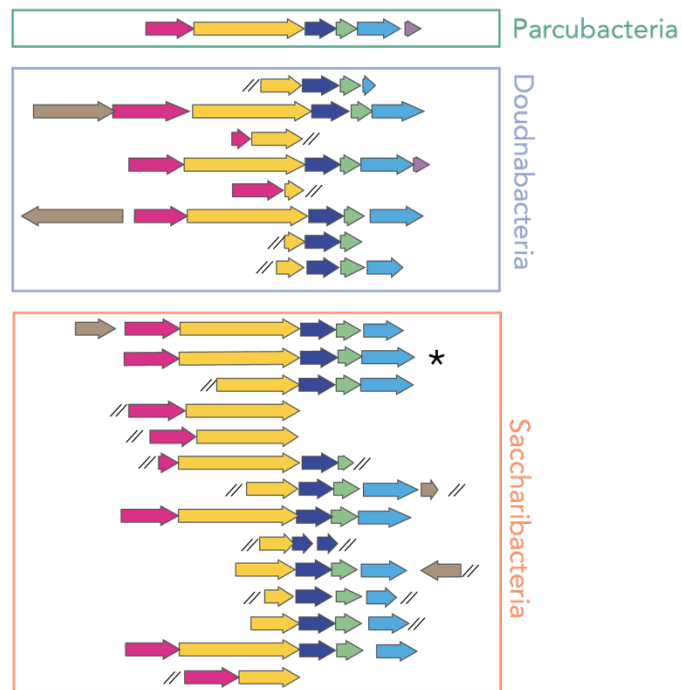
(a)



(b)



cyo
A B C D E [E.coli K12]



cytochrome bo3
ubiquinol terminal oxidase
(cyo) operon

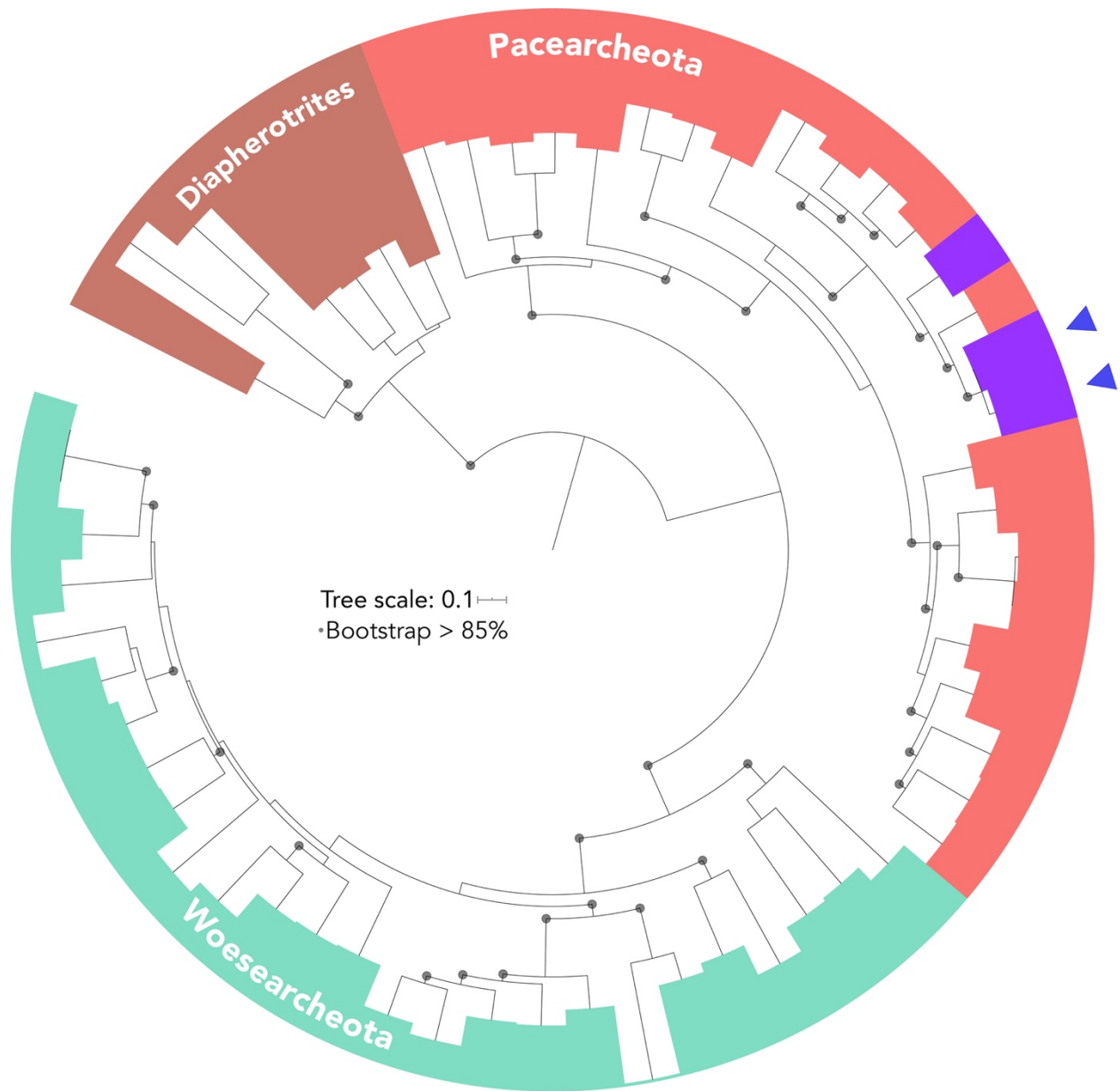
**T. rhizospherense* oxidoreductase hypothetical protein

Figure 2 | Soil CPR phylogeny and cytochrome operon synteny. Sequences assembled from the small concentrate metagenomes are shown in purple. **(a)** RpS3 tree of CPR bacteria rooted using RpS3 sequences that were assembled from bulk metagenomes, in light blue. Blue triangles denote draft genome recovered. Nodes with bootstrap values greater than or equal to 0.95 are marked as filled gray circles. **(b)** Phylogenetic relationships of cytochrome bo3 ubiquinol terminal oxidase subunit I across bacterial phyla. Circles overlaid on nodes correspond to support values (unfilled >0.50, filled >0.70). The asterisked sequences show the placement of *T. rhizospherense* CyoB, and its operon in (c) (6). Brackets next to tree tips correspond to phyla by color (green: Parcubacteria, blue: Doudnabacteria, orange: Saccharibacteria) and to sequence order in (c) synteny diagram. Tree rooted using a heme-copper oxidase superfamily member, the nitric oxide reductase (NorB). **(c)** Synteny diagram of cytochrome ubiquinol oxidase operon genes (*cyoA*, *cyoB*, *cyoC*, *cyoD*, and *cyoE*) with operon from *E. coli* K12 as reference. Scale bars correspond to the average number of substitutions per site across alignment.

Estimated count of CPR or nanoarchaea cells in soil

Bulk metagenome	Small particle concentrate metagenome
Total DNA extracted ~3 µg/g soil = 3×10^{-6} g/g soil	Total DNA extracted ~100 ng/g soil = 1×10^{-7} g / g soil
Relative abundance of most abundant CPR 10^{-6}	Relative abundance of most abundant CPR 10^{-4}
Mass of DNA/Mbp 10^{-15} g	
~Mass of 1 Cell ~ 5×10^{-15} g	
Max # cells per g soil $\frac{(10^{-6} \cdot 3 \times 10^{-6} \text{g})}{5 \times 10^{-15} \text{g}} = 600 \text{ cells}$	Max # cells per g soil $\frac{(10^{-4} \cdot 1 \times 10^{-7} \text{g})}{5 \times 10^{-15} \text{g}} = 2000 \text{ cells}$
Least abundant lineage 10^{-8}	Least abundant lineage 10^{-7}
Min # cells per g soil $\frac{(10^{-8} \cdot 3 \times 10^{-6} \text{g})}{5 \times 10^{-15} \text{g}} = 6 \text{ cells}$	Min # cells per g soil $\frac{(10^{-7} \cdot 1 \times 10^{-7} \text{g})}{5 \times 10^{-15} \text{g}} = 2 \text{ cells}$

Supplementary Table 1 | Estimated count of CPR or nanoarchaea cells in soil. *Left* shows approximation of the number of CPR or DPANN cells of a given lineage in a gram of soil using bulk metagenome parameters. *Right* provides the same estimate using parameters derived from the small particle concentrate metagenome.



Supplementary Figure 1 | Soil archaea phylogeny. Maximum-likelihood tree of RpS3 sequences of archaea. Purple colored sequences were assembled from the small concentrate metagenomes reported. Tree was rooted using RpS3 sequences of Diapherotrite phylum sequences. Blue triangles represent RpS3 sequences for which draft genomes were recovered. Nodes with bootstrap values, >85%, are marked as filled circles. Tree scale bar corresponds to the average number of substitutions per site across alignment.

3. Isotope-enrichment reveals active viruses follow microbial host dynamics during rewetting of a California grassland soil

Alexa M. Nicolas¹, Ella T. Sieradzki², Jennifer Pett-Ridge^{3,4}, Jillian F. Banfield^{2,5,6}, Michiko E. Taga¹, Mary K. Firestone^{2,6}, Steven J. Blazewicz^{3*}

¹Plant & Microbial Biology Department, University of California, Berkeley, CA, USA

²Environmental Science, Policy & Management Department, University of California, Berkeley, CA, USA

³Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, USA

⁴University of California Merced, Life & Environmental Sciences Department, Merced, CA, USA

⁵Earth and Planetary Sciences, University of California, Berkeley, CA, USA

⁶Lawrence Berkeley National Laboratory, Berkeley, California, CA, USA

*Corresponding author email: blazewicz1@llnl.gov

Released as a pre-print at bioRxiv, 2022.²

ABSTRACT

Soil viruses are highly abundant and ubiquitous, yet their impact on soil microbiome structure and function remains essentially unknown. We quantified how viruses and their hosts respond to the first rainfall of the year in seasonally dry soils – a moment commonly referred to as “wet-up”, when resident soil microbes are both resuscitated and lysed, with a disproportionately large effect on annual carbon turnover. We applied time-resolved metagenomics, viromics, and quantitative stable isotope probing to track spatiotemporal virus-host trends, and quantify cell death (and by proxy, the amount of biomass carbon released). Dry soil is a sparse yet diverse reservoir of putative virions, of which only a subset thrives following wet-up. A massive degradation of distinct viruses occurs within 24h post wet-up, but after one-week, viral biomass increased by up to seven-fold. Thriving viruses were not induced temperate phages. Our measures of viral-mediated microbial host death indicate that viruses drive a consistent rate of cell lysis after wet-up. These calculations – the first to demonstrate the quantitative impact of viral effects on soil microbiomes and carbon cycling – evidence that viruses significantly impact microbial community assembly. However, viruses do not appear to serve as top-down controls on soil microbial communities following wet-up.

²A preprint of this manuscript can be found online: doi.org/10.1101/2022.09.30.510406.

3.1 Introduction

Soil is a highly complex microbial environment and home to tens of millions of unique microbial populations (Sokol et al., 2022), with both biotic and abiotic factors influencing their assembly into complex microbial communities (Romdhane et al., 2022; Y. Yan et al., 2019; Zegeye et al., 2019). Moments of acute pressure on a community such as rapid environmental changes or fluxes in resource availability drive successional trajectories that result in differing community structures; the first rainfall following the dry season, wet-up, is such a moment (Barnard et al., 2020; Blazewicz et al., 2020; Nemergut et al., 2013; Placella & Firestone, 2013; Wang et al., 2020). However, the role of viruses in soil microbiome structuring remains unknown. Soil viruses are abundant and ubiquitous: a gram of soil can hold more than 10^9 viral-like particles (VLPs) (Sokol et al., 2022; K. E. Williamson et al., 2017). Recent metagenomic studies of soil viruses have expanded our appreciation of the many functions viruses may perform (Albright et al., 2022; Emerson, 2019; Pratama & van Elsas, 2018), but much remains unknown about how viruses impact soil microbiome turnover and community assembly, especially in response to a perturbation. In this study, we looked to a moment of annual transition in the soil microbiome to establish the role of viruses in soil microbial community assembly.

In arid and semi-arid soils that experience prolonged periods of water-deficit, the first rainfall serves as a pivotal moment which resuscitates activity in the soil microbial community and disproportionately contributes to annual carbon turnover (Barnard et al., 2020; Placella et al., 2012). Immediately following the wet-up of dry soil there is a rapid efflux of mineralized CO_2 , termed the “Birch Effect” (Birch, 1958). Previous studies suggest this CO_2 pulse is driven by the rapid growth and death of soil microorganisms (Barnard et al., 2020; Blazewicz et al., 2014; Fraser et al., 2016), but the mechanisms of mortality following rewetting are unclear. Microbial mortality after wet-up is not equitably distributed among bacterial phyla – Actinobacteria and Proteobacteria consistently have higher mortality rates (Blazewicz et al., 2020). Given the high degree of taxon-specific mortality quantified during wet-up, we hypothesized viruses act as a top-down control in community assembly and drive the differential mortality observed.

In marine systems, viruses are critically important to biogeochemistry, killing upwards of ~40% of bacteria daily and redistributing up to 55% of bacterial carbon in the proposed ‘viral shunt’ (Brum et al., 2015; Suttle, 2005; Wilhelm & Suttle, 1999). This outsized microbiome impact largely occurs in two additional ways: 1) through targeted predation which influences microbial composition, and 2) via alterations of community function caused by changes in community composition and expression of viral-encoded metabolisms (Brum et al., 2015; Hurwitz et al., 2013; Roux et al., 2016; Weitz et al., 2015). Despite the recent development of soil viromic methods (Han et al., 2017; Trubl et al., 2016), viral contributions to soil organic matter turnover or microbial mortality have yet to be quantified due to the massive diversity of soil microbes, inactivity of many of these microbes, and challenges of performing microscopy on viruses, particularly in soil (Ackermann, 2013).

In arid soil systems, abundant extracellular DNA and dormant cells can occlude distinguishing active microbial cells (Carini et al., 2016; Lennon & Jones, 2011; Rousk & Bååth, 2011), since up to 75% of cells present in soil may be inactive at any given time (Couradeau et al., 2019). Stable isotope probing (SIP) approaches enable the direct identification of active microbes through the

tracing of stable isotopes into microbial DNA and other biomolecules (Neufeld et al., 2007; Pett-Ridge & Firestone, 2017; Radajewski et al., 2000). Measuring the shift in DNA density following assimilation of a stable isotope, such as ^{18}O via H_2^{18}O (Schwartz, 2007), facilitates the quantification of growth and mortality rates of microbes, termed quantitative stable isotope probing (qSIP) (Blazewicz et al., 2020; Hungate et al., 2015; B. J. Koch et al., 2018). SIP techniques have proved powerful in identifying active microbes that metabolize soil organic matter and serve crucial roles in nutrient cycling (Greenlon et al., 2022; Starr et al., 2021), and has been used to identify active soil viral populations and viral-host linkages (Lee et al., 2021; Starr et al., 2021; Trubl et al., 2021).

The rapid growth and death of bacteria following wet-up suggests that wet-up may alter viral activities. Previous studies have characterized the immense taxonomic diversity and functions of soil viruses (Emerson et al., 2018; Paez-Espino et al., 2016). Soil DNA viruses encode a variety of metabolisms implicated in nutrient cycling including glycosyl hydrolases, chitin and polysaccharide degradation, ammonia oxidation, and sulfur metabolism (Bi et al., 2020; Emerson et al., 2018; Han et al., 2017; Reavy et al., 2014). Other work has shown that soil bacteria can maintain lysogenic viral populations in their genomes as prophages (Ghosh et al., 2008; K. E. Williamson et al., 2007). Pure culture research demonstrates these prophages often encode an integrase gene enabling phage recombination into the host chromosome (Correa et al., 2021; Nafissi & Slavcev, 2014). Prophages, upon certain cues, may excise from host chromosomes entering a lytic cycle, ultimately lysing their host cell. Induction of prophage – the switching from lysogenic to lytic cycle – can occur due to population stochasticity, by an external stress that may cause DNA damage, or through quorum sensing pathways (Ghosh et al., 2009; Howard-Varona et al., 2017). Temperate phages dominate certain marine systems and switching between lysogenic and lytic cycles corresponds to increased bacterial production (Brum et al., 2016). Similarly, integrase-encoding phages have been found at higher abundances in locations associated with more environmental stresses (Jurgensen et al., 2022). Given the unique challenges of the soil environment – extraordinary spatial heterogeneity, lack of mixing, microniches of low nutrient availability (Schimel, 2018) – some hypothesize that lysogeny may be a predominant lifestyle of soil phages (Emerson, 2019; Ghosh et al., 2008, 2009; Kimura et al., 2008; Knowles et al., 2016; Kuzyakov & Mason-Jones, 2018). In soil, environmental conditions that drive lysogeny and prophage induction largely remain unknown (Weitz et al., 2017; K. E. Williamson et al., 2007). Previous studies have shown that soil moisture modulates soil viral counts, composition, and activity (K. E. Williamson et al., 2005; R. Wu et al., 2021). Further, the summer dry-down marks a stretch of time where soil microbes maintain low activity and rates of replication (Barnard et al., 2013, 2020; Blazewicz et al., 2014; Placella & Firestone, 2013) and low host abundance is thought to favor conditions for lysogeny (Ghosh et al., 2009; McDaniel & Paul, 2005; S. J. Williamson et al., 2002). Thus, we hypothesized that rewetting of soil serves as an environmental inducer of viruses that had persisted through the summer-dry down as prophages.

Here, we quantified VLPs in soil and released during host infections using a combination of time-resolved viromics and qSIP-informed metagenomics to understand of the role of viruses in restructuring the soil microbiome following wet-up. We address the question of whether rewetting is a major inducer of prophages. The work builds on our previous efforts to link taxon-specific population dynamics with rainfall-induced carbon fluxes (Blazewicz et al., 2014, 2020). Our

analyses quantify contributions of viruses to microbial mortality, thus elucidate the contributions of viruses to the ‘Birch Effect’.

3.2 Results

Viral Succession

We established the set of MAGs and viral operational taxonomic units (vOTUs) to define the soil microbiome from 15 unfractionated metagenomes over five timepoints and 18 viromes over six timepoints (each in triplicate) generated from California grassland soil collected at the end of the dry season and wet-up in destructively-harvested microcosms (see Methods and our paired study Sieradzki et al. 2022 (Sieradzki et al., 2022)). We resolved 542 dereplicated metagenome-assembled genomes (MAGs): 503 from total soil metagenomes comprised of diverse soil bacteria and archaea; and 39 from small fraction metagenomes (Nicolas et al., 2021) i.e. viral-enriched metagenomes or viromes which represent candidate phyla radiation (CPR) bacteria. This study detected 26,368 dereplicated vOTUs from viromes (26,139) and metagenomes (229) (Figure 1). We consider vOTUs detected in viromes to be extracellularized VLPs, which are distinct from vOTUs detected in metagenomes where analysis cannot delineate between intracellular and extracellular viral populations (Supplementary Table 1). Prior to rewetting, our microbial community was dominated by Actinobacteria, Alphaproteobacteria, and Gammaproteobacteria (Sieradzki et al., 2022). Proteobacteria were abundant, highly active throughout wet-up, and responded early to rewetting. Actinobacteria also were a top-responding lineage to wet-up, but showed less overall activity despite their abundance (see Sieradzki et al., 2022 (Sieradzki et al., 2022) for more details on the microbial and function response to rewetting).

To explore spatial heterogeneity of viruses and potential hosts, we analyzed the number of unique vOTUs and MAGs detected in all three, just two, or only one of the three biological field replicates (Figure 1A). MAGs, predominantly representing bacteria, were more evenly detected (present in all three replicates) throughout soil replicates compared to vOTUs, and throughout time. In contrast, vOTUs were more likely to be detected in a single replicate than in all three, by an order of magnitude or greater (Figure 1A).

To visualize the spatial and temporal succession of microbial and viral populations, we calculated ordinations based on the Bray-Curtis dissimilarity of relative abundance of MAGs (Figure 2A left) and vOTUs (Figure 2A right). The first two ordination components of MAG abundances revealed that succession of potential hosts was most pronounced over time (52.54% variance explained) with a minor spatial component and explained in total over 70% of the variability. On the contrary, abundances of viruses had a significant spatial component, with the first two time points (3h and 24h) clustering with time 0h in two of the three plots. The first two ordination components of vOTUs represented a smaller fraction of the variance compared to potential hosts (35.6%).

Following rewetting, we found that viral richness significantly decreases, while viral biomass significantly increases. Early time points showed higher richness of vOTUs, but with lower abundances per vOTU (Figure 2B, Figure 2C). In contrast, later time points, in particular 168h, shifted to a less even viral community, with higher relative abundances of fewer vOTUs. Average

vOTU richness, as measured through the number of unique vOTUs per time point, decreased significantly after 3h (linear regression – pvalue: 0.02; rvalue: -0.53) (Figure 2B). While the number of vOTUs decreased by nearly 50% one-week post wetting, viral biomass increased by at least five-fold (Figure 2D). We found viral biomass, as proxied by calculating viral DNA through time, statistically significantly increased by the end of the time sampled (linear regression – pvalue: 2.1×10^{-3} ; rvalue: 0.68).

To further characterize viral temporal succession, we focused on the set of vOTUs that were present in at least three time points within each replicate. We termed these persistent vOTUs. Previously, our group observed that at sub-phylum taxonomic levels microbial response to wet-up is sequential with rapid, intermediate, and delayed responding groups (Placella et al., 2012). We predicted that persistent vOTUs would reflect host dynamics and show similar temporal response patterns. For each replicate-specific set of vOTUs we constructed hierarchically clustered heatmaps of relative abundance to establish vOTUs that behave similarly across wet-up (Supplementary Figure 1). This clustering demonstrated several broad patterns of abundance dynamics consistent across replicates: 1) vOTUs present early in wet-up (0h, 3h, 24h) that then disappear or drastically decrease in abundance i.e. early vOTUs; 2) vOTUs that appear for the first time 48h after wet-up, i.e. late responding vOTUs; 3) vOTUs that are present across all time points throughout the wet-up, i.e. ubiquitous vOTUs; 4) vOTUs that are present in dry soil (time 0) and at the end of the wet-up time course (168 h), but are not detected in all time points in between, i.e. 0h and 168h vOTUs. vOTUs that did not fit into these categories were included in the “other” category (Supplementary Figure 2A). Each of these defined response categories represented, at times, up to 15% of the total virome reads. Overall, we observed that viruses follow plot-specific successional patterns to wet-up.

We were curious whether late vOTUs (first detected as VLPs at 48 hours) were derived from viral infections occurring at earlier time points. We predicted that these vOTUs could be found in the metagenomes and tested this by asking if these vOTUs were inside microbial cells or harbored in host chromosomes. However, metagenome reads from time points 0h and 24h did not map to late vOTUs (Supplementary Figure 3). We were unable to trace vOTUs observed as VLPs following 48 hours from metagenomes collected prior to VLP detection.

In previous work and in our paired study, we observed that Actinobacteria and Proteobacteria dominate the microbial post wet-up response (Blazewicz et al., 2020; Placella et al., 2012; Sieradzki et al., 2022). We hypothesized that the vOTUs would reflect these host responses. While we appreciate that a phylum-level host analysis of viruses does not reflect known viral host ranges (which is often on a strain level) (Ignacio-Espinoza et al., 2020; K. E. Williamson et al., 2017), wet-up shows consistent phyla-level microbial responses and our host predictions for viruses were limited to a coarse taxonomic view. Thus, we examined the vOTU putative host composition within each taxonomic group (Supplementary Figure 2B) and found vOTUs infecting Actinobacteria dominated the groups responding immediately and after 24 hours. Alphaproteobacteria-infecting vOTUs were the second most numerous and were also captured in groups on both timescales. In contrast, the early vOTU category was equally dominated by Alphaproteobacteria- and Actinobacteria-infecting vOTUs. Generally, vOTUs with no host assignment (unknown) comprised 6 to 47% of total vOTUs, with the largest proportion found in the ubiquitous vOTU category. Overall, we did not find evidence for our hypothesis: vOTUs

sampled in each time frame were not predicted by the patterns of their hosts. Rather, vOTUs of prevalent hosts in our system (Actinobacteria and Proteobacteria) appeared as the most prevalent vOTUs in each response category.

Characterizing Active vOTUs with SIP metagenomics

After analyzing the dynamics of the total set of viruses following rewetting, we investigated the subset of viruses actively replicating at any time to evaluate the impact of viruses on their microbial host abundances. Actively replicating viruses here are defined by incorporating ^{18}O from the H_2^{18}O wet-up (Figure 1). Specifically, we applied qSIP to calculate isotopic enrichment (atom percent excess, APE) of DNA, which is a proxy for genome replication rate (Greenlon et al., 2022; Nuccio et al., 2022). For MAGs and vOTUs we constructed rank curves sorted by APE and time using taxa with APE significantly greater than zero (Figure 3). This enabled us to quantify each genome's isotope incorporation, or growth, through time and simultaneously the overall pattern per time point. 1) In the first time point, we show rank curves for genomes detected as enriched. In subsequent timepoints, the genomes from the first time point remain in the same rank position, and any new genomes appearing as enriched are presented in the same Figure to the right of the previously detected active genomes. 2) the disordering of previous time points' APE ranked curves in a time-dependent manner i.e., the MAG and vOTU APE change at a taxon-specific level through time resulting in the disruption of the descending rank pattern in later time points (Figure 3). The disordering seen in these plots represents the changes through time of a genome's enrichment. For instance, a genome first detected as highly enriched at 24 hours, may decrease in its enrichment by 72 hours and while its x-axis position remains the same throughout the time course, its APE value reflected on the y-axis thus becomes lower than its expected initial rank. Also, a genome that was previously detected as enriched, may become undetected, this is visualized in the density of a curve becoming sparser as it is missing the points that represent previously detected genomes.

For both MAGs (Figure 3A) and vOTUs (Figure 3B), we found the greatest number of newly isotope-enriched genomes at 48h. MAGs most isotopically enriched when they first appeared remain among the most enriched genomes if they are detected in later time points. Despite MAG and vOTU enrichment similarities, by 168 hours there is a higher degree of disordering amongst vOTUs than that of the MAGs. vOTUs most highly enriched at 24h and 48h ($\geq 20\%$ APE) (Sieradzki et al., 2022) in the time course appear much lower than their initial rank or are altogether undetected, and vOTUs first appearing as slightly enriched become more enriched at subsequent time points (increase in APE relative to their x-axis position) – the average change in rank from 24 hours for vOTUs is 38 and for MAGs is 9. Actinobacteria-infecting vOTUs are the most numerous active viruses across time points. However, the most highly enriched vOTUs vary through time and are predicted to infect Proteobacteria or have an unknown host (Figure 3B). Proteobacteria MAGs are also the most highly enriched microbial genomes, implying a high growth rate. Enriched archaeal viruses are detected at each timepoint with the largest number detected at 48h. Further, the ratio of isotopically-enriched viruses compared to microbes changes through time (Figure 1B, bottom panel): at 24 hours there are three times as many viruses as bacteria, at 48 hours only twice as many, and by 168 hours that number increases to 2.7 more active viruses to microbes.

Of the 177 total enriched vOTUs detected in metagenomes, 58 are found in VLP viromes, indicating that at least 33% of genomes replicating and assimilating isotope within host cells are also able to escape their host cell to be detected as putative virions.

Viral lifestyle characterization

To further assess how viruses may control host populations during soil microbiome assembly we investigated temperate capabilities of vOTUs. To test our hypothesis that wet-up induces temperate phages we searched for vOTU genomes common to viromes and metagenomes that contained bacterial flanking regions in metagenomes to discover and quantify lysogeny in our system. This strategy yielded just two candidates of vOTUs found both as VLPs in the virome and as putative prophage in metagenome contigs and only one of these candidates was successfully binned into a MAG. The unbinned sequence was predicted to be an Alphaproteobacteria, *Rhizobium leguminosarum*. The MAG identified as having a prophage was a Gammaproteobacteria of the genus *Pantoea*. In this study we did not further characterize this binned putative prophage since it was only detected in one of the three plots at two non-consecutive time points (0 and 24h).

Unsuccessful in finding vOTUs appearing as both integrated in MAGs or bacterial sequence and as VLPs, we examined our system for temperate phage by analyzing the genomic capacity of vOTUs to integrate, via an integrase gene, into a host chromosome (Figure 4). Of the total vOTUs detected 16.9 % (4,463) contained an integrase gene and were denoted as putative prophages. Concerned that vOTU contigs may be partial viral genomes left fragmented during assembly, we split the results of integrase-encoding and -absent viral genomes by vOTUs that circularize, implying genomic completeness, and vOTUs that are linear, and thus may or may not be complete. We reasoned that if the linear set of vOTUs followed similar trends to complete, circularized vOTUs, then the patterns displayed would be representative of biological trends rather than artifacts of incomplete assembly of linear vOTUs.

For both linear and circular vOTUs, non-integrase encoding vOTUs represented a greater relative abundance of the total virome community, more than twice as much at most time points (Figure 4A). Further, up to five times more circular and linear vOTUs did not contain compared to containing an integrase gene (Figure 4B). Integrase-containing vOTUs showed the same trend in relative abundance over time in both linear and circular vOTUs: a steady decline in relative abundance in the first 48 hours and then relatively constant abundances through 168 hours (Figure 4A). There does not appear to be an increase in integrase-encoding vOTUs following wet-up. The richness data underscore this. Both the number of integrase-encoding vOTUs and non-integrase containing vOTUs are consistent whether looking at linear or circular vOTUs. The non-integrase encoding vOTUs are not only much more numerous, but also, they reflect the pattern of total dereplicated vOTUs through time (Figure 2B) – a decrease in count through time.

We also tested whether vOTUs appearing after wet-up (0 h) were enriched in integrase genes, with the logic that vOTUs appearing for the first time following wet-up may be prophage excising from genomes. The post-wet-up vOTUs were not significantly enriched in integrase genes compared to vOTUs in the dry soil (Chi-square test – statistic: 1.6; pvalue: 0.2).

To further support the inference that wet-up does not induce temperate phages, we took a MAG-centric approach. We used bioinformatic tools that can predict prophages in the context of bacterial genomes, here MAGs, to discern whether, how many (counts), and how prevalent (abundant) prophages are in this system (Supplementary Figure 4). We looked at both aggregated relative abundance across all time points of MAGs by phylum and count of total MAGs with or without prophages. Most phyla contained a predicted prophage, but prophage-containing MAGs for each phylum were both fewer and less abundant than MAGs without detected prophages. For lineages other than Actinobacteria and Alphaproteobacteria, just one to six MAGs per phyla contained prophages. The aggregated relative abundance data underscores this point as the abundance of MAGs containing prophages (Supplementary Figure 4a) was lower than the proportion of the number of MAGs containing prophages (Supplementary Figure 4b). MAGs containing prophages comprised 15-26% of the total abundance of Actinobacteria and Alphaproteobacteria per time point. Based on coverage differences between the prophage sequence and the flanking host genomic regions (Kieft & Anantharaman, 2022) we identified only three prophages that seemed to be actively replicating. One of these prophages appeared active at all time points in at least two out of the three replicates, and was integrated into a MAG of Alphaproteobacteria genus *Microvirga*. The other two prophages were active only in a single replicate and in either one or two time points, and were integrated into genomes of Actinobacteria family Gaiellaceae and Alphaproteobacteria family Xanthobacteraceae, respectively.

To further investigate viral lifestyle and assess whether VLPs correspond to lysis we looked for evidence for non-lytic viruses i.e. inoviridae in the studied soils. We performed an HMM search using the pI-like ATPase protein family HMM, which is conserved in filamentous inoviruses (Roux et al., 2019), and found no evidence of inoviridae in these data.

Virus-host dynamics

We next examined the dynamics of viruses in a host-dependent manner to test whether and to what extent viruses may serve as a top-down control during community assembly. Given the prevalence of MAGs and vOTUs with a taxonomy or host-level taxonomy assignment of Actinobacteria and Proteobacteria, and given the previous high degrees of mortality and growth following wet-up calculated for bacteria of these phyla (Blazewicz et al., 2020), we focused analyses comparing vOTUs and MAGs to these lineages. Specifically, within Proteobacteria we focused on Alpha- and Gamma-proteobacteria as growth could not be detected for other classes within this phylum in our experiment. We analyzed APE for Actinobacteria, Alphaproteobacteria, and Gammaproteobacteria MAGs and found that individual MAGs mostly show a similar trend of increasing in enrichment in the first 48 hours (Supplementary Figure 5A). The Proteobacteria MAGs show a higher overall degree of enrichment for individual MAGs than the Actinobacteria (Supplementary Figure 5A). Similar to the temporal pattern of enrichment, the number of enriched MAGs increased over the first 48 hours, decreased through 72 hours, and mostly stabilized until the end of the time sampled for Actinobacteria and both Proteobacteria lineages (Supplementary Figure 5D). For both phyla, there were higher counts of enriched vOTUs than enriched MAGs. At 24 hours, for instance, there were more than twice as many vOTUs as MAGs for all lineages, and for Actinobacteria in particular, over fifteen times as many vOTUs were detected as enriched compared to MAGs.

Viral-OTUs, overall, showed a similar pattern of enrichment to that of MAGs: increasing over the course of the first 48 hours, decreasing in APE through 72 hours, and stabilizing or slightly increasing their APE by 168 hours (Supplementary Figure 5B). For putative Actinobacteria-infecting phages, the majority of vOTUs clustered into a low activity region (less than 20% APE (Sieradzki et al., 2022)) which is composed of two subgroups. A separate cluster of higher activity vOTUs (36% of the total) showed APE values consistently above 20 %, but did not follow clear trends through time. Alphaproteobacteria- and Gammaproteobacteria-infecting viruses showed similar trends in APE through time that appear as denser regions in their vOTU median APE plots.

To understand phylum-level dynamics between vOTUs and MAGs we examined the ratio of vOTU and MAG median APE values (Supplementary Figure 5C). Actinobacteria vOTUs had the highest ratio APE relative to hosts through time as compared to the Proteobacteria lineages, and only Actinobacteria viral enrichment was greater than MAG enrichment (ratio >1) observed in the first 48 hours. For all lineages, vOTU enrichment outpaces MAG enrichment (observed as positive slopes in Supplementary Figure 5C) between 24 and 48 hours, and again at 72 hours for Actinobacteria and Alphaproteobacteria.

The dynamics of MAGs and vOTUs are also reflected in the counts of active genomes detected through time (Supplementary Figure 5D). Viruses and MAGs showed similar temporal trends for isotopic enrichment (APE) and active genome richness (number of APE-detected genomes). We similarly analyzed other phyla using these metrics of MAG APE, vOTU APE, number of enriched genomes through time and found no discernable patterns.

Quantifying viral contribution to microbial mortality

We approximated the viral contribution to mortality following wet-up using VLP sequence data from this work and bacterial mortality estimates from corresponding SIP-density fractionated amplicon sequencing (Sieradzki et al., 2022). We modeled a range of viral burst sizes from one virus released per cell lysed (Kuzakov & Mason-Jones, 2018; K. E. Williamson et al., 2008) to 1000 viruses released per cell lysed (Sui et al., 2021). Taken cumulatively, looking at viral lysis of microbes at 3-, 24-, and 48-hours following rewetting, viruses account for minimally 0.03% (burst size of 1000) and maximally 29.3% (burst size of 1) of the microbial death observed (Figure 5B). At 168 hours cumulative viral contribution to microbial death ranges from 0.05 % to 46.6 % depending on estimated burst size. We observed at 24 hours the highest per day viral contribution to microbial mortality 0.01 % to 17.4 %.

3.3 Discussion

The combination of viromes and SIP enabled identification of pools of viruses and their dynamics during the community restructuring after soil was rewet after a long dry summer period. We 1) assessed viral biomass and richness following rewetting and posited that dry soil may serve as a seedbank of virions for the growing season; 2) analyzed dynamics of active viruses and found these more closely follow host populations rather than control them; 3) investigated viral life cycle and observed minimal evidence for lysogeny suggesting lytic lifestyle is predominant in soil viruses; 4) estimated viral contribution to microbial mortality to discover viral contribution to

steadily increases following rewetting; 5) examined viral succession and found that spatial heterogeneity drives temporal trajectories.

Dry soil serves as rich seedbank of viruses of which active viruses follow rather than control hosts

We observed a large decline in viral richness, and by extension diversity, in the first 24 hours following rewetting, which was inversely related to an increase in viral biomass (Figure 5A). This pattern was consistent across all soil replicates making it detectable despite the high degree of inter-sample heterogeneity. We thus observed that only a subset of viruses present in seasonally dry soil thrive following rewetting. Intriguingly, analysis of the total viral pool present after wet-up did not reveal patterns of succession correspondent with known host response previously reported or observed in our study (Supplementary Figure 2-3) (Blazewicz et al., 2020; Placella et al., 2012; Sieradzki et al., 2022). Shifting focus to the active subset of microbial and viral populations (Figure 3, Supplementary Figure 5) revealed vOTUs to behave more similarly to their active hosts based on the number, types, and degree of isotopic enrichment of genomes through time.

Marine system studies have shown that viruses follow their hosts' dynamics (Chow et al., 2014) and that high local diversity of viruses may be drawn from a seedbank of nearby viruses (Brum et al., 2015). Given the richness of viruses in dry soil, we hypothesize that dry soil may be a reservoir, or seedbank, of diverse viruses that become dynamic following rewetting. As suggested in Chow et al. 2014 (Chow et al., 2014), our isotope targeted results suggest that viral composition follows host dynamics, as opposed to viruses directly controlling host abundance. That is, the total viral pool does not necessarily mirror the total microbial pool, and therefore it is important to focus on the active subset of vOTUs and MAGs to appreciate host-virus dynamics in our soil system.

Our isotopic data revealed a broad diversity of active viruses following rewetting which represent only a subset of the detected viruses in the dry soil seedbank. To investigate viruses that could be important controllers of microbial populations, we categorized the active viruses based on their predicted host phyla. First, focusing on viruses that infect Actinobacteria, we found a range of virus activities. A subset of Actinobacteria viral populations appear to contain a small portion of highly active taxa (>20 % APE), while the majority of Actinobacteria-infecting viruses have a lower level of steady activity (Supplementary Figure 5B). We suspect this small highly active pool of viruses infects the similarly small pool of hosts that fall within the higher activity category. We find similarly structured subsets of varying activity in Alphaproteobacteria, Gammaproteobacteria and the viruses that infect them underscoring the diversity of viral and host activity and response. Across this diversity in activity, we found that viral APE increased faster than microbial APE in the first 48 hours suggesting that these infections happened faster than microbial genome replication (Supplementary Figure 5C). Previously, our group showed in a carbon-SIP experiment that phages can be the most enriched genomes in a soil system (Starr et al., 2021). We find here too, Actinobacteria- and Proteobacteria-infecting phages to be the most isotopically enriched viruses that we could identify following wet-up. That is, we can observe in real time, phage replication outpacing bacterial growth in a time- and taxonomic-dependent manner.

We predicted that we would see evidence of virus control on microbial host populations in coupled, offset predator-prey dynamics, but our data suggest that viruses reflect host community dynamics

rather than serve as a top-down control on microbial community composition. The relationship of APE through time did not reveal the expected canonical predator-prey paired cycles. Rather, we see that viral enrichment and counts per phylum broadly follow the same trends through time as their bacterial counterparts. Future studies of this system may choose to sample more finely resolved time points (Van Goethem et al., 2019) to see if predator-prey oscillations may exist more clearly, for example, at 72 hours when we see both virus and host enrichment decrease. However, predator-prey oscillations may only be detected at the species or strain level (Ignacio-Espinoza et al., 2020).

Viruses exhibit successional patterns in response to perturbation

Remarkably, we found a consistent successional pattern of our total detected viruses across all plots. However, these successional patterns appeared decoupled from their hosts. Clusters of vOTUs responded similarly across plots (Supplementary Figure 2; Supplementary Figure 3A) – vOTUs found early in wet-up, late-responding vOTUs, vOTUs present throughout the entire time course, and vOTUs that were detected at the start and end of our sampling but not consistently throughout. Within these response groups we observed plot-specific fluctuations and found that the late category of viral response only applies to two of three plots. However, separating viruses into response categories of early, late, etc. echoed the microbial response to rewetting, which shows taxon-specific growth and mortality patterns (Blazewicz et al., 2020; Placella et al., 2012). We thus hypothesized that these viral response categories would align with previously observed phylogenetic trends for rapid, intermediate, and delayed microbial, i.e., putative host, responses to resuscitation. However, at the phylum level, we did not observe host-specific trends in viral response groups (Supplementary 2B). Rather, host signal within each response category showed the same trends – viruses of Actinobacteria and Alphaproteobacteria hosts responded early, late, and were ubiquitous across wet-up.

While a clear host signal associated with different viral response groups was not apparent, we found that of Betaproteobacteria-infecting viruses, the mode falls in the early responding category, which is also when Betaproteobacteria have previously been reported to have the highest mortality rate (Blazewicz et al., 2020). These results suggest that mortality of Betaproteobacteria immediately following wet-up is driven, in part, by viral lysis. Although it is unclear whether virions detected early in our time course represent recent host infections or whether these virions have persisted from the dry season, our lack of persistent vOTUs suggests virions detected may indeed derive from recent host infections. Our study provides a first glimpse into fine spatiotemporal viral dynamics in soil and enabled us to categorize distinct viral responses to rewetting.

Given the higher spatial homogeneity in microbial distribution and response to rewetting compared to viral distribution (majority of MAGs appear in all samples), it is tempting to hypothesize that this indicates microbes in soil are widely susceptible to infection by many distinct viral populations. However, the lack of specific host-virus linkages in our system meant we were unable to investigate viral host breadth.

Spatial heterogeneity drives viral population dynamics

Our data show spatial heterogeneity is an important feature in understanding the response of viruses to rewetting. While a rapid decline in viral diversity, increase in viral biomass, and successional patterns emerged as trends across the studied soil plots, the majority (59%) of viruses appeared in fewer than three of our total 18 viromes and most do not appear in all biological replicates (Figure 1A). To account for the high degree of spatial heterogeneity we conducted our analyses within each plot as opposed to across all plots which allowed us to uncover emerging trends across space that averages would have precluded. However, comparing viruses to their hosts at this spatial scale posed a challenge because microbes had a more coherent and uniform response than viruses to wet-up. Further, the patchiness of viruses precluded following subsets of viruses through space and time to find statistically meaningful trends. The heterogeneity of soil viruses suggests that infection may not be as uniform across space as potential hosts are. Perhaps, this is the effect of rewetting: virions persisting in soil, more often than not, may be decoupled from their host dynamics, and the addition of water may exacerbate the hurdle of host-viral physical interactions as previous research has predicted increased water potential is associated with lower rates of direct cell-cell contact in soil (Tecon et al., 2018). On the other hand, wet-up may serve to increase connectivity between soil microsites and relieve dispersal limitations of viruses, supporting host connections. As research has shown in water-undersaturated soils, dispersal is a constraint (Jiao et al., 2020). Our observed increase of viral biomass following wet-up suggests water would support physical interactions between viruses and their hosts.

The viral heterogeneity observed throughout our field site (Santos-Medellín et al., 2022) also underscores our own assumptions in the design of this experiment: that the field plot layout, constructed with our knowledge of microbial distribution across a soil site would enable similar trends for viruses. Rather, our fine-time resolved study and concurrent work studying heterogeneity on a field scale at longer time intervals (Santos-Medellín et al., 2022) have compelled us to question what scale of sampling for viruses represents biological replicates. Our results suggest that future studies should consider the distinct scales of microbial and viral populations in soil to have a holistic understanding of microbial communities (Wilpiseski et al., 2019).

We found that each soil plot contained a unique standing community including a seedbank of virions. These viral populations may degrade or thrive following rewetting, but after one week post perturbation we did not observe that viral composition across plots became more similar. Rather, the virions present at the end of the dry season determined the trajectory of the plot-specific viral composition. That is, at this timescale, wet-up is not a perturbation at which we observed a convergence of viral populations across space despite host uniformity at the same sampling (Figure 2).

Scant evidence for lysogeny suggests other lifestyles are dominant

Critical to understanding how viruses impact the soil microbiome during this rapid moment of community response is understanding the mechanism of how viruses live and survive in soil – an analysis of viral life cycle. An emerging idea in soil viral ecology suggests that lysogens are prevalent in soil microbes, and further, lysogeny is a favorable lifestyle for viruses both to remain in association with a host and to weather conditions difficult for growth i.e. dry soil (Emerson, 2019; Howard-Varona et al., 2017; Jansson & Hofmockel, 2020; Kimura et al., 2008; Kuzyakov

& Mason-Jones, 2018; Pratama & van Elsas, 2018; Sokol et al., 2022; K. E. Williamson et al., 2017; Wilpiseski et al., 2019). We tested this hypothesis looking specifically at wet-up as an environmental perturbation that could induce prophage and found no evidence to advance this hypothesis.

Our data suggests that vOTUs, which cannot recombine into a host chromosome through an integrase, dominate this system. We found no clear evidence showing wet-up induces a significant number of prophages. The viruses that do increase in abundance following the first 24 hours are predominantly non-integrase encoding. We hypothesized that late-responding vOTUs may also provide a window into viral life strategies (Supplementary Figure 2A) – that examining virome late-responders in metagenomes would reveal these viruses were cell-associated prior to appearance as VLPs. As late-responding viruses did not appear in metagenomes at earlier time points at higher abundances than when they were detected in viromes (Supplementary Figure 3), and prevalent MAGs did not maintain detectable prophages (Supplementary Figure 4), we conclude that temperate phage did not dominate this soil system and they were not the predominant responding viruses to wet-up.

Whether all detected VLPs in our system signify cell lysis remains unclear. While we did not find evidence of inoviridae – viruses that can replicate and release virions without lysing their hosts through a chronic infection (Roux et al., 2019), other non-lytic lifestyles may be possible. Further, while neither the capacity to integrate into a host chromosome nor persistence as a prophage is prevalent in our wet-up system, we were not able to rule out other mechanisms of viral maintenance in a host cell when a virus is not lytically programmed i.e. pseudolysogeny (Miller & Day, 2009). Perhaps evidencing the prevalence of pseudolysogeny or some other cell-associated state of soil viruses, the majority (67%) of viral populations detected as active were only found in metagenomes (direct DNA extraction from soil) rather than viromes. For a virus to be detected as enriched it must be infecting host cells to replicate, thus assimilating isotope – most of these active viruses may not be detected as putative virions because of their host-associated state.

Viral contribution to microbial mortality increases though time post wet-up

Microbial death is critical to the flow of nutrients through biomes (Sokol et al., 2022). A significant feature of soil wet-up is high rates of microbial mortality (Barnard et al., 2020; Blazewicz et al., 2014, 2020). Our previous work demonstrated rapid bacterial mortality following wet-up, with 25% of bacteria degrading in the first 3 hours and continued mortality observed through 72 hours (Blazewicz et al., 2014). We hypothesized that the initial burst of mortality in the first 3 hours was driven by osmotic lysis due to the rapid change in osmotic pressure, while mortality after 3 hours was driven by viruses or other biological factors (Barnard et al., 2020). Others have also speculated that viruses play a role in observed mortality following rewetting (Naylor et al., 2020). We found that viruses play a minor role in mortality in the first 3 hours, when osmotic lysis is likely the dominant mechanism, and the contribution to mortality by viruses increases through time, more than doubling after the first 24 hours (Figure 5B). We estimate that viruses can drive a sizeable portion of microbial mortality in soil – by 168 hours viral-mediated mortality may contribute 0.05 % to 46.6 % of total microbial death dependent on burst size estimated. While our evidence suggests viruses play an important role in microbial mortality following wet-up, we infer that this influence is minimized as wet-up led to a rapid decline of viral particle abundances. If viral

biomass continues to increase following 168 hours post wet-up, we posit that the viral contribution to microbial mortality may be higher, and perhaps contribute to microbial mortality on the scale of what is observed in marine systems – lysing up to 20 – 40 % of bacteria daily (Jover et al., 2014). There may also be other circumstances in the soil microbiome where viruses play a larger role in mediating microbial death, perhaps in the rhizosphere (Starr et al., 2021). Further studies are required to understand a more representative genome length of soil viruses, range of viral burst sizes in soil, and how long VLPs persist in soil to understand whether virions represent recent infections or whether these calculations require SIP techniques to capture an active fraction of viruses. Moreover, viral predation has impacts beyond contribution to mortality broadly. Even low viral turnover rates of key lineages may profoundly impact strain selection and functional potential of soil microbiomes.

Current challenges in the state of the field – linking host and viruses

Using mixed methods and multiple databases we were able to determine at least a domain or phylum-level taxonomy for all but 443 vOTUs (marked “unknown” host). Despite having some degree of host taxonomy resolution for an overwhelming majority of vOTUs detected in our system, we found few CRISPR spacer matches (Burstein et al., 2016) and thus were only able to directly link vOTUs to specific MAGs with high confidence in a few instances. We also used a network-based analysis to connect MAGs to vOTUs modelling canonical predator-prey dynamics, but were unsuccessful in ground truthing these linkages (data not shown). While our methods enabled us to ascertain a host taxonomy for vOTUs and to achieve results that aligned with microbial dynamics on a phylum-level, the current state of host-viral linkage techniques did not allow us to directly implicate viruses as preying on specific MAGs and track predator-prey dynamics on a strain level. However, our host predictions revealed that prevalent viruses in our system infected Actinobacteria and Proteobacteria, which reflected our microbial data: the most prevalent viruses infected the most prevalent hosts.

Others studies have connected viruses to their hosts by combining SIP techniques using specific isotopically-labelled substrates and CRISPR matches which allowed linkages of methanotrophic bacteria to their phages (Lee et al., 2021). Further, virus-host linkages may be missed due to spatial scaling issues discussed here which metagenomic studies on soil aggregates or aggregates may reconcile (Wilpiseski et al., 2019).

Conclusion

We probed the wet-up of a seasonally dry soil to understand the role of viruses in microbiome assembly. We found that dry soil is a low biomass yet highly rich seedbank of unique virions, and that resuscitation of the microbiome yields a bloom in biomass of a subset of viruses. Succession of the total pool of viruses is driven by spatial heterogeneity and nearly 50% of viral richness is lost in the first 24 hours following rewetting. We observed that, on a phylum level, active viral populations appear to follow temporal trends of their microbial hosts rather than serving as top-down controls. Further, contrary to the current paradigm, our data show that lysogeny is not as prevalent in our system as previously hypothesized underscoring that the reservoir of soil viruses may be as virions rather than as inserts in microbial genomes. Our combined use of viromics, metagenomics, and isotopes facilitated estimation of the viral contribution to microbial mortality

following rewetting. While our conservative calculations indicate viruses may not be the primary mechanism of observed mortality during microbiome assembly following rewetting, viruses appear to play an important and dynamic role in the microbial response following wet-up.

3.4 Materials and Methods

Field sample collection

Topsoil samples (0-15 cm, roughly 0.5 m³) from replicate field plots were collected from the Hopland Research and Extension Center (HREC) in Northern California, which is unceded land of the Shóqowa and Hopland People, on August 28th, 2018 (only 100% precipitation plots were used for this study, see our paired study Sieradzki et al. 2022 (Sieradzki et al., 2022) for full detail). During the growing season the same field plots were covered by mixed grassland flora with *Avena sp.* (wild oat) as the most abundant plant. Soil was collected before the first rainfall event of the season and average gravimetric soil moisture was 3%. The dry soil was transferred to Lawrence Livermore National Laboratory (LLNL) where it was homogenized and sieved (2mm) to remove large rocks and roots.

Wet-up experiment

5 g sieved soil from each of 3 field replicates was weighed into separate 15 mL Nalgene flat bottom vials. Soil was wetted to 22% average gravimetric soil moisture with either H₂¹⁸O (98 at% ¹⁸O) or H₂¹⁶O (natural abundance SIP controls) (Figure 1), after which the vials were sealed inside 500 mL mason jars and incubated at room temperature in the dark. Three incubation treatments were established, each in triplicate: 1) H₂¹⁶O amended (used for DNA extraction, shotgun sequencing, unfractionated DNA metagenomics and as a fractionated DNA control for qSIP); 2) watered with H₂¹⁸O for DNA extraction, density centrifugation, shotgun sequencing, and metagenomics (isotopically-labeled fractionated DNA for qSIP); 3) watered with H₂¹⁶O for viral-like particle (VLP) suspension, concentration, DNase treatment followed by DNA extraction and shotgun sequencing to generate viromes (unfractionated DNA for viromics). Parallel jars were destructively harvested at 0, 3, 24, 48, 72, and 168 hours following the water addition; at each harvest, soil vials were frozen in liquid nitrogen and then stored at -80 °C.

Viral metagenomes (“viromes”)

Soil viromes were generated as described in Nicolas et al. 2021 (Nicolas et al., 2021) to assess viruses detected as VLPs i.e. extracellularized. Briefly, two 5-gram soil microcosms from each of three field replicates watered with H₂¹⁶O were combined for 10 g soil virome extractions. Soil was buffered with a potassium citrate solution and the effluent was passed through a 0.2-μm filter, concentrated, and treated with Dnase to remove extracellular DNA. Prior to DNA extraction using a phenol:chloroform strategy, we used an iron flocculation step to concentrate 0.2-μm filtered and concentrated samples as described in Nicolas et al. 2021 (Nicolas et al., 2021). DNA was sequenced on an Illumina Novaseq (2x150 cycles) (UC Davis). Sequencing was performed to attain an average 40 Gb per sample. Illumina adapters and phiX were removed from resulting sequences using BBTools (<https://jgi.doe.gov/data-and-tools/bbtools/>) and then quality trimmed using Sickel (Joshi & Fass, 2011). Sequences were assembled with MEGAHIT version 1.2.9 (D. Li et al., 2015); we co-assembled all biological replicates per time point using default parameters. Reading frames were predicted for each contig using Prodigal (Hyatt et al., 2010) and were

annotated using sequence similarity searches with USEARCH (Edgar, 2010) to KEGG (Ogata et al., 1999), UniRef (Suzek et al., 2007), and UniProt (Magrane & UniProt Consortium, 2011).

Metagenome assembled genomes (MAGs) from viromes

To generate microbial bins from the size-fractionated metagenomes, i.e. viromes, we used multiple binning methods. Contigs were binned using MaxBin (Y.-W. Wu et al., 2014), MetaBAT (Kang et al., 2015), CONCOCT (Alneberg et al., 2014), and manually based on coverage, GC content, and taxonomy using in house tools on ggKbase (<https://ggkbase.berkeley.edu>). DASTool (Sieber et al., 2018) was used to predict best bins and bins were further manually curated on ggKbase to remove highly divergent sequences based on coverage, GC, and taxonomy. MAGs from viromes were clustered and dereplicated alongside MAGs from metagenomes (discussed below) using dRep (Olm et al., 2017) requiring at least 75% completeness and no more than 10% contamination, yielding 39 genomes from candidate phyla radiation bacteria and archaea.

Viral prediction and generation of viral operational taxonomic unit (vOTU) set

To establish a set of viral genomes from our metagenomic efforts we used multiple methods of prediction and dereplicated vOTUs across samples. Viruses were predicted from metagenome and virome contigs. Viruses were predicted using: virstorter1 (Roux et al., 2015) (all categories), virstorter2 (Guo et al., 2021) (dsDNA and ssDNA only), VIBRANT (Kieft et al., 2020), DeepVirFinder (Ren et al., 2020) (score ≥ 0.9 ; pvalue ≤ 0.5), and Seeker (Auslander et al., 2020). From the set of virome and metagenome contigs explored 3,979,792 were predicted to be viral. Of these, 777,725 contigs were designated viral by more than one tool, and 32,660 of this subset were either circularized contigs by VRCA (Crits-Christoph et al., 2016) (i.e. implied to be a complete genome, or $\geq 10,000$ base pairs long). This set of size-filtered, circularized, more-than-once predicted viruses was dereplicated using MMseqs (Hauser et al., 2016) (95% ANI, 85% breadth required of sequences) to 26,368 viral operational taxonomic units (vOTUs). All vOTUs were annotated using DRAM-v (Shaffer et al., 2020), and specifically searched for integrase genes.

DNA density fractionation

DNA was extracted from three biological replicates at each time point using a modified phenol-chloroform protocol adapted from Barnard et al. 2015 (Barnard et al., 2015) as described in (Sieradzki et al., 2022). For density gradient SIP processing, 5 μg of DNA in 150 μL 1xTE was mixed with 1.00 mL gradient buffer, and 4.60 mL CsCl stock (1.885 g mL^{-1}) with a final average density of 1.730 g mL^{-1} as described in Sieradzki et al., 2022. The fractions for each sample were binned into 5 groups based on density (1.6400-1.7039 g/mL , 1.7040-1.7169 g/mL , 1.7170-1.7299 g/mL , 1.7300-1.7449 g/mL , 1.450-1.7800 g/mL), and fractions within a binned group were combined and sequenced. Unfractionated DNA (100 ng) from each soil sample was also saved for shotgun sequencing.

All sequencing was performed on an Illumina Novaseq 2x150 cycles. Pre-aliquoted unfractionated DNA from each sample was sequenced to an average of 20 Gb per sample, and DNA from each density fraction was sequenced to an average of 10 Gb per fraction, yielding a total of 227 metagenomes. Illumina adapters and phiX were removed using BBTools and reads were quality trimmed using Sickle using default parameters.

Metagenome assembly, annotation, and binning

To establish MAGs from metagenomes we approached assembly in two ways: 1) co-assembly of unfractionated triplicates for each time point 2) co-assembly of the 5 density fractions from each time point and each replicate. We used MEGAHIT version 1.2.9 for assemblies (D. Li et al., 2015) details described in (Sieradzki et al., 2022). MetaWRAP version 1.3.2 (Uritskiy et al., 2018) was used to map reads from unfractionated metagenomes to contigs from each assembly, bin each assembly with MetaBAT2 (Kang et al., 2015) and MaxBin2 (Y.-W. Wu et al., 2014) and refine bins. Genomic bins from all samples as well as bins from soil metagenomes collected at the same site in the preceding winter (Greenlon et al., 2022) were then dereplicated at 95% average nucleotide identity (ANI) using dRep version 3.0.1 (Olm et al., 2017) to create a final collection of high quality 503 metagenome assembled genomes (MAGs) that were on average 87% complete and 5% redundant. Taxonomy was assigned to bins using GTDB-tk version 1.5.1 (Chaumeil et al., 2020).

To calculate abundance measurements for MAGs, reads from unfractionated samples were mapped back to the MAG collection using bbmap with minid=0.98 (<https://jgi.doe.gov/data-and-tools/bbtools/>). Mapping statistics were then used to calculate coverage at reads per kilobase million (RPKM), mapped read counts and breadth of coverage using CoverM (<https://github.com/wwood/CoverM>), and were normalized to sequencing depth.

Calculating relative abundances of the vOTU collection

To quantify abundance of vOTUs in viromes and metagenomes and to assess activity via qSIP we mapped reads from these sequencing efforts to our vOTUs. All reads from viromes, metagenomes, and SIP-fractionated metagenomes were mapped to the set of vOTUs using bbwrap.sh at 98% minimum identity. Samtools (H. Li et al., 2009) sort and index functions were applied to resulting bam files. We used CoverM (<https://github.com/wwood/CoverM>) to calculate breadth and read count per vOTU per sample. Read counts were divided by total reads per sample to calculate relative abundance. For virome relative abundance we considered vOTUs present in a sample if they had at least 80% breadth covered (fraction of genome covered by reads). For use with the qSIP pipeline, we found that few vOTUs had reads that mapped stringently (98% identity) and in all three replicates of multiple density fractions in both ^{18}O and ^{16}O conditions and therefore chose a lower breadth cutoff, 50%, for detection of vOTUs in SIP fractions (described below).

Ordination plots

To understand compositional differences through time responding to rewetting we constructed ordination plots. For MAGs we constructed a matrix of metagenome relative abundance per sample per genome. Similarly for vOTUs, we generated a matrix of relative abundance per sample per genome, using vOTU relative abundance as calculated from our virome samples. We used the python scikit package (Virtanen et al., 2020) to calculate Bray-Curtis dissimilarity and to perform principal coordinates analysis (PcoA) to visualize the dissimilarity between samples.

Virus-host linkage and assigning viruses a host taxonomy

To link vOTUs to potential host microbial MAGs, we used multiple approaches, starting with CRISPR spacer searches. We predicted CRISPR arrays using minced (<https://github.com/ctSkennerton/minced>) (Bland et al., 2007) in the original set of MAGs (not dereplicated) from this study, in MAGs from another qSIP study conducted in a soil from our HREC field site (Greenlon et al., 2022), and in our viromes. 84 MAGs and 32 vOTUs contained

CRISPR arrays with at least 1 spacer. We aligned the MAG spacers against the vOTUs using BLASTn (Altschul et al., 1990), and vice versa, and screened for hits with no more than one mismatch and no gaps between the full spacer and the vOTU. By these criteria, only one vOTU spacer matched a MAG in our dataset. This vOTU appeared in two of three plots and in two time points for each plot. In contrast, 25 MAG spacers from 6 MAGs matched 19 vOTUs. From these linkages we assigned host taxonomies to vOTUs using the associated MAGs GTDB taxonomy (Chaumeil et al., 2020).

To assign host taxonomy to additional vOTUs we combined these CRISPR spacer searches in MAGs, with a broader set of CRISPR spacer matches, and contig taxonomy predictions. Using Kaiju (Menzel et al., 2016), we assigned a consensus host taxonomy per vOTU. In addition to MAG and vOTU spacers, we also predicted spacers in unbinned metagenome sequences using minced and established contig taxonomy for these unbinned sequences using Kaiju. We used the same BLAST parameters to match these unbinned CRISPR array spacers to our dereplicated vOTU set. Next, we queried the comprehensive spacer database generated in Shmakov et al. 2017 (Shmakov et al., 2017) against our vOTU sequences using the same BLAST parameters. Each spacer from this database (Shmakov et al., 2017) included an NCBI taxonomy that was assigned to matching vOTU protospacers.

We also used vOTU sequences to inform host taxonomy. Previously, Al-Shayeb et al. 2020 and others assigned host phyla to phages based on whether ORF annotations in phages were predominantly of one phylum (Al-Shayeb et al., 2020). Here, we used Kaiju on vOTUs to assess taxonomy. Most vOTUs were assigned to microbial taxonomies that presumably reflect shared host-virus genome/gene content. We used this microbial taxonomic assignment to set a putative host. For vOTUs predicted to have viral taxonomies we captured host lineages using the Virus-Host Database (Mihara et al., 2016).

Lastly, we created a database of all vOTUs, their host prediction, and the source of the host prediction – CRISPR spacer-based taxonomy from GTDB for MAGs, unbinned spacers with Kaiju taxonomic predictions, NCBI taxonomies from spacers matched from the spacer database used, or Kaiju taxonomy predictions per vOTU. We used this database to determine the consensus taxonomy between these sources closest to a species level by searching for the greatest common denominator. This allowed us to find putative hosts for 93.1% of total vOTUs at the domain level, and 92.8% at the phylum level.

Examination of evidence for viral induction

To quantify viruses in our system that were present both as prophages in MAGs and as VLPs in the virome we mapped virome reads to metagenomes and looked for alignment between metagenome sequences and vOTUs. Specifically, we mapped all virome reads to metagenome assembled sequences at 90% average nucleotide identity. We used coverM (<https://github.com/wwood/CoverM>) to generate breadth and coverage calculations for mapped reads and identify metagenome contigs that had virome reads mapping with at least 1000 bp covered by reads, but not mapping with 100% breadth. We believed this implied that reads mapped to a potential lysogenic viral genome embedded in the host genome. Using BLASTn (Altschul et al., 1990), we aligned the set of metagenome contigs that met these criteria against our vOTUs with a minimum E-value of 10^{-100} . We subsetted the resulting matches to vOTUs that matched a

metagenome contig with greater than 90% nucleotide identity, aligned with a metagenome sequence that had more than 10 Kbp of flanking region around the alignment, and an aligned section of at least 10 Kbp. This resulted in 11 vOTUs representing candidate lysogenic viruses. Using ggKbase (<https://ggkbase.berkeley.edu/>), we manually analyzed each of these vOTUs and their corresponding metagenome contigs for aligned and unaligned regions consistent with phage (enriched for hypothetical genes, viral structural genes, integrase genes). This resulted in two sequences that appeared to be lysogenic: found both as vOTUs in viromes and in metagenomes in the context of bacterial sequences.

Assessment of prophages in MAGs

Prophages were identified from our set of MAGs by virSorter version 1 categories 4 and 5 (Roux et al., 2015), or by VIBRANT (Kieft et al., 2020). Active prophages, i.e. prophages assessed as replicating more than their flanking host genome, were identified by running PropagAtE (Kieft & Anantharaman, 2022) with metagenome read mapping files to MAGs previously generated. PropagAtE relies on differential coverage between prophage and adjacent host regions.

Viral response categories

To establish viral response patterns through time, we established the set of 6,840 vOTUs present in a specific plot for three or more time points; we refer to these as “persistent vOTUs.” Persistent vOTUs represented an average of 11% of the total virome reads through time. We generated hierarchically clustered heatmaps (cluster maps) per plot in seaborn (Waskom, 2021) using relative abundances per time point for each vOTU. With cluster map observations and k-means clustering, we established 5 response categories based on vOTUs presence or absence in given time points: 1) ‘early vOTUs’ present at 0h, 3h, or 24h post wet-up; 2) ‘late responding vOTUs’ present at 48h, 72h, or 168h post wet-up; 3) ‘ubiquitous vOTUs’ present in all time points; 4) ‘vOTUs present at 0h and 168h’ and one other time point; 5) ‘other’ vOTUs that did not fall into these categories. We summed vOTU relative abundances across these categories for each time point per plot. We used the seaborn relplot function (Waskom, 2021) to visualize response categories through time per plot. To measure the host signal by response category, we summed the unique vOTUs of a given bacterial putative host phylum or domain per category. We visualized host abundances per response category by overlaying seaborn swarmplots on boxplots where each point corresponds to a field plot.

Construction of integrase relative abundance and richness barplot

We used the seaborn relplot function (Waskom, 2021) to construct graphs of vOTUs that did (4463 unique vOTUs) or did not contain integrase genes. We plotted the set of vOTUs split into these two categories and by circular versus linear, by their aggregated relative abundance through time and by vOTU richness per category through time.

Quantitative stable isotope probing

Atom percent excess (APE) isotopic enrichment values were calculated using quantitative stable isotope probing (qSIP) (Greenlon et al., 2022; Hungate et al., 2015; Sieradzki et al., 2022) by normalizing MAG and vOTU relative abundance to the quantity of DNA per density fraction (Greenlon et al., 2022). Both bacterial and viral genomes were required to be detected in all three replicates of each labeled ^{18}O treatment and in three replicates of unlabeled (^{16}O) treatments for the bootstrapped median atom percent excess (APE) to be calculated per genome. In the qSIP output

we removed genomes with a lower APE median confidence interval < 0 (i.e., not statistically significant), and ranked the remaining 75 MAG and 177 vOTU genomes first by time and then by descending APE value. We identified the first timepoint each genome appeared as enriched and constructed ranked APE curves using the seaborn python package relplot (Waskom, 2021). To compare MAG and vOTU enrichment dynamics, we calculated the mean differences in rank of isotopically enriched genomes (relative to each previous time point). We plotted median APE values per MAG and vOTU through time for our most prevalent vOTUs and hosts – Actinobacteria, Alphaproteobacteria, and Gammaproteobacteria – using seaborn line relplots. We calculated the ratio of vOTU to MAG enrichment by summing the median APE values of vOTUs and MAGs per phylum and using MAG aggregated enrichment as the denominator.

Viral contribution to microbial mortality

We modeled the viral contribution to microbial mortality using a range of potential burst sizes (virions per infected cell) to relate the number of approximated infectious viral particles through time to observed microbial cell death counts. Counts of microbial cell mortality were derived from our previous qSIP analysis of corresponding amplicon sequence data (Sieradzki et al., 2022). We used the median rate of rRNA copies lost per gram soil per day and converted this rate to per gram soil by multiplying by the number of days. To convert rRNA copies to cells lost, we used a conservative assumption of six rRNA copies per microbial cell (Klappenbach et al., 2000). The approximate loss of microbial cells per gram of soil was used to represent the total observed microbial mortality. To calculate the contribution of viruses to observed mortality, we estimated the number of total virions per time point, using the mean aggregated relative abundance across plots (total reads mapped to our vOTU dataset). We used the fraction of viral reads to estimate viral DNA of total extracted DNA per time point per plot. With our estimate of viral DNA and a mean genome length of 2.1254×10^4 bp, we used equation 1 to estimate plasmid copy number to calculate virions per time point:

$$\text{Virions} = \frac{X \text{ ng} \times 6.0221 \times 10^{23} \frac{\text{molecules}}{\text{mole}}}{N \times 660 \frac{\text{g}}{\text{mole}} \times 1 \times 10^9 \frac{\text{ng}}{\text{g}}} \quad (1)$$

$X = \text{Calculated viral DNA}$
 $N = \text{Average calculated genome length (bp)}$

As an average viral burst size in soils is unknown, we converted number of virions to cells lysed by using a range of viral burst sizes from one virus released per cell (Kuzakov & Mason-Jones, 2018; K. E. Williamson et al., 2008) to 1000 (Sui et al., 2021). Virions scaled to cells dead by the burst size range served as a numerator for viral contribution to microbial mortality calculated with equation 2. We next captured cumulative viral contribution to microbial mortality through time by taking the cumulative sum of all previous time points for each time point.

$$\text{Viral contribution to microbial mortality} = \frac{\text{Virions} \times \text{burst size}}{\frac{\text{degraded rRNA copies}}{6} \times \text{Days}} \quad (2)$$

Our calculations made the following assumptions: 1) All virions detected lyse their host. 2) Virions calculated for each time point represent new infections rather than virions persisting through time i.e., virions were released from newly lysed cells. 3) Our VLP-virome extraction protocol efficiently extracted 100% of soil viruses, i.e., viruses detected represent all possible soil viruses. 4) The mean vOTU genome length represented an accurate approximation of soil viral lengths irrespective of short read metagenomics and potential assembly fragmentation. Our data supported the second assumption given that only a minority of vOTUs appeared in multiple time points (Supplementary Figure 2). Further, we performed a sensitivity analysis of the impact of genome length on our calculation and found a statistically significant relationship between genome length and percent viral contribution to mortality (linear regression on log transformed data, p-value: 6.63×10^{-07}) – changing viral genome length by an order of magnitude led to an order of magnitude decrease in the percent contribution to mortality.

Acknowledgements

This research was supported by the U.S. Department of Energy (DOE) and Office of Biological and Environmental Research (BER) under awards DE-SC0020163 and DE-SC0016247 to M.K.F. Work conducted at Lawrence Livermore National Laboratory was supported by the U.S. DOE, BER, and Genomic Science Program Lawrence Livermore National Laboratory (LLNL) as part of ‘Microbes Persist’ Soil Microbiome Scientific Focus Area SCW1632 under the auspices of the U.S. DOE under contract DE-AC52-07NA27344. We acknowledge funding support to J.F.B. from the Chan Zuckerberg Biohub and the Innovative Genomics Institute at UC Berkeley.

We would like to thank Rachel Hestrin and Alexander Jaffe for helpful conversations regarding data analysis and visualization, and interpretations of isotopic data. We thank Katerina Estera-Molina for her expertise and management of the Firestone Lab field work site at the Hopland Research and Extension Center (HREC). Members of the Taga and Firestone labs provided helpful commentary and criticisms of this work. In particular, Maggie Yuan engaged in important conversations on statistical analyses. Rohan Sachdeva and Shufei Lei of the Banfield lab equipped this project with the necessary computational power and data management via the biotite server and ggKbase, respectively. Support staff at LLNL enabled the large-scale wet-up microcosm experiment and sample extractions, particularly Xiao Bin Max Li and Marissa Lafler.

We acknowledge that HREC sits on traditional, unceded land of the Pomo Indians. Research conducted at the University of California, Berkeley took place on the territory of xučyun (Huichin), the ancestral and unceded land of the Chochenyo-speaking Ohlone people, the successors of the sovereign Verona Band of Alameda County. This land was and continues to be of great importance to the Muwekma Ohlone Tribe and other familial descendants of the Verona Band.

J.F.B. is a founder of Metagenomi. The other authors declare no competing interests.

3.5 Figures

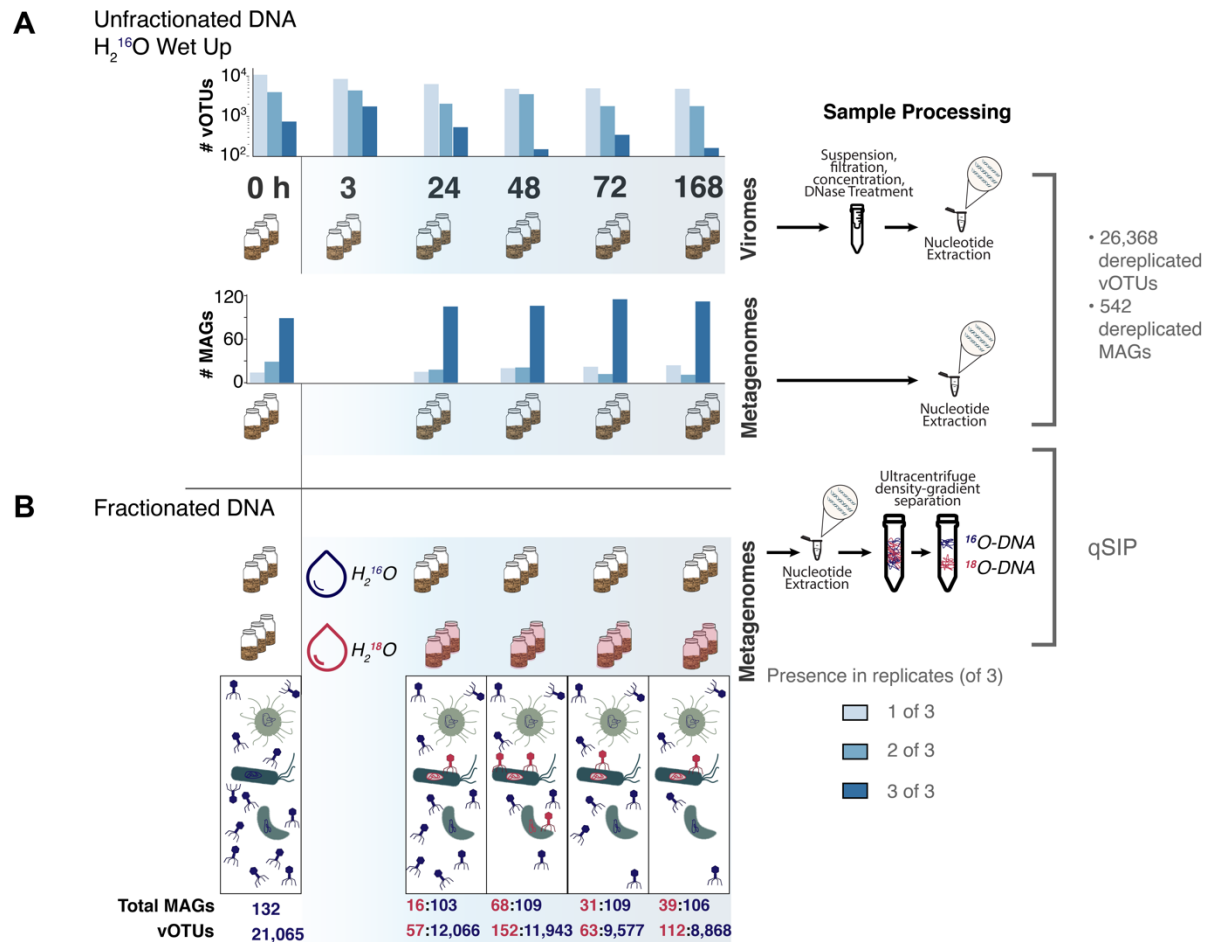


Figure 1 | Experimental overview and analysis of spatial heterogeneity of recovered Genomes from time-resolved cellular and viral size-fraction metagenomes following soil wet-up. (A) Triplicate sampled soil microcosms were destructively harvested before (0h) or following addition of natural abundance water ($H_2^{16}O$) (3h, 24h, 48h, 72 and 168h) for the generation of viromes and metagenomes. Microcosms fated for metagenomic sequencing did not include a 3h time point. Associated barplots visualize the number of (top) recovered vOTUs (log scale) or (bottom) MAGs found in one (lightest blue), two, or three (darkest blue) of three replicate microcosms. (B) Stable isotope probing (SIP) incubation: $H_2^{16}O$ or $H_2^{18}O$ (in red) were added to triplicate microcosms (no 3 h time point), which were then density-separated prior to shotgun sequencing. Below is a cartoon demonstrating that viral isotopic incorporation relies on microbial assimilation of ^{18}O into DNA. The associated counts show the total number of enriched MAGs and vOTUs (in red) relative to the total number of MAGs and vOTUs (blue) detected at each time point.

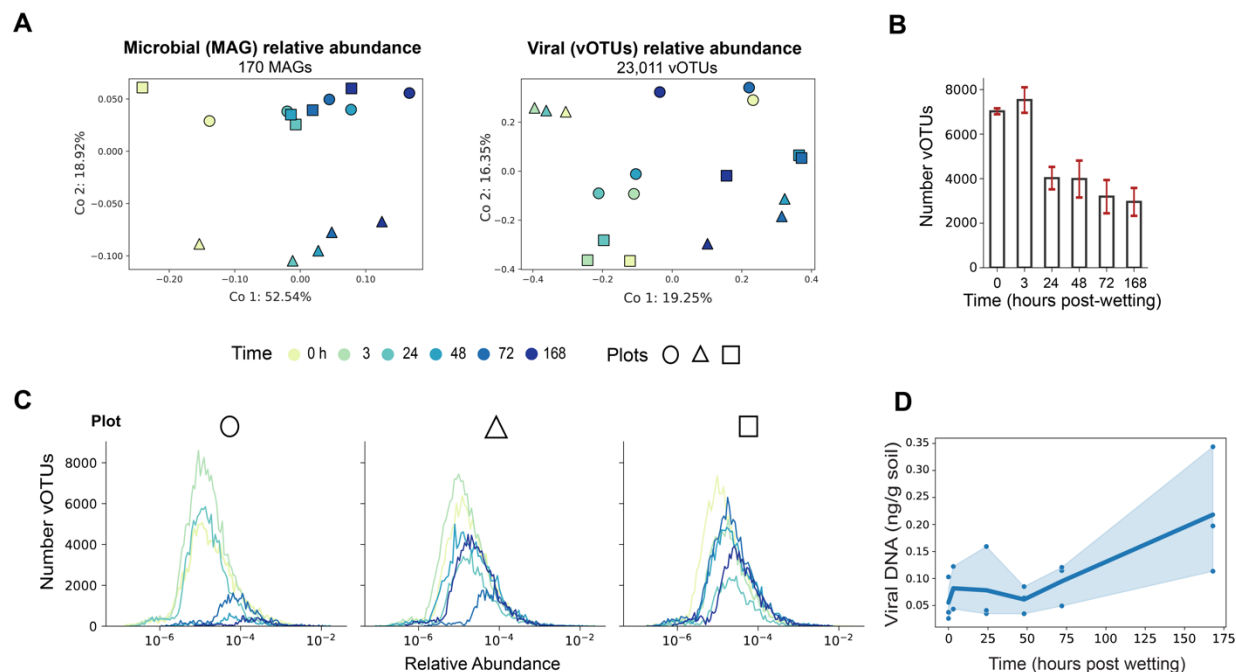


Figure 2 | Temporal compositional changes in microbial and viral communities following soil wet-up. (A) PcoA plots of MAGs (left) and vOTUs (right) calculated based on Bray-Curtis dissimilarity colored by time with different marker shapes signifying distinct field plots (biological replicates). **(B)** Barplot of the mean total number of unique vOTUs per time point. Error bars (red) show standard error of mean across all three field plots. **(C)** Distribution of vOTU relative abundances per time point (colored curve) graphed by field plot (signified by shapes used in A). The y axis is the number of unique vOTUs detected per time point per plot. **(D)** Viral biomass approximated by total extracted viral DNA (ng) per gram of soil plotted through time. Each point represents a distinct microcosm. The shaded area around the line represents a 95% confidence interval.

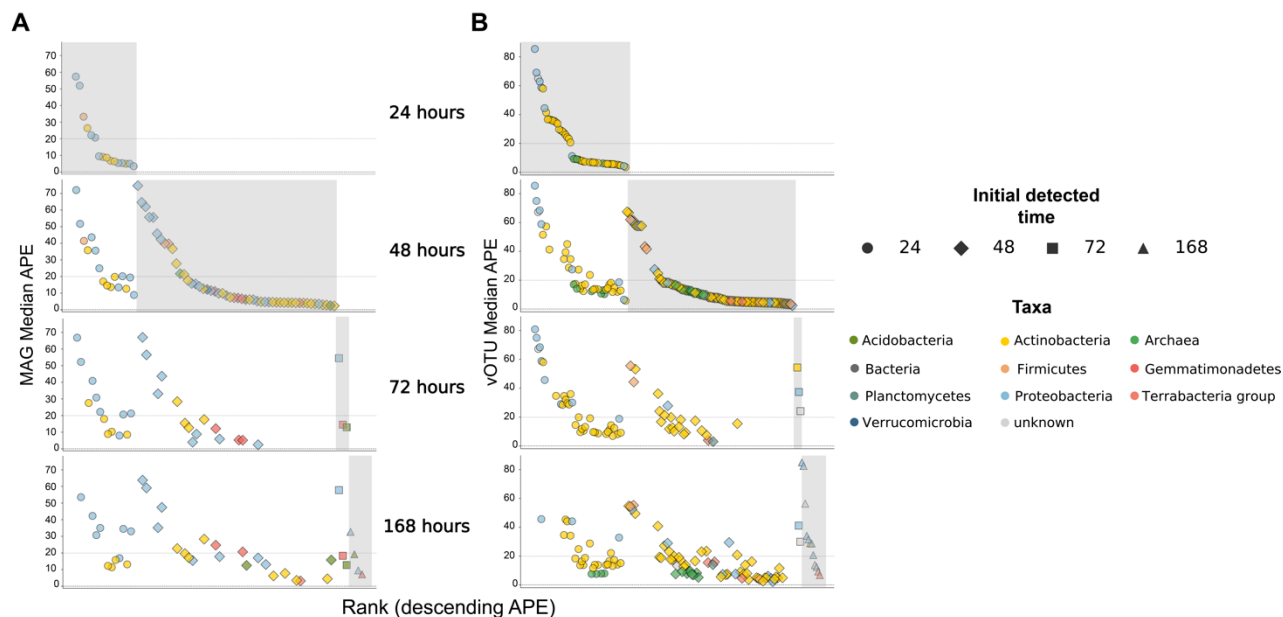


Figure 3 | Changes in abundance of isotope-enriched (growing) viruses and microbes through time following soil wet-up. Rank abundance curves of atom percent excess (APE) ^{18}O wherein the x axis shows absolute rank of genomes sorted first by time and then by descending APE value, and the y axis shows the median APE per genome of (A) microbes (MAGs) and (B) viruses (vOTUs). Colors represent taxonomic groups. Shapes indicate the time point a genome first appeared as isotope-enriched. Gray regions represent genomes appearing as newly enriched in a given timepoint.

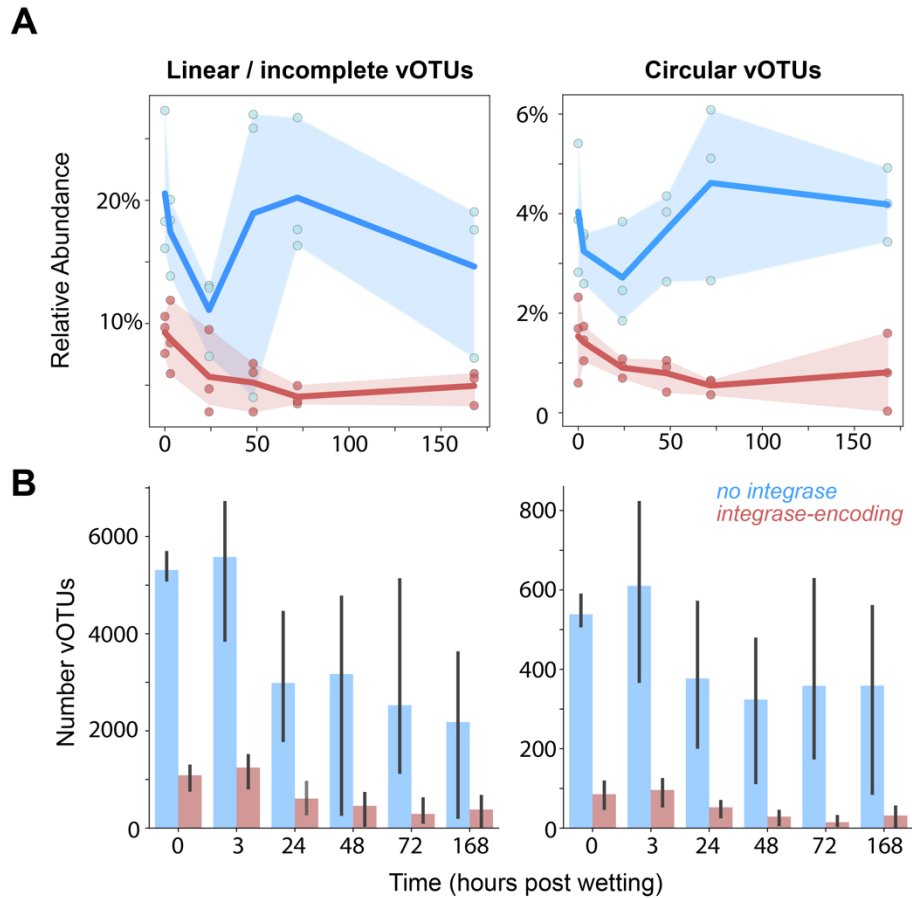


Figure 4 | Relative abundance and richness of integrase-encoding vOTUs through time following soil wet-up. vOTUs were split by whether they encode an integrase gene (red) or not (blue) and whether the vOTU is predicted to circularize (right set of graphs) or is linear and perhaps fragmented (left set of graphs). **(A)** Aggregated relative abundance (percent of total sample reads) of integrase-encoding or not integrase-encoding vOTUs. **(B)** Counts of integrase-containing and non-integrase containing vOTUs.

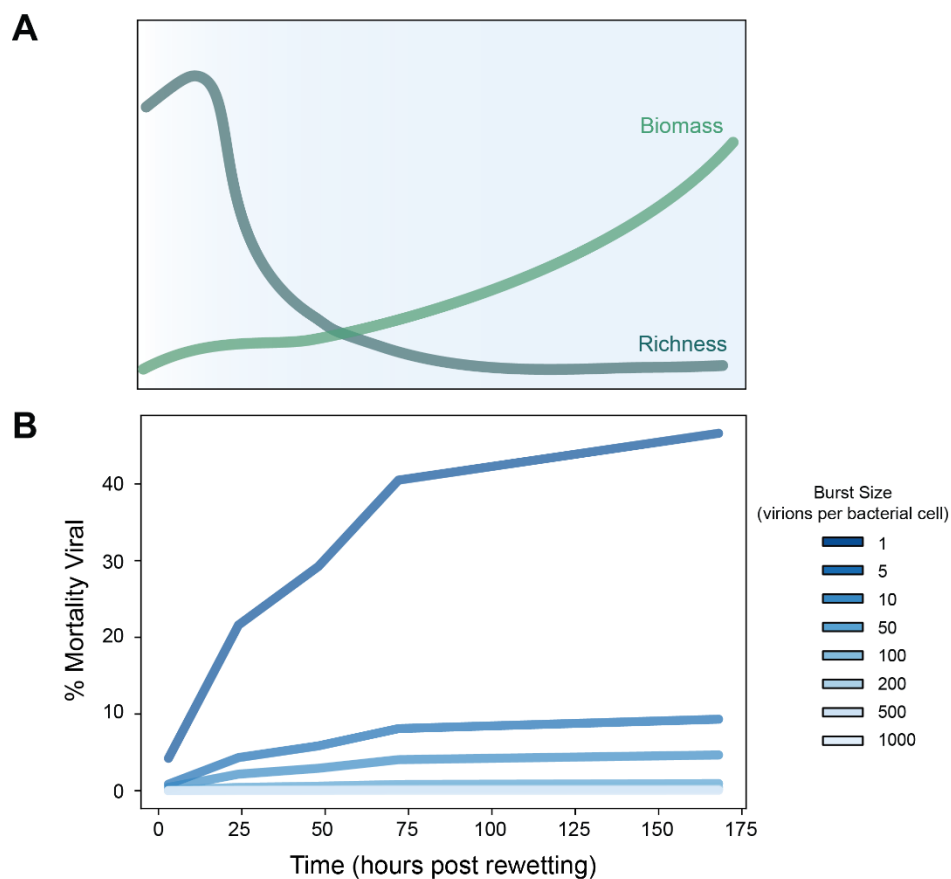
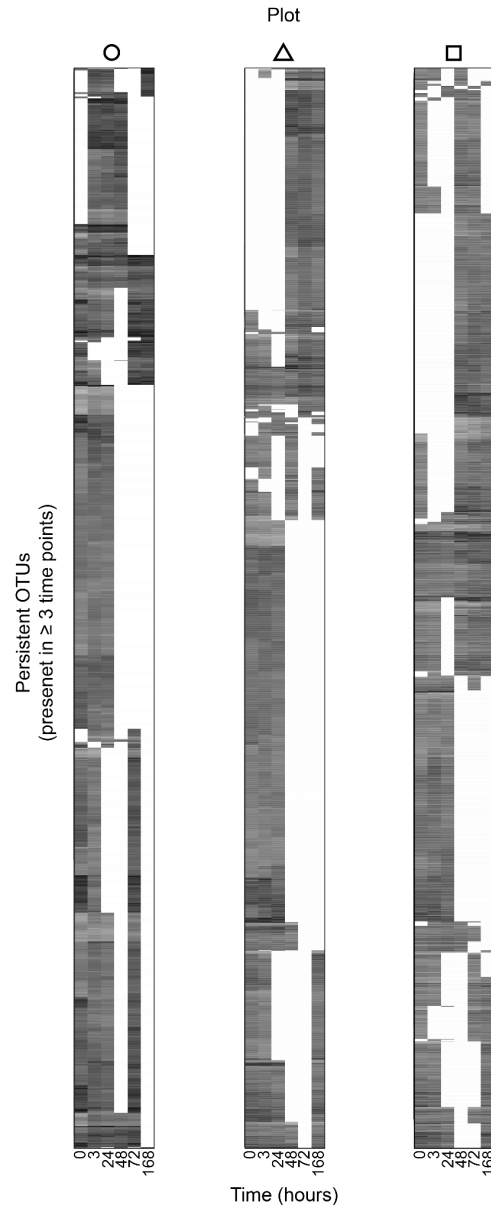


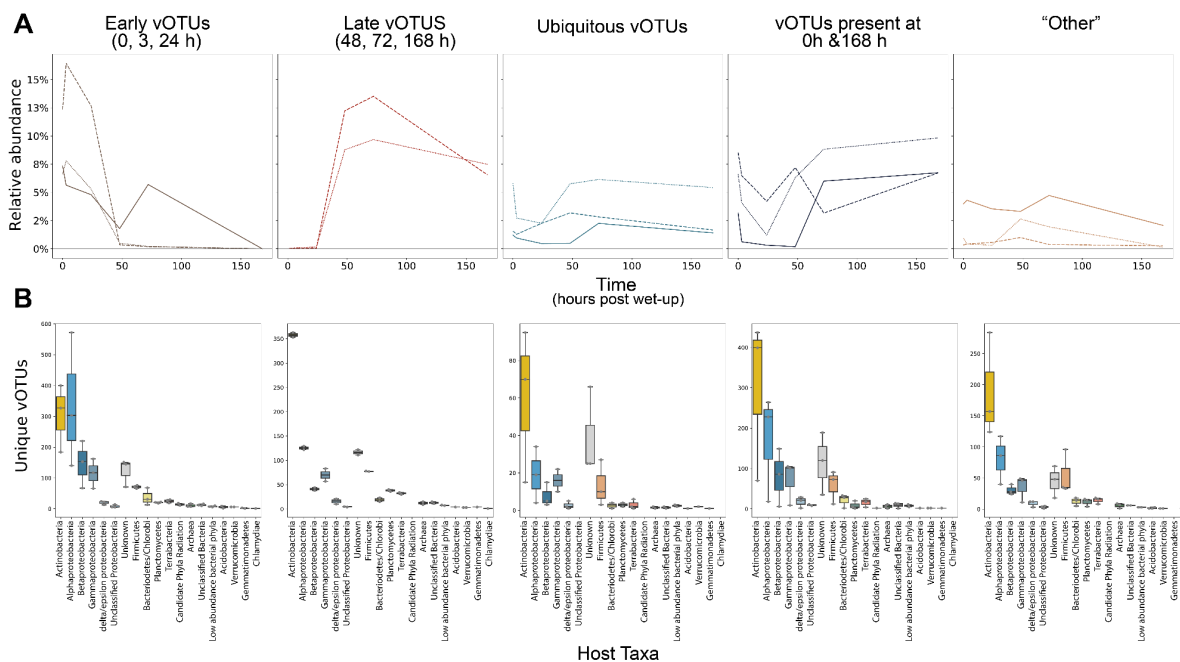
Figure 5 | Viral response to wet-up and corresponding contribution to microbial mortality. (A) Conceptual model of changes in viral richness (gray-blue) and biomass (green) over time. (B) Modelled contribution of viruses to microbial mortality where each line graphed represents an estimated percent contribution for a different viral burst size shown in the legend ranging from 1 virion per lysed microbial cell to 1000 virions per lysed microbial cell.

Metagenome type	Method	Presumed composition
Unfractionated metagenome	Direct DNA extraction from soil (Figure 1A)	Bacterial, archaeal, eukaryotic, and viral sequences from MAGs and vOTUs (which may be intracellular or extracellularized).
SIP-fractionated metagenome	DNA extracted from soil and ultracentrifuged in a CsCl gradient to capture distinct DNA-density fractions for quantitative stable isotope probing (qSIP) (Figure 1B)	H ₂ ¹⁸ O-incorporating MAG and vOTU genomes in heavier DNA fractions i.e., considered active.
Virome	Small-fraction metagenome, i.e. viral-enriched sequences: soil buffered, 0.2 µm filtered, concentrated, treated with DNase. DNA extracted from 0.2 µm effluent (Figure 1A)	Sequences from < 0.2 µm membrane-bound or encapsulated genomes i.e. Viral-like particles (VLPs) or virions; and ultrasmall cells such as CPR bacteria and nanoarchaea

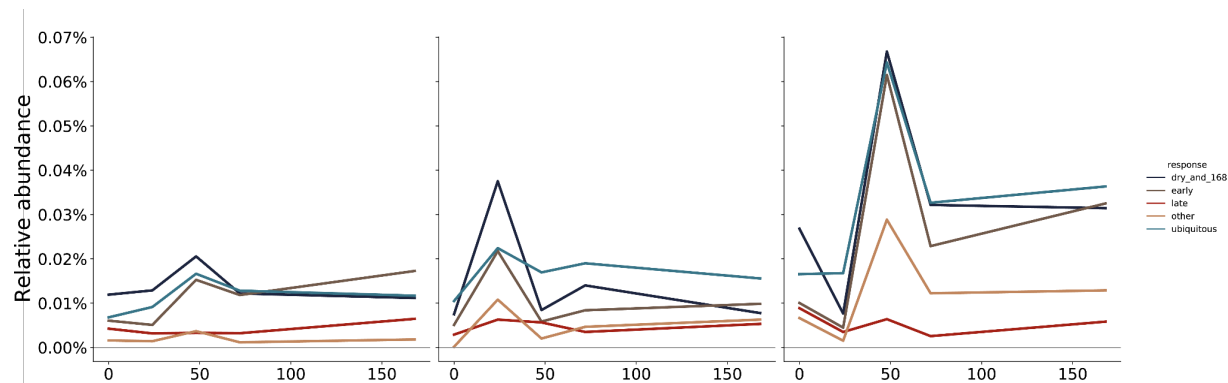
Supplementary Table 1 | Mixed ‘omic methods enable detection of genomes and distinguishing as abundant, active, and virions.



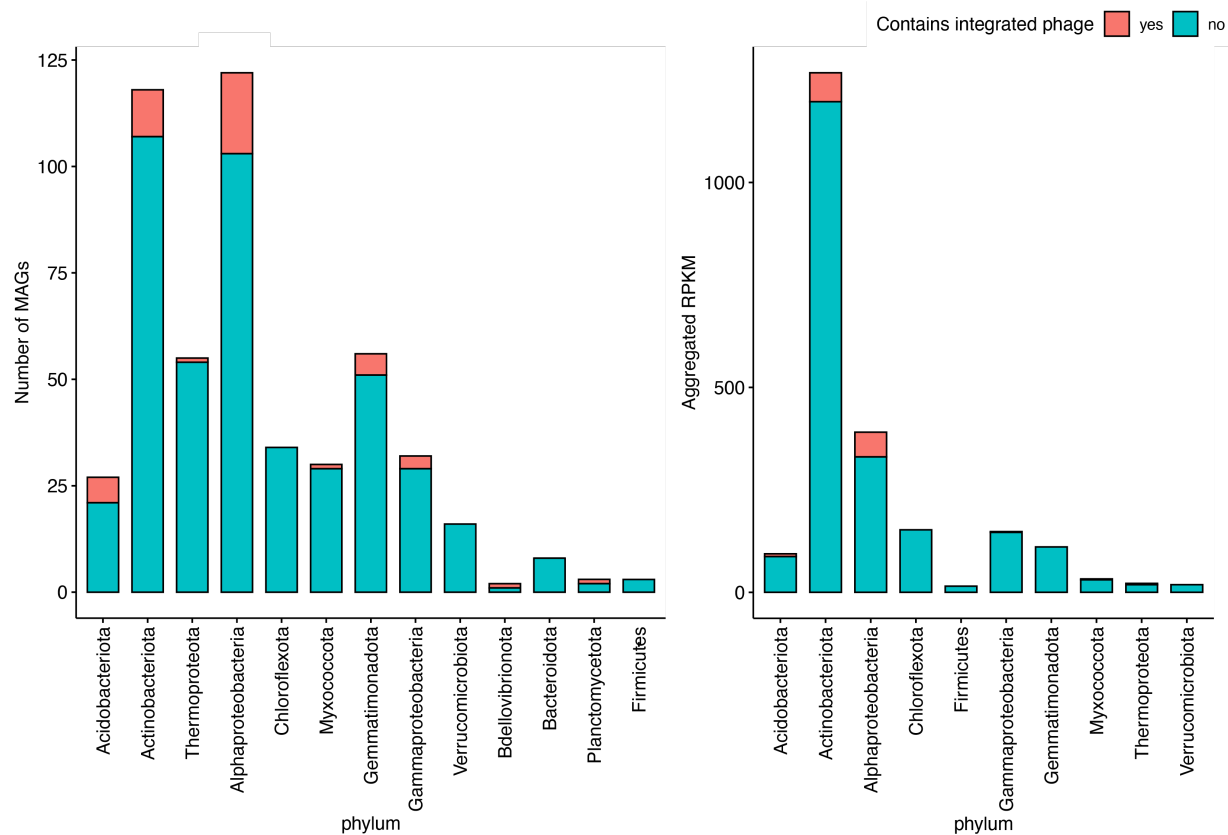
Supplementary Figure 1 | Temporal changes in relative abundance of persistent viral populations (vOTUs). Each heatmap corresponds to one of the three field plots sampled and is hierarchically clustered according to vOTU relative abundance. Time is shown on the x-axis of each plot (columns); Each row (y-axis) corresponds to a unique vOTU and its relative abundance through time. All vOTUs were present in each given plot in at least three time points, i.e., “persistent” vOTUs. Circles, triangles, and squares correspond to the field plots in the PCoA legend (Figure 2).



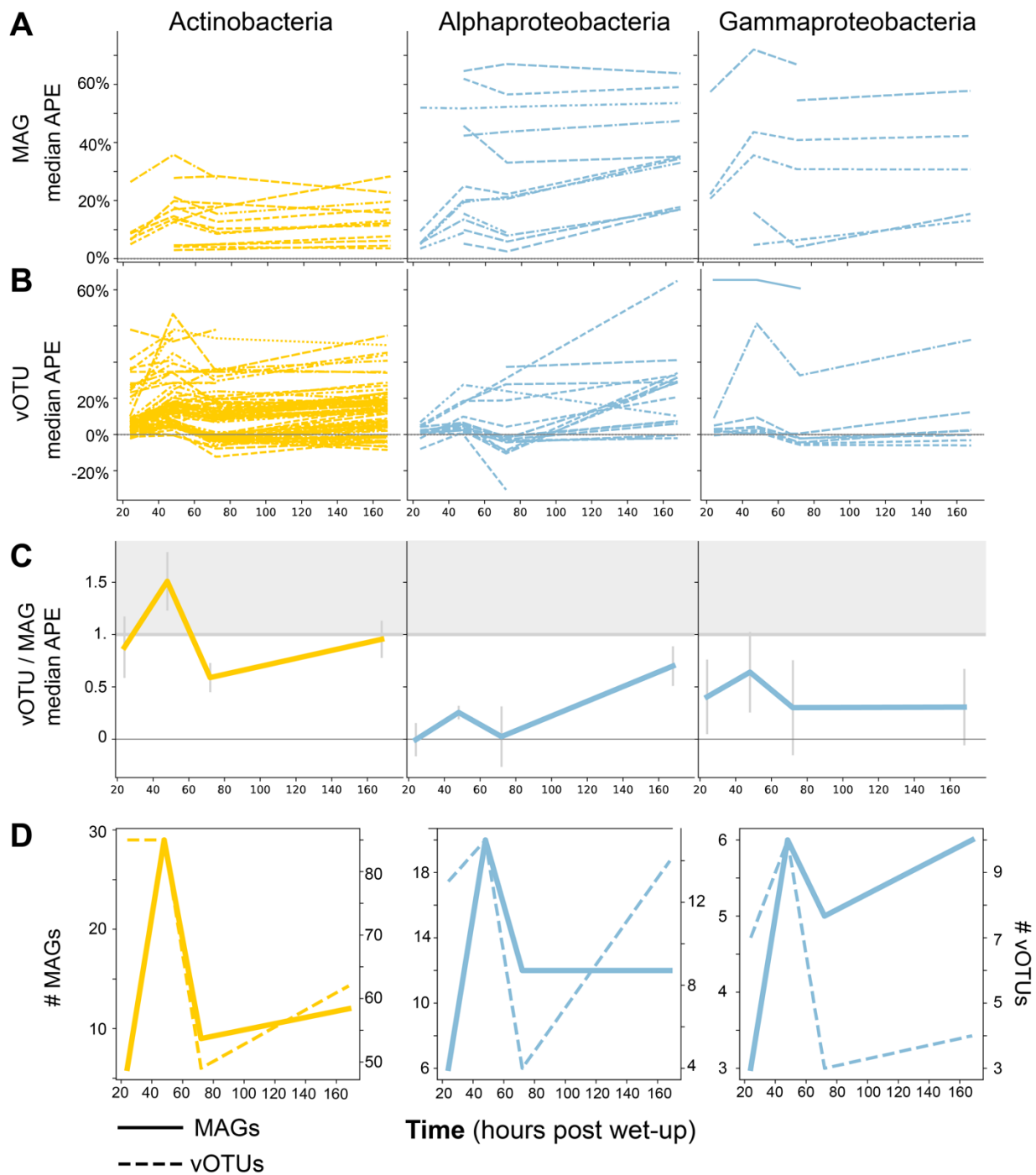
Supplementary Figure 2 | Viral dynamics following soil wet-up classified by temporal response. (A) Each plot shows the aggregated relative abundance per plot of the persistent vOTUs that respond early (0, 3, 24h); late (48, 72, 168h); ubiquitously (at all time points); present at 0h, 168h, and one other time point; and other. Relative abundance is shown as the percent of total reads per virome. (B) Boxplots graphed per response category showing the range of vOTU counts (y axis) separated and colored by predicted host taxonomic groups (x axis) across all three plots. The x axis is sorted by rank, but with all Proteobacteria groups adjacent to one another. The box plots represent 75% of the data with the median as a line, and whiskers represent 90% of the data.



Supplementary Figure 3 | Field plot-specific viral responses in unfractionated metagenomes. Each graph represents unfractionated metagenome reads from one field plot mapped to the set of viruses of a specific response category defined in the virome (Sup. Fig. 2). Relative abundance represents the percent of total aggregated reads mapped to vOTUs of a given category.



Supplementary Figure 4 | Prophage detected in bacterial genomes (MAGs). For each MAG of a given phylum the number of MAGs (left) and the relative abundance in RPKM of those MAGs is shown according to whether the MAG contains an integrated phage (red) or not (blue).



Supplementary Figure 5 | Temporal isotopic enrichment of MAGs and vOTUs by phylum. Each column of graphs represents the phylum-specific enrichment response to wet-up. **(A)** Median APE per isotope-enriched MAG through time. **(B)** Median APE per isotope-enriched vOTU through time. **(C)** The mean ratio of vOTU/MAG median APE through time. Error bars represent the standard error calculated from propagating the error of median isotope enrichment per genome and of the mean enrichment ratio. The gray region indicates ratio >1. **(D)** The number of MAGs (solid line) or vOTUs (dashed line) detected as isotope-enriched through time.

4. Corrinooids influence community assembly in a soil-derived enrichment

Alexa M Nicolas¹, Zachary F Hallberg¹, Rohan Sachdeva^{2,3}, Hans K Carlson^{4,5}, Jillian F Banfield^{2,3,6,7}, Michiko E Taga^{1*}

¹Plant & Microbial Biology Department, University of California, Berkeley, CA, USA

²Innovative Genomics Institute, University of California, Berkeley, CA, USA

³Earth and Planetary Sciences, University of California, Berkeley, CA, USA

⁴Lawrence Berkeley National Laboratory, Berkeley, California, CA, USA

⁵Energy & Biosciences Institute, University of California, Berkeley, Berkeley, CA, USA

⁶Environmental Science, Policy & Management Department, University of California, Berkeley, CA, USA

⁷The University of Melbourne, Melbourne, Victoria, Australia

*Corresponding author email: taga@berkeley.edu

Earth's soils are habitats for microbial communities that drive biogeochemical cycling, plant growth, and carbon storage and persistence. To probe the interactions between microbes that underlie the soil microbiome we used a class of key metabolites related to vitamin B₁₂, known as corrinooids. Only a subset of the microbes that use corrinooids can produce them, and bacteria in culture exhibit preferences for different corrinooid types in their environment. We combined multi-omic approaches in soil and in soil-derived enrichments to test how corrinooids, as a model metabolite, impact community structure and function. In soil, 16S rDNA amplicon sequencing revealed that corrinooid addition led to compositional changes that are corrinooid-specific, but transient and most pronounced just three days after addition over a 50-day incubation. In both metagenome-assembled genomes (MAGs) from soil and soil-derived enrichments we find few genomes to be predicted corrinooid producers while predicted corrinooid use is widespread across MAGs. Methionine synthesis is the most prevalent corrinooid-dependent metabolism in enrichments and in soil. Based on genome abundance (metagenomes) and total significantly differentially expressed genes (metatranscriptomes) we found many genomes appear to have a preference for B12 and cobinamide. Further, we observed significant downregulation of corrinooid independent ribonucleotide reductases, a critical step in DNA synthesis, in corrinooid treated enrichments. Overall, corrinooids are able to shift community composition in soil and community assembly of soil-derived enrichments. These findings enable us to hypothesize interactions that may occur in soil based on the sharing of this widely-used nutrient. By simplifying a complex system to the lens of a single shared nutrient and testing how this nutrient impacts community structure and function in the environment and in a simplified enrichment we provide a framework for teasing apart the complex interactions that occur in environmental microbiomes.

4.1 Introduction

Soils maintain thousands of microbial species and billions of cells in each gram (Delgado-Baquerizo et al., 2018; Sokol et al., 2022). These communities drive biogeochemical cycling, plant growth, and carbon storage and persistence. Diverse interactions between microbes comprise these communities (Bahram et al., 2018; Ling et al., 2022; Young & Crawford, 2004). Critical to understanding the important processes soil microbial communities perform is thus mapping the varied interactions among microbes. To probe microbial interactions that underlie the soil microbiome we used a class of key metabolites related to vitamin B₁₂, known as corrinoids. Corrinoids serve as cofactors for a variety of metabolic processes performed by microbes in soil and other environments, such as catabolism of particular carbon sources, nucleotide metabolism, and methionine synthesis (Lu et al., 2020; Matthews, 2009; Roth et al., 1996; Shelton et al., 2019). Only bacteria and archaea are known to biosynthesize corrinoids and only a subset of these microbes that use corrinoids can produce them (Roth et al., 1996; Warren et al., 2002). Importantly, corrinoids are structurally diverse, and bacteria studied to date in isolation can use only some of the corrinoid forms that are available in their environment (Heal et al., 2017; Kennedy et al., 2022; Matthews, 2009; Men et al., 2014; Mireku et al., 2017; V. Sharma et al., 2019; Sokolovskaya et al., 2019; Woodson & Escalante-Semerena, 2004). Whether this specificity impacts bacterial growth and metabolism on a community scale has not been examined.

While it is well accepted that substrate additions and growth conditions (i.e. anaerobic vs. aerobic) impact microbial community composition and function *in situ* and environmentally (Baker et al., 2022; Carlson et al., 2020; Chuckran et al., 2021; Justice et al., 2017), how nutrients such as vitamin cofactors that do not become integrated into biomass, i.e. non-assimilatory nutrients, may impact community composition and function is less well studied (Kelly et al., 2019; V. Sharma et al., 2019). Corrinoid requirements have been investigated in greater depth in non-soil communities. In marine systems, cobalamin is a limiting nutrient for phytoplankton growth, implicating corrinoids in affecting the ocean food web (Bertrand et al., 2007; Sañudo-Wilhelmy et al., 2014; Seth & Taga, 2014). Requirements for specific corrinoids have been illustrated in several pairs of aquatic microbes engaged in mutualistic corrinoid cross-feeding (Croft et al., 2005; Heal et al., 2017; Helliwell, 2017; Iguchi et al., 2011). Animal gut environments have also been shown to contain six or more corrinoids (Allen & Stabler, 2008; Hallberg et al., 2022; Liu et al., 2022), and corrinoid-sharing relationships have been observed between human gut bacteria *in vivo* (Degnan, Barry, et al., 2014; Degnan, Taga, et al., 2014). Previously, research showed that addition of cobalamin to trichloroethene (TCE)-degrading enrichments alters community structure, and that the TCE-metabolizing bacterium *Dehalococcoides mccartyi* has stringent requirements for specific corrinoids (Men et al., 2015; Yi et al., 2012). Taken together, these results provide support for our hypothesis that corrinoids impact community composition in diverse environments. However, this hypothesis has not yet been tested in soil.

To tease apart microbial interactions in such a complex system we developed a simplified framework centered on corrinoids as a key metabolite. We hypothesized that adding excess corrinoids would change the corrinoid profile each microbe senses and stimulate growth of microbes that preferentially use the added corrinoid over environmentally available corrinoids. Underlying our hypothesis is the assumption that corrinoids in general, and specific corrinoid types in particular, are limiting nutrients in soil. We thus tested the effect of exogenous corrinoid

additions in soil microcosms. To further elucidate how corrinoid addition and corrinoid type connect microbes in soil we established soil enrichment cultures (Goldford et al., 2018) and, following one week of passaging, added one of four distinct corrinoid types to test how corrinoids may shift community function and composition during community assembly (Figure 1). Previous work in a synthetic co-culture showed that symbiosis between a B12-producing bacterium and a B12-dependent alga could be decoupled with addition of B12 (Kazamia et al., 2012). We predicted that corrinoid addition could decouple interactions between corrinoid producing microbes, i.e. producers, and corrinoid dependent microbes, dependents, in a complex community.

We combined analyses from a grassland soil environment – using metagenomics on soil samples without added corrinoids (Nicolas et al., 2022; Sieradzki et al., 2022) and 16S rDNA amplicon sequencing on corrinoid-perturbed microcosms – with 16S rDNA amplicon sequencing, metagenomics, and metatranscriptomics on soil-derived enrichment cultures with added corrinoids to understand how corrinoid preferences of individual organisms scale to impact community composition and function. We found that different corrinoid types impact community composition transiently in soil microcosms and subtly in enrichment cultures. Further, we observed that different corrinoid types elicit significant differential gene expression on a community-wide scale that may downregulate corrinoid-independent pathways in specific organisms.

4.2 Results

Analysis of MAGs predicts widespread corrinoid sharing in soil

We established soil as an environment where potential corrinoid-producing and corrinoid-dependent microbes reside and interact by assessing corrinoid producing and using capacity of 446 MAGs assembled from our Northern California grassland fieldsite at Hopland Research and Extension Center (HREC). These MAGs were sampled at the end of the dry season and were subject to the addition of water in the lab to simulate the annual rewetting event these microbes experience in their native environment. Based on the six marker genes (three each anaerobic and aerobic) established in Shelton et al. 2019 (Shelton et al., 2019), we found that only a subset of our MAGs were likely to be corrinoid producing (< 20%). The vast majority of these MAGs (> 70%) encoded at least one corrinoid-dependent enzyme (Figure 2A). Methionine synthase was the most prevalent corrinoid-dependent enzyme detected (Figure 2C). The laboratory rewetting of soil microcosms was also performed with H_2^{18}O enabling us to capture a temporal view of isotopic enrichment as a proxy for activity of these MAGs (discussed further in chapter 3) (Nicolas et al., 2022; Sieradzki et al., 2022). Actively replicating microbes assimilate ^{18}O from H_2^{18}O in their DNA. We observed that the amount of active predicted corrinoid producers changes through time in a highly dynamic moment in the soil microbiome (Figure 2B). Through time it also appears that the ratio of predicted producers to non-producers is somewhat dynamic.

Impact of corrinoid treatment on soil microbial composition

Given the prevalence of corrinoid-dependent metabolisms in soil microbes, we wanted to understand whether addition of corrinoids directly to soil could alter soil community composition i.e. whether addition of a non-assimilatory trace nutrient may reveal community-wide impacts. We

supplemented destructively harvested soil microcosms with water or one of six structurally distinct corrinoids – B12, cobinamide (Cbi), Factor A, [Cre]Cba (Cre), pseudocobalamin (pB12), or [5-OHBza]Cba. Following DNA extraction and 16S rDNA amplicon sequencing from 150 soil microcosms and establishing operational taxonomic units of microbes, here zOTUs, we analyzed the impact of corrinoid treatment effects on each taxonomic level. Based on PCoAs of Bray-Curtis dissimilarity (Figure 3), we observed the strongest treatment effects 3 days after corrinoid addition on all phylogenetic levels – from phylum to genus. In particular, we observed that samples amended with cobinamide, B12, or no corrinoid clustered closest together and the first two principle components explained at least 30% of the variance in data. On the genus level, all but our 50-day time point showed a treatment effect. On a higher phylogenetic level, day 30 also showed a statistically significant treatment effect (ANOSIM $p \leq 0.05$, test statistic ≤ 0.15), but qualitatively was less distinct by PCoA. By day 50 we did not observe corrinoid addition to have a qualitative or statistically significant impact on soil community composition, suggesting that the effect of corrinoid amendment on community composition was transient.

Establishing soil enrichment cultures to analyze corrinoid-related dynamics more granularly

We established simplified soil-derived communities in M9 minimal media to further test our hypothesis that corrinoid addition and type could structure microbial communities and elucidate in a simplified system the corrinoid-related treatment effects observed in our soil microcosms. After inoculating soil into our culture media we allowed a community to assemble for one week before splitting and passaging our enrichments, and allowing them to reestablish for one hour (Table 1). Following one hour, we added one of four corrinoids: cobinamide (Cbi), cobalamin (B12), pseudo-cobalamin (pB12), Cresolylcobamide (Cre), or no corrinoid. These corrinoids were chosen because we observed in soil microcosms that the treatments seemed to partition between no corrinoid, B12, or pB12, compared to other corrinoids. We next sampled our enrichments 3 hours after corrinoid addition (172 hours total in culture) for metatranscriptomes, and again at 1 week after corrinoid addition (336 hours in culture or 2 weeks) to generate metagenomes. For metatranscriptomes, we chose a 3 hour time point based on previous research that has shown that soil microbes can respond quickly to environmental changes (Chuckran et al., 2021; Placella & Firestone, 2013) and to see direct effects of corrinoid addition on gene expression while mitigating indirect influences on other organisms. For each sampled time point shown in Table 1, we sequenced 16S rDNA to understand treatment impact on community composition. For our 172-hour time point there was very low biomass as proxied by extractable DNA, thus we were not able to establish metagenomes at this time point. However, we were successful in attaining extracted RNA quantities high enough to proceed with metatranscriptomics for all but one replicate each of our Cre and pB12 conditions.

To test whether the treatment effects observed in soil extend to culture conditions we first looked at whether there were marked shifts in community composition following corrinoid addition. We assessed total community composition using our 16S rDNA amplicon sequencing dataset and found that our soil enrichments were predominantly comprised of Proteobacteria (137 zOTUs across all samples) as the most prevalent detected zOTUs at over 90% total abundance. We constructed ordination plots based on Bray-Curtis dissimilarity (Figure 4), iterating through each taxonomic level. Across all samples we detected 214 unique zOTUs. Appreciable treatment related differences were observed only 1 week after corrinoid addition (2 weeks in culture, 336 hours). At

the zOTU level we noticed a pattern of clustering where Cbi, B12, and no corrinoid samples of the same replicate clustered in higher proximity than Cre and pB12. However, neither ANOSIM (p-value 0.93) nor PERMANOVA (p-value 0.991) tests noted these observations as statistically significant. At 2 weeks, at the phylum level corrinoid addition led to a statistically significant change in the community (ANOSIM p-value 0.01; test statistic 0.27).

We sought to examine changes at the phylum level to understand why corrinoid amendment of our cultures was significant only at this high order taxonomy (Figure 5). Proteobacteria did not show appreciable differences across corrinoid conditions. We observed that corrinoid addition may impact the rarer phyla in our enrichments. Within the Firmicutes (30 zOTUs) and Actinobacteria we noticed differences across corrinoid conditions between 3 hours after corrinoid addition and 1 week after addition where B12 addition showed the highest relative abundance of these phyla. For Actinobacteria this represented two distinct zOTUs: *Nocardia coeliaca* and *Microbacterium oxydans*.

Since we hypothesized that the addition of corrinoids would disentangle established interactions between producing microbes and organisms that relied on exogenous corrinoids, we generated 24 metagenomes from our enrichment cultures prior to corrinoid addition and at 1 week following addition. From these samples we resolved 28 near-complete dereplicated MAGs to analyze their genomes for corrinoid biosynthesis and dependent metabolisms (Figure 6). 13 of these genomes represented Proteobacteria and the majority of these were Gammaproteobacteria. We resolved two Actinobacteria genomes. We found one genome of the predatory bacteria phylum, Bdellovibrinota, of which there was no corresponding zOTU. This organism only appeared in one replicate, but given its likely distinct predatory role (Jurkevitch et al., 2000) further underscores that there may be important replicate-specific dynamics at play. Despite some compositional differences between our MAG and 16S rDNA amplicon sequencing data we found that on average there were 30 to 80 zOTUs per sample and resolved 28 bins overall. Given that the majority of zOTUs detected and the highest frequency of zOTUs per sample were Gammaproteobacteria which are typically overrepresented in 16S sequencing studies due to their, on average, higher 16S copy number and divergence of sequences (Acinas et al., 2004; Větrovský & Baldrian, 2013), our metagenomic data and 16S profiling seem parable.

In order to understand the dynamics of MAGs in our enrichments, we classified each organism by its ecological role related to our treatment: corrinoid producers, corrinoid dependents, and corrinoid independent organisms. Corrinoid producers contained at least one of six marker genes – three aerobic and three anaerobic steps of corrin ring biosynthesis (though HMM models do not predictably distinguish between anaerobic and aerobic homologs). Each of these marker genes was previously shown to be present in 99% of predicted corrinoid producers (Shelton et al., 2019). Corrinoid dependent organisms encoded at least one corrinoid-dependent metabolism. Corrinoid independent MAGs were not detected to have any biosynthesis components or dependent metabolisms. All of our 5 predicted producer genomes (Table 2) also contained corrinoid-dependent metabolisms, consistent with Shelton et al. 2019 (Shelton et al., 2019). Of our 5 producers, four were Proteobacteria: three rhizobia and one burkholderiales; and one was a Bacilli genome (Firmicutes). 22 of the MAGs were corrinoid-dependent and one MAG, a *Microbacterium* (Actinobacteria), was corrinoid-independent.

While predicted corrinoid producers were a subset of all genomes, corrinoid dependent organisms were the mode. Of the 22 dependent genomes and five predicted producers we sought to uncover what corrinoid-dependent metabolisms were functionally possible in our enrichments (Figure 7). As in our set of 446 MAGs from a previous experiment, we found that methionine synthase was the most prevalent corrinoid-dependent metabolism in MAGs from our enrichments, suggesting that our enrichments were representative of soil microbes in terms of the most prevalent corrinoid use. Of our 28 MAGs, only three appeared to not encode any methionine synthase, and one of these genomes was the corrinoid independent Actinobacteria. We observed more organisms encoded a corrinoid-dependent methionine synthase than the corrinoid-independent analog. We further inspected methionine synthase distribution in enrichment MAGs to discern if a subset of our dependent genomes were true auxotrophs, reliant on an externally available nutrient for growth (Zengler & Zaramela, 2018). Of the 24 genomes encoding a corrinoid-dependent methionine synthase, 17 also encoded the corrinoid-independent analog. One dependent MAG only encoded the corrinoid-independent methionine synthase, but maintained other corrinoid-dependent metabolisms. Of the seven MAGs that only encoded the corrinoid-dependent methionine synthase, two were producers and five were not predicted producers suggesting these are auxotrophs of methionine or corrinoids.

Similarly, examining propionate metabolism in our MAGs, we observed that eleven organisms only encoded methylmalonyl-CoA-mutase (MCM) which is a corrinoid-dependent enzyme and no enzymes associated with the corrinoid-independent methylcitrate pathway. Three organisms encoded both corrinoid-independent and -dependent genes involved in the conversion of propionyl-CoA to succinyl-CoA, and five MAGs appeared to only maintain corrinoid-independent metabolism. On the other hand, considering nucleotide metabolism, the corrinoid-independent ribonucleotide reductase (RNR) was more prevalent in MAGs than the corrinoid-dependent homolog. Only four organisms encoded corrinoid-dependent RNR, but three of these MAGs did not also encode the corrinoid-independent RNR. Two of these three MAGs were predicted producers and one, *Edaphocola* MAG of the phylum Bacteroidota appeared to be a corrinoid or nucleotide auxotroph. Nine organisms were detected to only encode the corrinoid-independent RNR, all of these MAGs contained other corrinoid-dependent metabolisms, and one was a predicted corrinoid producer.

Our experiment was based on the assumption that microbes in our enrichment cultures would be able to sense corrinoids in their environment, whether produced by neighboring cells or supplemented in the medium (Men et al., 2014). We identified corrinoid transporters in two predicted Gammaproteobacteria corrinoid-dependent genomes, one of the *Pantoea* genus (*Pantoea brenneri*) and one of *Serratia*. We detected more cobalt transporters in MAGs, ten: in each of our five predicted corrinoid-producing genomes and five in predicted dependent genomes.

After defining the corrinoid-based ecological role for each of our MAGs, we examined the dynamics of MAGs in our enrichments prior to corrinoid amendment and one week after corrinoid amendment (2 weeks in culture). We assessed dynamics in two ways: (1) examining the mean genome abundance in samples by their ecological role; (2) analyzing the total composition of producers and dependents genomes (aggregating MAGs abundance by ecological role) (Figure 8). Despite replicate-dependent compositional differences observed in our 16S rDNA amplicon sequencing, these analyses of MAG community composition by corrinoid production and use

revealed a degree of consistency between replicates. Mean genome abundance analysis showed that prior to corrinoid addition, as the enrichment community assembled, predicted producer and dependent genomes were on average equivalently abundant. By day 7 in culture, dependent genomes, on average, became much more abundant. This trend persisted into the second week of enrichment, after corrinoid addition and was apparent in all corrinoid and non-corrinoid supplemented cultures. Our aggregated abundance data exaggerated the trend featured in our mean genome abundance data: dependents (more numerous in our dataset, n=22) made up a much larger portion of the enrichment community. Thus, by abundance, we observed predicted corrinoid producing genomes, though fewer, are not on average more abundant than dependents nor do organisms that serve the role of corrinoid producer comprise a more substantial portion of the total community i.e. punch above their numbers. Corrinoid addition does not change the mean abundance of producer or dependent genomes, nor the portion of the total enriched community defined in these roles. The hypothesis previously observed in co-culture experiments (Kazamia et al., 2012) did not prove to be true in our more complex enrichment communities.

While corrinoid addition did not discernibly alter the ratio of predicted corrinoid producers or dependent genomes on net or on average, we reasoned that corrinoid addition may impact MAGs individually. We hypothesized that individual genomes may show relatively higher abundances with corrinoid addition compared to no corrinoid addition and that these changes may be corrinoid specific (Kennedy et al., 2022; Mok & Taga, 2013; Sokolovskaya et al., 2019; Yi et al., 2012). Further, we hypothesized that the impact of corrinoid addition may be observed acutely on the level of gene expression (Johnson et al., 2009; Kennedy et al., 2022; Men et al., 2017). We thus analyzed our metagenomic data and metatranscriptomic data in tandem to determine the impact of corrinoid addition on individual MAGs (Figure 9). To reveal corrinoid-specific preference in terms of an organism's growth we used MAG fold difference, the relative abundance of a MAG in a given corrinoid condition compared to its relative abundance without corrinoid supplementation. To discern differences in the activity of a given MAG in a corrinoid-dependent manner we assessed the total number of significantly differentially expressed genes of a MAG (relative to genome size) in each corrinoid condition compared to no corrinoid based on our metatranscriptomic data. Since we sampled our cultures 3 hours after corrinoid addition for metatranscriptomics, we hypothesized that differential expression specific to a given corrinoid condition would portend relative abundance changes we could observe 1 week after corrinoid addition in our metagenomes. Thus, we visualized these genome abundance and differentially expressed gene data in tandem to compare how expression level changes may influence growth differences.

For predicted corrinoid-producer MAGs, two of five showed a degree of preference for growth with supplemented corrinoid (>100% growth in corrinoid amended conditions). Four producers showed significantly differentially expressed genes, and for all four of them a suite of genes appeared differentially expressed with B12 addition. One producer was not present in all replicates and thus did not show significantly differentially expressed genes 3 hours after corrinoid addition. Notably, the *Herbaspirillum* genome showed the highest number of genes differentially expressed in B12 and the highest growth compared to no corrinoid in this condition, suggesting this *Herbaspirillum* genome may have a preference for B12 over Cbi, and perhaps over Cre and Pse (which only had 2 of 3 replicates). The corrinoid independent *Microbacterium* was more abundant without corrinoid supplement (< 100% fold change relative abundance in all treatments) and also did not have corresponding differentially expressed genes.

For predicted corrinoid-dependent genomes, sixteen genomes had metagenomic fold difference treatment data with paired metatranscriptomic data that showed significantly differentially expressed genes. Eight dependents showed some preference for corrinoid amendment – had a higher relative abundance in corrinoid amended condition compared to no corrinoid (fold difference > 100%). Of these genomes, significant differential expression of a suite of genes was most apparent with Cbi and B12 addition. Even corrinoid-dependent genomes that did not show a higher relative abundance with corrinoid addition, exhibited significantly differentially expressed genes in the presence of added corrinoid. In particular, the *Serratia* genome demonstrated the highest number of differentially expressed genes of all genomes with Cbi addition, perhaps suggesting a preference for Cbi. *Serratia* also showed a number of differentially expressed genes with B12 and pB12 addition. Of dependents that did not have paired RNA data, *Bacteriovorax* and *Flavobacterium* each only appeared in the third replicate. Thus, dependent genomes showed differentially expressed genes 3 hours after corrinoid addition that were corrinoid specific. However, differentially expressed gene changes per corrinoid condition did not portend growth differences per condition at 2 weeks. Many dependents were equivalently abundant with or without different corrinoid additions.

Given that the majority of MAGs revealed distinct patterns of significantly differentially expressed genes based on corrinoid addition, we endeavored to understand to what extent these genes were involved in corrinoid-dependent metabolisms or biosynthesis. In addition to the role corrinoids serve as enzymatic cofactors, corrinoids are also known to bind riboswitches and regulate gene expression (Kennedy et al., 2022). Corrinoid riboswitches are often found upstream of corrinoid-dependent enzymes, corrinoid biosynthesis genes, and corrinoid-independent processes for which a corrinoid-dependent enzyme also exists (Degnan, Barry, et al., 2014; Kennedy et al., 2022). We hypothesized that addition of a corrinoid would downregulate corrinoid independent homologs/analogues and upregulate corrinoid-dependent versions of these pathways. We also predicted that we may see transporters downregulated with the supplementation of corrinoid (Degnan, Barry, et al., 2014). In predicted producer genomes, we expected that, given the expected cost of corrinoid production (Sokolovskaya et al., 2020), we would see organisms turn to environmentally provided corrinoids rather than produce corrinoids *de novo*. Previous research has shown that the majority of corrinoid producers maintain corrinoid transport systems (Degnan, Barry, et al., 2014), so like dependents, producers may use exogenous corrinoids.

The vast majority of detected differentially expressed genes did not play a known role in corrinoid metabolism or biosynthesis. Of the total genes significantly differentially expressed 12% were ribosomal proteins, 10% were transporters, 5% were transcriptional regulators. We observed three instances of corrinoid-related metabolisms differentially regulated with the addition of corrinoid (Figure 10). Two were independent RNRs in the same corrinoid-dependent genomes that contained corrinoid transporters. Both corrinoid-independent RNRs in *Serratia* and *Pantoea* showed significant downregulation in the presence of a corrinoid. The downregulation of their expression of corrinoid independent RNRs, 3 h after corrinoid addition, also appears to confirm the ability to import exogenous corrinoids. Both of these *Proteobacteria* appear to downregulate their independent RNRs in response to cobinamide, B12, and pB12, though significant differential expression is not detected in all conditions. However, *Serratia* does not appear to downregulate its RNR in response to Cre whereas *Pantoea brenneri* may, suggesting their independent RNRs differ

in corrinoid-specificities in regulatory elements, perhaps in upstream riboswitches (Kennedy et al., 2022; Rosnow et al., 2018). Interestingly, a corrinoid-dependent RNR was not detected in either of these genomes.

We also observed a corrinoid-dependent methionine synthase in the predicted dependent MAG, *Edaphocola*, that was, perhaps counterintuitively, upregulated in no corrinoid condition compared to corrinoid addition, and specifically was significantly downregulated with the addition of B12. There have been previous reports in pure culture research of corrinoids repressing regulation of corrinoid-dependent enzymes (Kennedy et al., 2022), but there remains no emerging theory to explain these observations. This *Edaphocola* genome seemed to be a corrinoid or nucleotide auxotroph given it only encoded a corrinoid-dependent RNR. Irrespective of differential expression, we observed that corrinoids are likely produced by predicted corrinoid-producer genomes since metatranscriptomes mapped to biosynthesis genes.

4.3 Discussion

We analyzed community composition and function with respect to corrinoids as a model metabolite in the soil microbiome taking a reductive approach – first examining corrinoid sharing capacity and treatment impact in soil and next investigating corrinoid-based community dynamics in a soil enrichment culture. We found soil MAGs maintained the potential to use and produce corrinoids during a highly dynamic moment in annually dry grassland soil. We thus tested whether corrinoids could impact microbial composition in soil through addition of distinct corrinoid types and found the treatment effect to be immediate (at 3 days) and transient. From our findings of function and impact of corrinoids in soil microbes we developed a reduced system to disentangle corrinoid-based microbial interactions. Our reduced complexity system of soil-derived enrichment cultures with subsequent corrinoid addition enabled insights on how soil microbes respond to and use dissimilatory metabolites in a community context. We found that our predictions based on previous pure culture research did not necessarily scale to our observations of how corrinoids impacted communities. We determined that corrinoid treatment impacted individual organisms differently and further individual microbes' corrinoid preferences may not necessarily lead to composition differences, that there may be a degree of community resilience to corrinoid addition in culture not similarly observed in soil, and that corrinoid addition and type may have secondary impacts on communities.

Individual microbial corrinoid preferences impact communities

We hypothesized that the unique corrinoid preferences of organisms shown in pure culture studies (Kennedy et al., 2022; Matthews, 2009; J. Yan et al., 2016; Yi et al., 2012) would show a community wide impact in the form of changes in community composition and on the scale of differential gene expression. From the treatment effect of corrinoid addition in soil, and smaller treatment effects observed in our soil enrichment cultures we found that, indeed, corrinoid addition and corrinoid type impact community composition. We were also able to glimpse in our metagenomic relative abundance and expression data how individual corrinoid preferences may alter community composition at-large with different MAG relative abundance dependent on the type of corrinoid added. For instance in the case of the predicted corrinoid producer,

Herbaspirillum sp004563925, which potentially showed specificity for and higher abundance with exogenous B12 addition, compared to the predicted corrinoid dependent MAG, *Stenotrophomonas* sp002836635, which showed a promiscuity for corrinoid type via equally increased abundance in all corrinoid addition conditions compared to no corrinoid, but showed a higher amount of significantly differentially expressed genes with B12 addition. Thus, corrinoids can differentially regulate genes of organisms in a community context. Interestingly, these expression differences do not have an obvious correspondence with compositional changes or known corrinoid function. Perhaps, this is due to a degree of resilience of communities to corrinoid addition.

Community resilience to corrinoid addition in culture

While we observed significant changes in soil microbial community composition in response to corrinoid addition, the effects in soil enrichment cultures were not as pronounced. Our metatranscriptomic data expression data clearly showed that different corrinoid treatment led to differential expression of genes, but these expression differences at 3 hours did not translate into or portend compositional changes one week after corrinoid addition. While in soil microcosms, corrinoid addition led to significant compositional changes at each taxonomic level, 16S rDNA amplicon sequencing of enrichments did not reveal similarly significant treatment impacts at 3 hours or 1 week following corrinoid addition. Enrichment microbial communities were much more resilient to compositional changes with different corrinoid additions than soil communities. Perhaps, in liquid culture as compared to in soil, corrinoid producers were able to provision sufficient corrinoid to meet corrinoid dependent genomes' requirements, and without the same spatial constraints as in soil, dependent organisms perhaps experienced a sufficient pool of accessible corrinoids without additional exogenous corrinoid. This compositional resilience also may be due to organisms' capacity to remodel corrinoids i.e. replace a corrinoid's lower ligand for preferential use by an organism (Hazra et al., 2013; Mok et al., 2020; Shelton et al., 2019; Yi et al., 2012), and would suggest organisms of our soil enrichments may be highly versatile in which corrinoids they use.

We hypothesized that, as shown in co-culture (Kazamia et al., 2012), addition of corrinoid could decouple interactions between producers and dependents and lead to a compositional shift to predominantly corrinoid dependent organisms. However, our analyses of composition by ecological role showed that corrinoid addition did not alter the relative abundance of producer or dependent genomes or the overall producer to dependent community composition. This underscores our findings that our soil enrichment communities were resilient when challenged with corrinoid additions. In our system, the observed decoupling of corrinoid-based interactions in co-cultures did not scale to a multi-member soil-derived culture.

Another explanation of the observed resilience to large-scale compositional changes of predicted producers and dependent organisms could be that corrinoids are a co-limiting nutrient. Previous research showed that in a marine environment B12 was a co-limiting nutrient with iron (Bertrand et al., 2007). Particularly when looking at predicted corrinoid dependent organisms' relative abundance with corrinoid added compared to without corrinoid there are several organisms where there appears to be a higher abundance without corrinoid added (*Stenotrophomonas*, *Pedobacter*, *Edaphocola*, *Pedobacter* sp004364225). For *Edaphocola*, in particular, given its downregulation by B12 of its corrinoid-dependent methionine synthase, this may also be due to suppression as

certain corrinoid types may also suppress or impair function (Men et al., 2014; Mok & Taga, 2013). Thus, co-limitation or growth suppression by corrinoids may prevent apparent compositional changes in the abundance of predicted producers and dependents with corrinoid addition.

Predicting corrinoid producers and ability to sense exogenous corrinoids

Our study of soil enrichment metagenomes relied on genome annotations to delineate producer and dependent organisms. In defining corrinoid potential producers we used marker genes identified in Shelton et al., 2019 (Shelton et al., 2019) and only 5 MAGs contained at least one of these marker genes. However, other organisms categorized as potential corrinoid dependent MAGs encoded genes involved in *de novo* corrinoid biosynthesis. Perhaps, for this subset of dependents this signifies the capacity to exchange and modify corrinoid precursors. Even in Shelton et al., 2019 (Shelton et al., 2019), 17 % of complete bacterial genomes analyzed contained partial corrinoid biosynthesis pathways and these often lacked complete modules of the production pathway. Further, we were not able to detect corrinoid transporters in all dependent MAGs, though protein sequence alone has not always proved a reliable method for annotating corrinoid transporters (Degnan, Barry, et al., 2014). While significant downregulation of expression of corrinoid independent RNR in two dependent Proteobacterial genomes corresponded with detection of transporters, we could not determine to what extent corrinoids specifically differentially regulate other dependent genomes. Previous research has shown that organisms capable of *de novo* corrinoid biosynthesis also encode functional transporters (Degnan, Barry, et al., 2014) suggesting that organisms may preferentially use exogenous corrinoids rather than expend the energy in *de novo* production. However, we did not detect corrinoid transporters in our predicted producer MAGs.

Downstream and secondary impact of corrinoid additions

While corrinoids serve as a cofactor in a diverse array of metabolic processes and regulate a vast set of known processes (Rosnow et al., 2018; Roth et al., 1996; Shelton et al., 2019; Sokolovskaya et al., 2020), the extent of corrinoids used by microbes as a cofactor or as a regulator remains unknown (Romine, Rodionov, Maezato, Anderson, et al., 2017; Rosnow et al., 2018; Sherlock & Breaker, 2020). Specifically, of genes we observed to be significantly differentially expressed, it appeared we only found three instances of transcript level changes in corrinoid-related metabolisms. To what extent other differentially expressed genes may be regulated by corrinoids directly or indirectly remains unknown. We may not appreciate the full breadth and impact of corrinoids on gene regulation in our soil-derived enrichments. We also observed a resilience in our enrichment communities to corrinoid addition that did not align with our hypothesis that addition of corrinoids could release a nutritional limitation and allow dependent organisms to outpace producers in growth. This observed resilience may also underscore other nutritional exchanges or tertiary interactions maintaining sampled MAGs as interconnected.

Conclusion

Our study combined 16S rDNA amplicon sequencing in soil and soil-derived enrichments with metagenomics and metatranscriptomics to study the impact of corrinoids, widely used cofactors, on soil microbial interactions. We hypothesized that previously observed organisms' specificity

for distinct corrinoid types would predict compositional community changes following addition of specific corrinoids to soil microcosms and soil-derived enrichment microbial communities. In soil, we observed this to be true: soil treated with corrinoids exhibited significant but transient impacts, with Factor A, Cre, pB12, or 5'-hydroxybenzimidazolylcobamide shifting the community. B12 and cobinamide addition had little effect. In soil-derived enrichment cultures, the corrinoid type had a stronger effect on the abundances of bacteria from rarer phyla, but corrinoid addition did not significantly change the overall composition of enrichments in terms of the total abundance of predicted corrinoid producer and predicted corrinoid dependent organisms. However, a subset of individual organisms showed distinct responses to different corrinoid types. Organisms such as *Herbaspirillum sp004563925* increased in abundance more in response to B12 addition and exhibited differential expression of a suite of genes and others, such as *Serratia*, showed a high degree of differentially expressed genes with cobinamide treatment but this did not correspond to greater abundance. Another observed pattern was organisms that increased in abundance with corrinoid addition in a non-specific manner, suggestive of corrinoid promiscuity (e.g., *Stenotrophomonas sp002836635*). Overall, we found that methionine synthesis is the predominant corrinoid-dependent process and that that addition of a trace non-assimilated nutrient can have transient community-wide impacts. The simplified system thus provides a unique framework for understanding the impacts of a model metabolite on soil microbial abundance and expression in community.

4.4 Methods

Direct corrinoid additions to soil microcosms

Soil was sampled from the top 10 cm of a seasonally dry grassland at the Hopland Research and Extension Center (HREC) on April 22, 2021, sieved in the field to 2 mm, and brought back to the lab where it was prewetted to 20% field moisture. In order to test whether direct addition of corrinoids to soil influenced community composition we set up 155 1g soil microcosms to be destructively harvested. Our microcosms represented five timepoints: day 0 with 22 μ L dH₂O addition or no addition, day 3, day 10, day 30, and day 50. To each 1 g soil microcosm, 22 μ L containing 10 nM of one of 6 corrinoids or water alone was added to each of five replicates of all time points except day 0 resulting in a final corrinoid concentration of 220 pmol/g soil. We added one of the following structurally distinct corrinoids: cobalamin (B12), cobinamide (Cbi), [2-methyl]adeninylcobamide (Factor A), Cresolylcobamide (Cre), pseudocobalamin (pB12), 5'-hydroxybenzimidazolylcobamide ([5-OHBza]Cba). Commercially available cobalamin was purchased from MilliporeSigma. Non-commercially available corrinoids were produced as described in (Crofts et al., 2014; Sokolovskaya et al., 2019). At each time point, soil microcosms were transferred to a -80°C freezer for subsequent DNA extraction.

Soil-derived enrichment cultures

To establish soil-derived enrichment cultures, 3 g of soil collected from HREC was added to a sterile flask and suspended in 1x PBS supplemented with 2.24 mM sodium pyrophosphate (25 mL). The soil suspension was mixed on a magnetic stir plate for 30 minutes at room temperature, and allowed to settle for 30 minutes. To a flask with modified M9 medium (50 mL, composition

below), 1 mL of soil suspension was added. Samples were prepared in triplicate, along with a sample containing uninoculated medium. Triplicate aliquots (0.5 mL) of each sample were next collected in 1.7 mL microfuge tubes at 3 and 7 days (72/168 h) post-inoculation for nucleotide extractions. Aliquots were centrifuged at 20,817 g for 5 minutes, the supernatant was removed by suction, and frozen at -80 °C. At 168 h post-inoculation, the enrichment was supplemented with modified M9-medium (5.3 mL). Triplicate aliquots were again taken at 169 hours post-inoculation.

To shift our established soil-derived enrichment cultures into corrinoid media at 169 hours post initial inoculation we transferred 0.5 mL samples of each of the three enrichments (and the uninoculated control) to a 14-mL polypropylene round-bottom tube (Falcon) containing modified M9-medium (4.5 mL) and corrinoid (if applicable, to 10 nM final concentration) in triplicate. Each replicate was supplemented with one of: B12, Cbi, [Cre]Cba, pB12, or no corrinoid. Samples were incubated at room temperature and aliquots (3x0.5 mL) taken at 172 h and 336 h post-enrichment inoculation.

Nucleotide extraction from soil and enrichment cultures

For all DNA extractions of soil microcosms and soil-derived enrichment cultures we used the Qiagen DNeasy PowerSoil Pro Kit (250) (Catalog Number: 47016). For enrichment cultures we added solution CD1 directly to 0.5 mL of the enrichment culture and carried out subsequent extraction steps as described. For RNA extractions we used the Qiagen RNeasy Mini Kit (Catalog Number: 74104) with the on-column DNase digestion using RNase-Free DNase Set (Catalog Number: 79254) and Sigma 425-600 μ m acid-washed glass beads (G8772). To lyse cells during RNA extraction we used a FastPrep FP120 (MP Biomedicals), using the following settings: 5.0 m/s for 20 s (first pass), 5-min rest period on ice, and 4.5 m/s for 20 s (second pass) as described in Beltrame et al. 2015 (Beltrame et al., 2015). All extractions were conducted alongside water-only extraction controls which yielded undetected DNA or RNA concentrations and failed sequencing quality controls during library preparation.

Community composition assessed by 16S rDNA amplicon sequencing

DNA template was added to a PCR reaction to amplify the V4/V5 16S gene region using the 515F/296R primers based on the Earth Microbiome Project primers (Parada et al., 2016; Quince et al., 2011) but with in-line dual Illumina indexes (Price et al., 2018; Sharpless et al., 2022). The amplicons were sequenced on an Illumina MiSeq with 600 bp v3 reagents. Reads were processed with custom Perl scripts implementing Pear for read merging (J. Zhang et al., 2014), USearch (Edgar, 2010) for filtering reads with more than one expected error, demultiplexing using inline indexes and Unoise (Edgar, 2016) for filtering rare reads and chimeras. 16S sequences in the relative abundance table were searched against the RDP database (Cole et al., 2014) to assign taxonomy.

Sequencing, assembly, establishing MAGs, and readmapping

Metagenomic and metatranscriptomic sequencing was performed only on samples from our soil-derived enrichment cultures. All metagenomic and metatranscriptomic 150 paired end shotgun sequencing was performed by Novogene on an Illumina sequencing platform. For details about

the soil metagenomics analyzed in this study see Chapter 3. Metagenomic sequencing yielded 16.7 - 22.1 Gbp raw data and metatranscriptomic sequencing yielded 8.4 - 15.3 Gbp raw data. For metagenomes and metatranscriptomes, data was trimmed and cleaned as described in Chapters 2 and 3. For metagenomic assembly, triplicate samples of the same time point and corrinoid addition were co-assembled using `idba_ud` with 64 threads, using the “pre-correction” flag and minimum kmer size of 40 and maximum of 120. Samples were prepped for upload to ggKbase as previously described in chapters 2 and 3. To bin metagenome-assembled genomes we mapped individual sample read files to assemblies using `bowtie2` with default parameters. We calculated the total reads per sample mapping to assemblies as ranging from 42% to 53%. We used the `jgi_summarize_bam_contig_depths` script to summarize the reads per sample from bam files that mapped to assemblies. The output was used for binning using `metabat` and `MaxBin` (Kang et al., 2015; Y.-W. Wu et al., 2014). We also binned MAGs using `CONCOCT` (Alneberg et al., 2014). We used `DASTool` to choose representative bins which calculated 106 high quality bins (Sieber et al., 2018). To dereplicate bins across assemblies we used `dRep` with the secondary cluster ANI threshold of 0.98 (Olm et al., 2017). We further manually refined dereplicated bins on ggKbase (ggkbase.berkeley.edu) to reduce contamination and remove scaffolds from bin with outlier normalized GC and coverage. We ran `checkM` on these manually curated and dereplicated bins to proceed with the 28 MAGs predicted to be $\geq 85\%$ complete with $\leq 15\%$ contamination (Parks et al., 2015). To establish MAG taxonomy we used the GTDB classify workflow and to annotate genomes we ran `DRAM` (Chaumeil et al., 2020; Shaffer et al., 2020).

To map metagenomic reads to MAGs we used `Bowtie2` (Langmead & Salzberg, 2012). To filter reads to only those with a minimum percent identity of 95 % and to calculate MAG coverage per sample we used `coverM` (github.com/wwood/CoverM). We found that, on aggregate, total reads per sample mapping to dereplicated MAGs represented 5 (3 day sample) to 37 % (14 day sample) of total reads. We used `bowtie2` to map metatranscriptomic reads to dereplicated MAGs and `featureCounts` with `prodigal` gff files to establish read counts per ORF (Liao et al., 2014). Resulting metatranscriptomic read counts per MAG ORF were used to create a sample by gene ID matrix for `DEseq` input. `DEseq` was run to compare differential expression of genes between no corrinoid and corrinoid addition (B12, Cbi, pB12, Cre) treatments 3 hours after corrinoid addition in established enrichment cultures (Love et al., 2014).

Analysis and data visualization

All analyses were performed using Python 3.8.2 (docs.python.org/3.8/) and `pandas` version 1.5.1 (The Pandas Development Team, 2022). Graphs were generated using `seaborn` version 0.11.0 (Waskom, 2021). All ordination plots examining corrinoid treatment impacts on community composition proxied by 16S rDNA amplicon sequencing in soil and in enrichment cultures were performed as described in chapter 3. To test corrinoid treatment significance we ran `ANOSIM` tests using the `scikit-bio` Python package on Bray-Curtis distance matrices from our amplicon sequencing data, iterating through each phylogenetic level (Virtanen et al., 2020).

Corrinoid annotations in MAGs and subsequent ecological role analyses

In order to establish whether each MAG was a predicted corrinoid producer, predicted corrinoid user, or predicted to be corrinoid independent we searched `DRAM` (Shaffer et al., 2020) `pfam` and

KOfam (Aramaki et al., 2020) annotations for corrinoid biosynthesis and dependent functions enumerated in Shelton et al., 2019 (Shelton et al., 2019). We searched the KEGG database (genome.jp/kegg/kegg2.html) and InterPro (ebi.ac.uk/interpro/) for pfam and KOfam numbers that aligned with genes in Shelton et al., 2019 (Shelton et al., 2019). We then used these corrinoid-involved HMMs to search DRAM output. Predicted corrinoid producers were defined as having at least one of the 6 aerobic or anaerobic marker genes established in Shelton et al. 2019. Predicted corrinoid dependent genomes encoded at least one corrinoid-dependent pfam or KOfam. Predicted corrinoid independent MAGs encoded no biosynthesis markers or corrinoid-dependent enzyme annotations.

Having defined each MAG based on its ecological role with respect to corrinoids, we analyzed relative abundance data to understand how predicted corrinoid producer/dependent community composition changed through time and with corrinoid amendment. We calculated relative abundance based on coverM (github.com/wwood/CoverM) reads mapped to a given MAG divided by the total reads generated per sample and filtered with a 50% breadth threshold per genome. First we looked at the relative abundance taking the mean of MAGs predicted to be corrinoid producers, dependents, or independents. Next, we aggregated relative abundance for MAGs, grouping MAGs by predicted corrinoid producer, dependent, or independent to test whether corrinoid addition impacted the composition of the total predicted producer and dependent MAGs.

To explore how corrinoid additions impact MAG gene expression we assessed significantly differentially expressed genes from our DESeq output (adjusted p value ≤ 0.05). We summed the total significantly differentially expressed genes per genome and normalized these counts by genome size (bp) to assess how many genes were differentially expressed per MAG per corrinoid condition. To understand whether differentially expressed genes were involved in corrinoid biosynthesis, were corrinoid dependent, or were corrinoid independent homologs/analogs we aligned differentially expressed ORFs with our corrinoid biosynthesis and dependent genes previously defined. We plotted the three resulting ORFs based on their DESeq calculated normalized gene counts per corrinoid treatment.

4.5 Figures

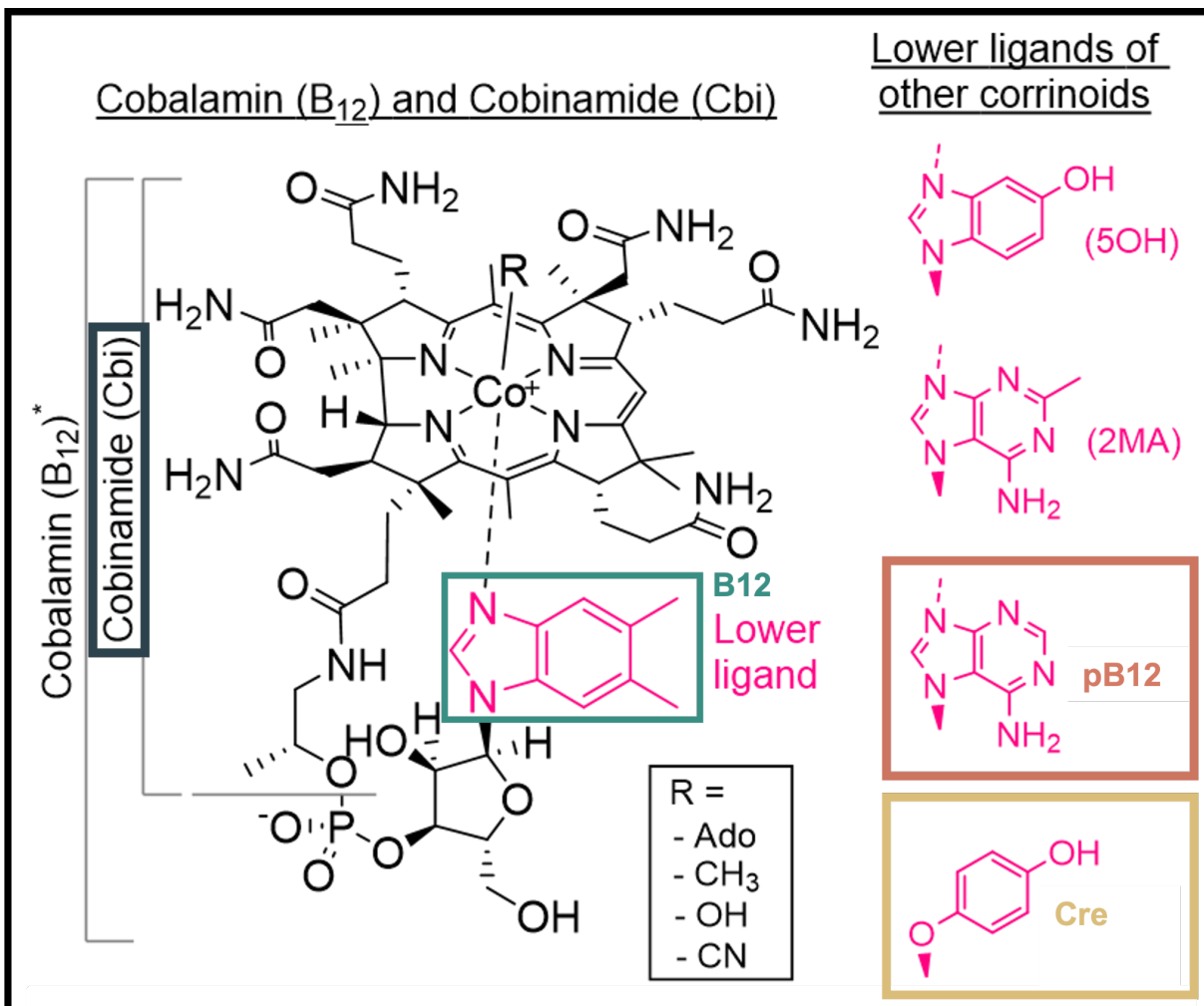


Figure 1 | Corrinoid structure overview. Shown on left is cobalamin (B₁₂), a complete corrinoid that contains dimethylbenzimidazole as the lower ligand. Corrinoids contain one of four upper ligands, represented by R that coordinate a cobalt ion, and may contain a lower ligand. Four examples of lower ligands are shown to the right. Each colored box represents a corrinoid added to the soil-enrichment cultures (Cresolycobamide or Cre, pseudocobalamin or pB₁₂, cobinamide or Cbi and cobalamin or B₁₂).

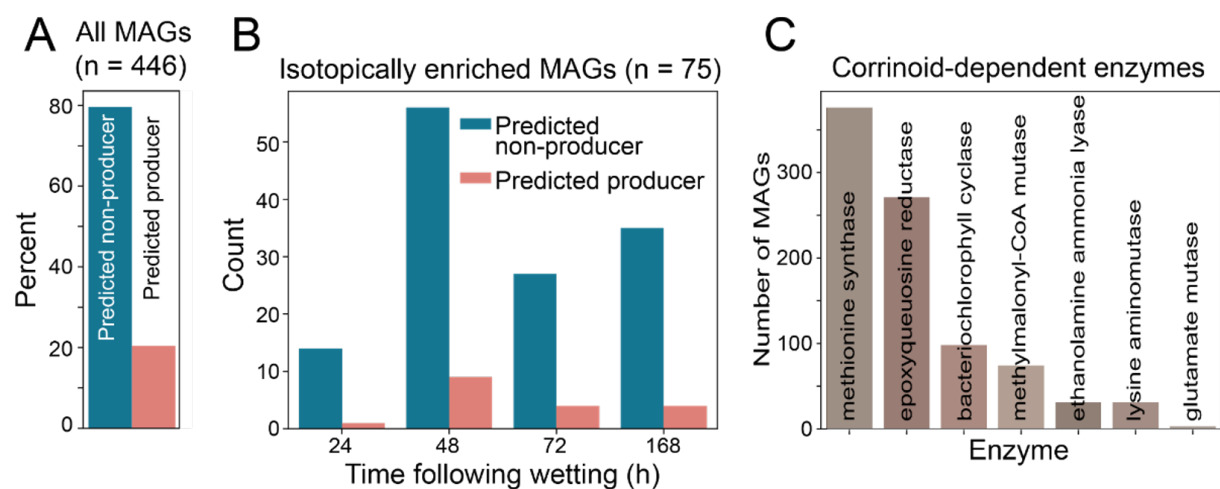


Figure 2 | Predicted corrinoid production and dependent metabolism in soil MAGs. (A) Out of 446 MAGs the percent predicted producer and non-producer. (B) Of 75 detected isotopically enriched MAGs, the count of predicted producers and non-producers through time following soil rewetting. (C) Counts of observed corrinoid-dependent enzymes detected in 446 soil MAGs.

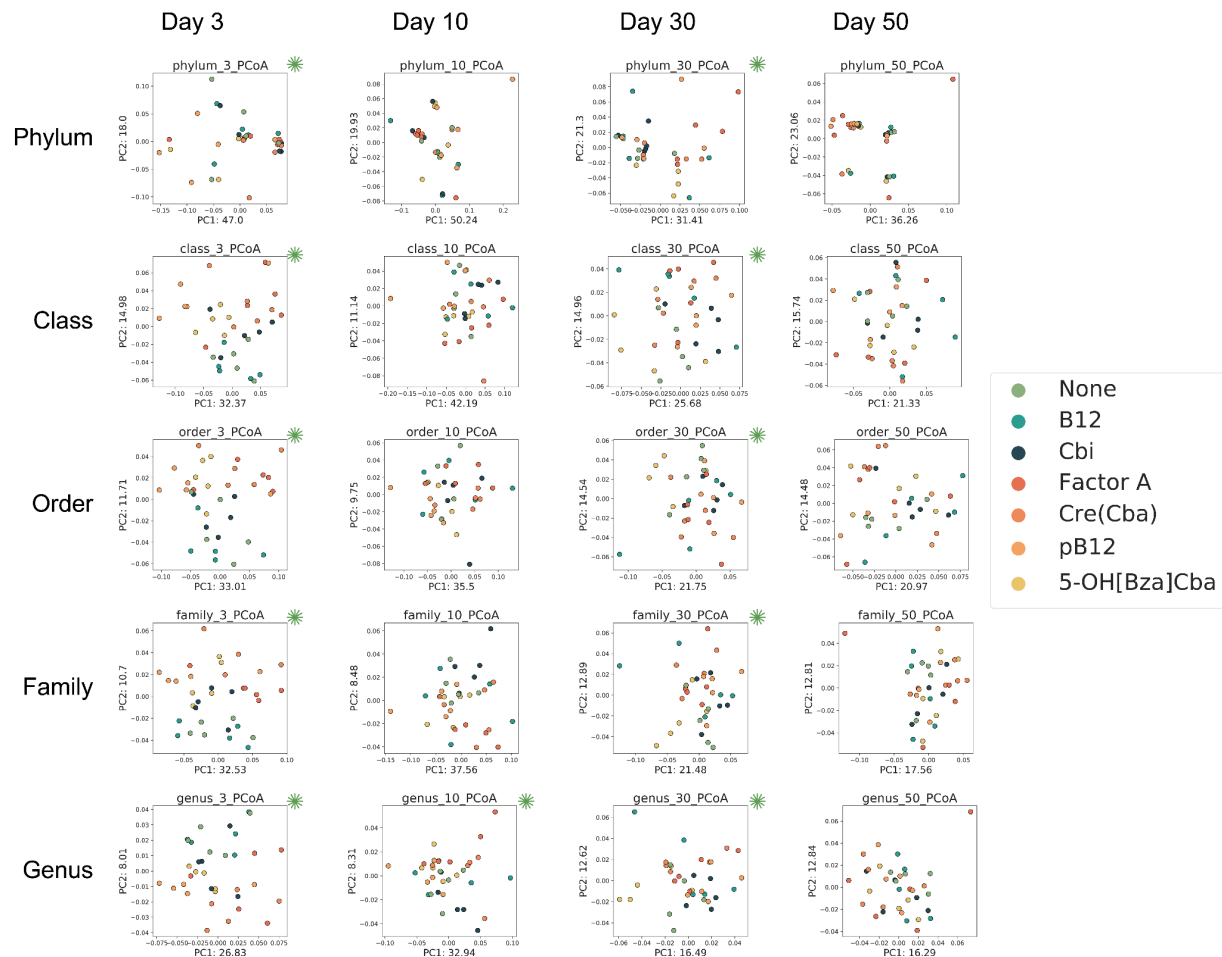


Figure 3 | Temporal microbial composition changes in soil microcosms following corrinoid addition. Each plot represents a PCoA of the Bray-Curtis dissimilarity of zOTUs aggregated at the phylum, class, order, family, or genus taxonomic level per time point (3-50 days) following corrinoid addition. Axes indicate variance explained by each of the first two principal components. Colors indicate corrinoid type added to each soil microcosm. Five replicates per treatment were sampled. Green asterisks denote statistical significance by ANOSIM test (p -value ≤ 0.05).

Treatments	Timepoints (days)	16S	Metagenome	Metatranscriptome
No corrinoid	3 (72h)	3	3	3
No corrinoid	7 (168h)	3	3	3
No corrinoid	7 + 1 hour (169h)	3	3	0
Corrinoid addition				
No corrinoid	7 + 4 hours (172h)	3	0	3
Cbi	7 + 4 hours (172h)	3	0	3
B12	7 + 4 hours (172h)	3	0	3
pB12	7 + 4 hours (172h)	3	0	2
Cre(Cba)	7 + 4 hours (172h)	3	0	2
No corrinoid	14	3	3	0
Cbi	14	3	3	0
B12	14	3	3	0
pB12	14	3	3	0
Cre(Cba)	14	3	3	0

Table 1 | Soil-derived enrichment and shift to corrinoid addition experimental design. In the columns that indicate sequencing method – 16S, metagenome, metatranscriptome – counts reflect the number of replicates sampled per treatment.

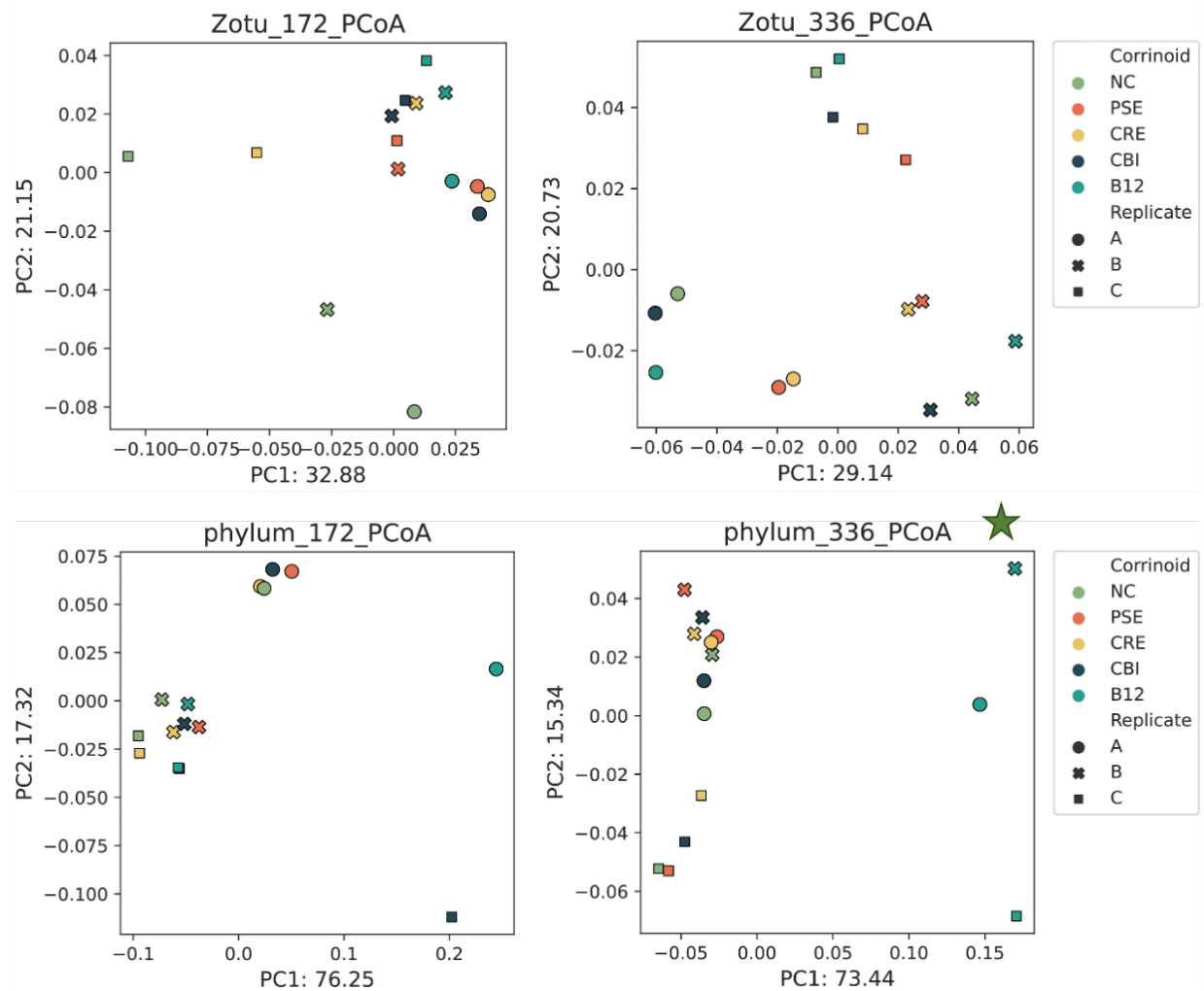


Figure 4 | Microbial community composition in soil-derived enrichments following corrinoid addition. All graphs are PCoAs of Bray-Curtis dissimilarity aggregated at the zOTU level (top graphs) or phylum level (bottom graphs). The left graphs are 172 hours of enrichment which is 3 hours after corrinoid addition; right graphs are 336 hours of enrichment, 1 week after corrinoid addition. Colors represent corrinoid type, shapes represent replicate. Green star signifies statistical significance based on ANOSIM test (p -value ≤ 0.05).

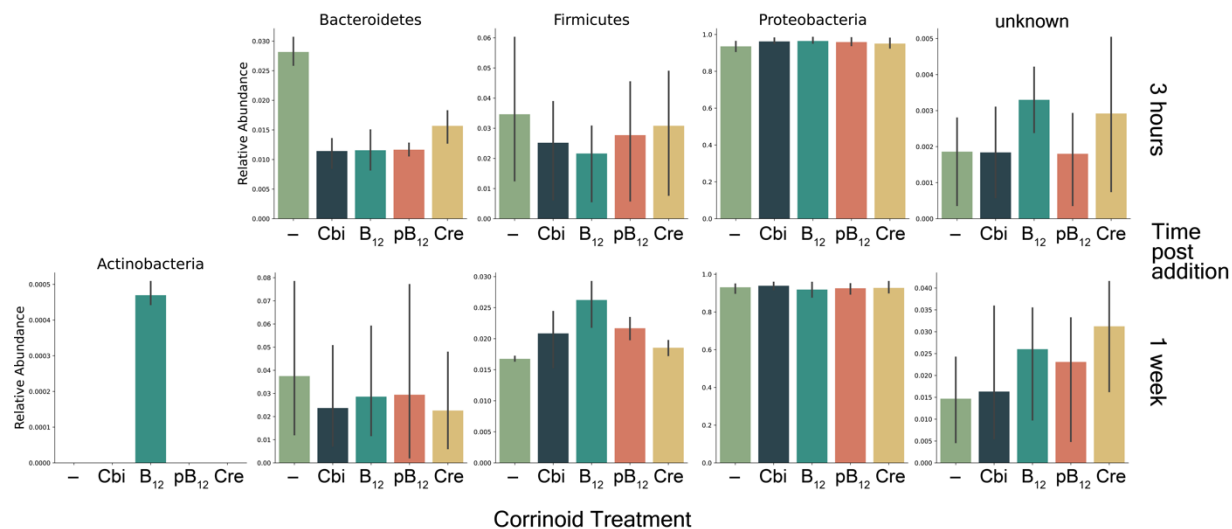


Figure 5 | Phylum level view of soil enrichments based on 16S rDNA sequencing. Each plot represents the relative abundance of a given phylum per corrinoid treatment at 3 hours (top row) and 1 week (bottom row) following corrinoid addition. Note the differences in scale in the y axes.

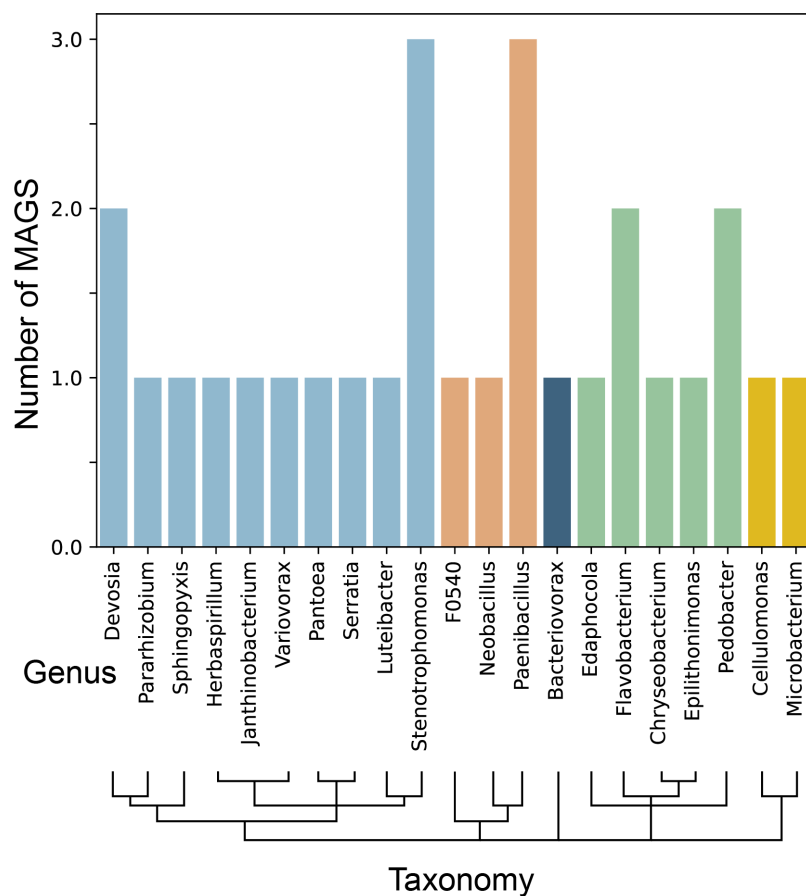


Figure 6 | Genus level taxonomy of soil enrichment MAGs. Light blue bars represent Proteobacteria, orange are Firmicutes, dark blue is Bdellovibrionota, green are Bacteroidota, and yellow are Actinobacteria. Beneath the x axis, representing genera of MAGs sampled, is a diagram illustrating the higher order taxonomic relationships among genera – the deepest nodes represent MAGs of the same family, the shallowest represent phyla.

phyla	class	order	family	genus	species
Proteobacteria	Alphaproteobacteria	Rhizobiales	Devosiaceae	Devosia	
Proteobacteria	Alphaproteobacteria	Rhizobiales	Devosiaceae	Devosia_A	
Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Pararhizobium	Pararhizobium sp900473045
Proteobacteria	Gammaproteobacteria	Burkholderiales	Burkholderiaceae	Herbaspirillum	Herbaspirillum sp004563925
Firmicutes	Bacilli	Paenibacillales	Paenibacillaceae	Paenibacillus	

Table 2 | Predicted corrinoid producer taxonomy. Taxonomic classifications of the five MAGs predicted to contain the capacity for *de novo* biosynthesis of corrinoids are shown.

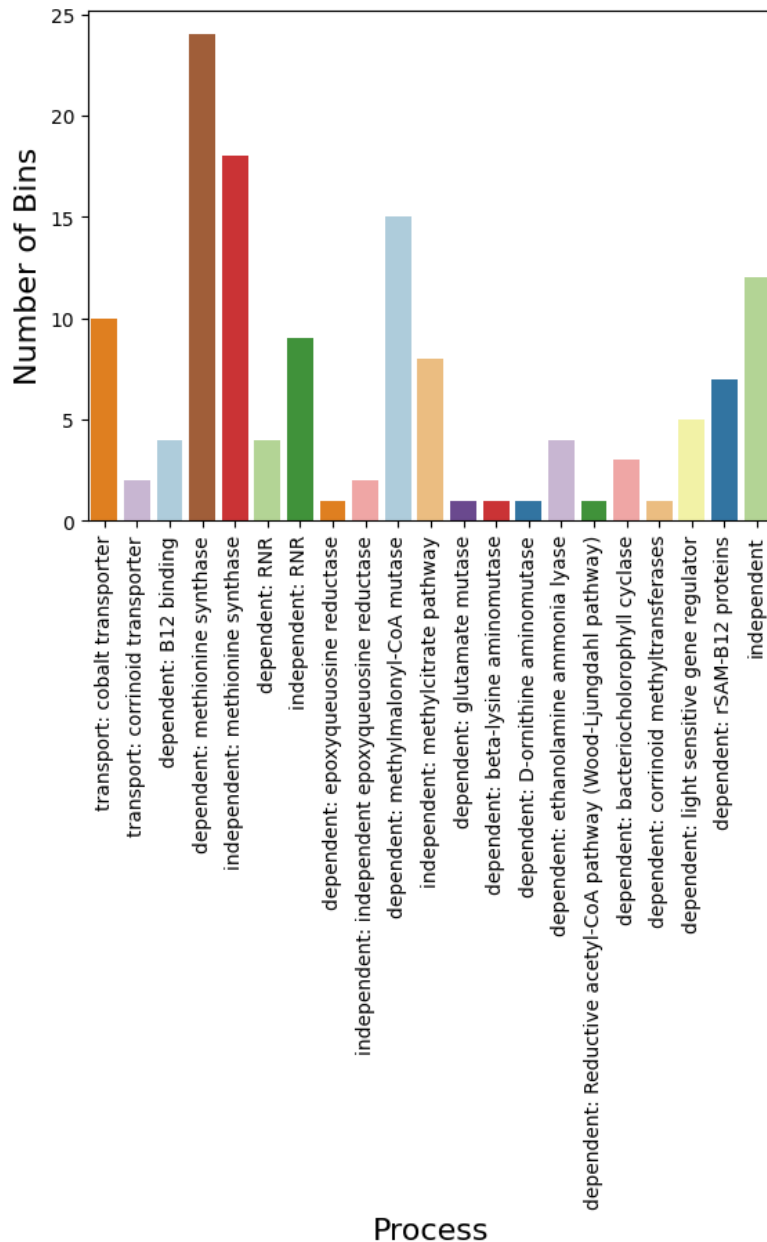


Figure 7 | Corrinoid-dependent and -independent processes in soil enrichment MAGs. The number of bins in the combined metagenomic dataset containing predicted corrinoid transporters, corrinoid-dependent enzymes, and corrinoid-independent alternative enzymes is shown.

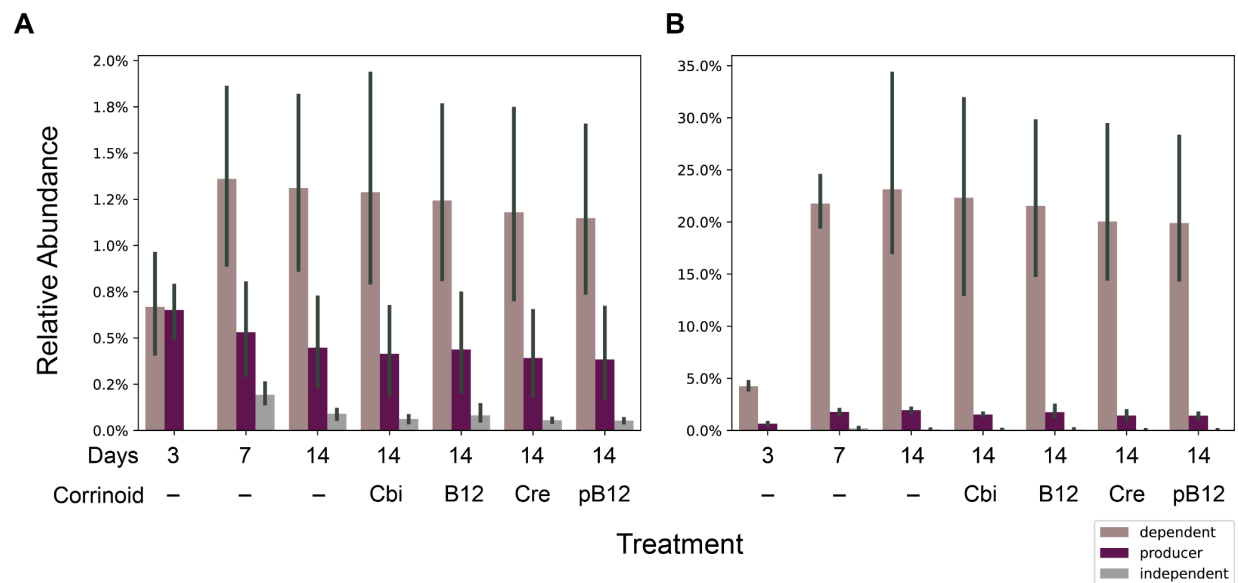


Figure 8 | Composition of soil enrichment cultures based on predicted corrinoid ecological role. (A) The mean relative abundance per MAG is shown. **(B)** The aggregated relative abundance based on a MAG's corrinoid-based ecological role.



Figure 9 | Corrinoid impact on relative abundance and differential gene expression. Each plot represents, per genome (colored by phylum), the fold differential relative abundance (upper rows) calculated by the MAG relative abundance per corrinoide condition compared to the relative abundance without corrinoide added at 1 week post corrinoide addition. Lower rows of graphs per MAG show the suite of significantly differentially expressed genes in a corrinoide condition compared to no corrinoide (adjusted p -value ≤ 0.05) at 3 hours post corrinoide addition calculated per base pair MAG. Of note: for Cre and pB₁₂ treatments, metatranscriptomes were produced in duplicate, not triplicate. Results are split by predicted corrinoide producers, predicted corrinoide dependents, and independents. Where there is no bottom row for MAGs (no corresponding plot of total differential expression), no genes were detected as significantly differentially expressed.

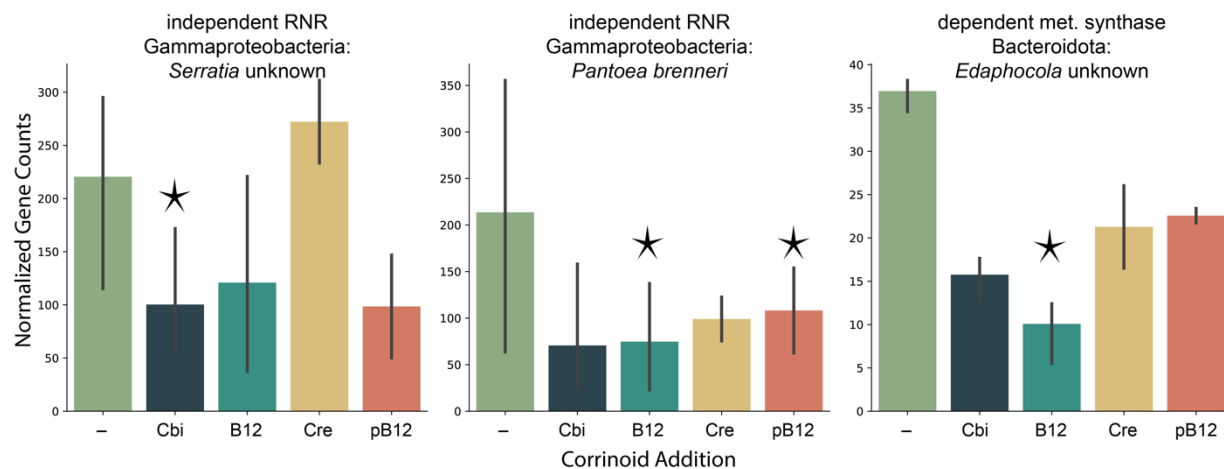


Figure 10 | Differential gene expression of corrinoid regulated processes. Each graph represents an ORF with a predicted corrinoid-related function that showed significant differential gene expression in at least one corrinoid treatment condition compared to no corrinoid added, at 3 hours after corrinoid addition. First two graphs are corrinoid independent ribonucleotide reductases (RNRs), last graph is for a corrinoid-dependent methionine synthase.

Concluding remarks

This dissertation seeks to disentangle interactions in the soil microbiome and move forth the field of soil microbial ecology from cataloging who is in soil and how organisms may live, towards how microbes may generate, distribute, and share resources. The past decade has marked an extensive effort by scientists to enumerate the soil microbiome (Delgado-Baquerizo et al., 2018; Paez-Espino et al., 2016), and more recently has begun to work toward understanding the role of microbes in major nutritional cycles (Diamond et al., 2019; Woodcroft et al., 2018) and diverse functional capacities (Crits-Christoph et al., 2018; Sharrar et al., 2020). Analyses of the soil microbiome have largely focused on the net possible set of interactions in the form of networks generated through microbial co-occurrence patterns and how specific environmental treatments impact network complexity (Baker et al., 2022; Ling et al., 2022; Yuan et al., 2021). Simultaneously, the painstaking work of soil microbial isolations to study how organisms differentially use soil substrates and ground truth functional capacities predicted in genomes continues (Baran et al., 2015; da Rocha et al., 2015; Zhelnina et al., 2018).

In sum, my research connects distinct aspects of the soil microbiome through the perspective that interactions can be inferred in genomic capacities. Through studying CPR bacteria and DPANN, the ultrasmall cells that are rare members of the biosphere, and emphasizing their reduced genomic inventory, I highlight how these overlooked lineages may serve an important role in cycling nutrients through the microbial loop (Pomeroy et al., 2007). I quantified the contribution of viruses to soil microbial death during a rapid moment of cellular lysis, and elucidated the dynamics of viruses on a fine timescale, which furthers a more comprehensive understanding of the role of viruses in soil. In using corrinoids as a model nutrient to probe the soil microbiome, I observed how a non-assimilated nutrient can have broader impacts on soil than on soil-enrichment culture communities and that the effects of corrinoid addition target microbes differently at the level of RNA and DNA.

Each of these research projects has involved developing new methods and frameworks to studying the soil microbiome. I streamlined a protocol for generating small fraction metagenomes that enrich for membrane-bound or encapsulated genomes smaller than 200 nanometers which enabled access to the nanoscale sized microbes of soil, simultaneously capturing more viruses (100x), and resolving a greater number of genomes of CPR and DPANN than extracting DNA directly from soil. Importantly, other studies using similar soil protocols computationally ignore small bacteria and archaea in downstream analyses. I then combined these viromic protocols with SIP metagenomics to uniquely analyze the isotopically enriched, or active, soil viruses. Here, an interesting open question remains – how do small fraction metagenomes and traditional metagenomic extractions differentially sample soil viruses. Finally, in developing a reduced approach to study nutritional interactions in the soil microbiome – moving from soil to soil treated with our model nutrient to simplified soil communities in culture with our nutrient added – I demonstrated how a complex community can be simplified to analyze the effects of individual key metabolites.

As the field of soil microbial ecology moves from cataloging our soils' microbiomes to testing interactions, the field will contend with many remaining challenges. An understated component of soil's complexity is its opacity which has led to the development of transparent soils and other

systems for real-time imaging of microbe-microbe and microbe-plant interactions (Downie et al., 2012; Sasse et al., 2019; K. Sharma et al., 2020). However, the challenge of isolating soil microbes (Steen et al., 2019) remains a major barrier in testing interactions *in vitro*. Transparent soil and other newly developed methods to image and analyze interactions still rely on soil isolates. For instance, kChip is a platform for high throughput arraying of isolates in co-cultures in various media conditions, which despite its high throughput capacity necessitates isolation campaigns (Kehe et al., 2021). Similarly, dilution to extinction experiments to reduce soil microbes to multi-member consortia depend on culture conditions to test interactions (Ossowicki et al., 2021). Enrichment cultures, too, as used in the fourth chapter, require choice carbon and nutrient conditions (Goldford et al., 2018), but provide a path forward for deconstructing and reconstructing communities (Coker et al., 2022), especially for community editing to target specific microbes in these contexts (Rubin et al., 2021). In the near future, the field will develop a deeper understanding of microbial interactions through cultivation-based experiments, but will ultimately be limited by the task of isolation itself.

Beyond physical methods to test microbial interactions, there remain many difficulties in linking organisms *in silico*. As discussed in the third chapter, linking a specific virus to a host genome in soil can only be performed with high confidence in very few instances. Further, research has seen that CPR live in close association with a host (C. He et al., 2021; X. He et al., 2015), but ascertaining host from metagenomic data alone remains an untested hypothesis (Jaffe et al., 2021). In terms of the nutritional interactions studied in chapter four, we still do not understand whether corrinoid producers and dependent microbes are coupled. This may rely on how accessible corrinoids are throughout soil, whether they are exported out of corrinoid producing organisms or whether cell death maintains an accessible corrinoid pool, i.e., to what extent corrinoid producers act as specific providers to other organisms.

Ultimately, merging soil microbial ecology with mechanistic chemical ecology to understand microbial interactions will have to contend with the scales of soil. The studies that comprise this dissertation amass thousands of unique soil aggregates. We will need to develop a collective understanding and new set of methodologies to ask what happens not in the aggregated collection of soil pores, but what happens in each unique soil aggregate and what are the scales of distinct interactions within aggregates. We will need to zoom into soil further to unravel the experience of a microbe in its native environment.

Early in my doctoral studies, I assumed I could unify all these still disparate elements of the soil microbiome – viruses, bacteria, nutrients. I believed I could link viruses to their specific hosts and approximate microbial cell death, predict the subsequent release of newly accessible nutrients into the environment, and hypothesize how that would impact neighboring organisms. I wonder now, how far into the future we may have this resolution in such a complex microbial ecosystem as soil.

References

- Abramson, J., Riistama, S., Larsson, G., Jasaitis, A., Svensson-Ek, M., Laakkonen, L., Puustinen, A., Iwata, S., & Wikström, M. (2000). The structure of the ubiquinol oxidase from *Escherichia coli* and its ubiquinone binding site. *Nature Structural Biology*, 7(10), 910–917. <https://doi.org/10.1038/82824>
- Abreu, N. A., & Taga, M. E. (2016). Decoding molecular interactions in microbial communities. *FEMS Microbiology Reviews*, 40(5), 648–663. <https://doi.org/10.1093/femsre/fuw019>
- Abrudan, M. I., Smakman, F., Grimbergen, A. J., Westhoff, S., Miller, E. L., van Wezel, G. P., & Rozen, D. E. (2015). Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proceedings of the National Academy of Sciences*, 112(35), 11054–11059. <https://doi.org/10.1073/pnas.1504076112>
- Acinas, S. G., Marcelino, L. A., Klepac-Ceraj, V., & Polz, M. F. (2004). Divergence and Redundancy of 16S rRNA Sequences in Genomes with Multiple *rrn* Operons. *Journal of Bacteriology*, 186(9), 2629–2635. <https://doi.org/10.1128/JB.186.9.2629-2635.2004>
- Ackermann, H.-W. (2013). Sad State of Phage Electron Microscopy. Please Shoot the Messenger. *Microorganisms*, 2(1), 1–10. <https://doi.org/10.3390/microorganisms2010001>
- Al-Shayeb, B., Sachdeva, R., Chen, L. X., Ward, F., Munk, P., Devoto, A., Castelle, C. J., Olm, M. R., Bouma-Gregson, K., Amano, Y., He, C., Méheust, R., Brooks, B., Thomas, A., Lavy, A., Matheus-Carnevali, P., Sun, C., Goltsman, D. S. A., Borton, M. A., ... Banfield, J. F. (2020). Clades of huge phages from across Earth's ecosystems. *Nature*, 578(7795), 425–431. <https://doi.org/10.1038/s41586-020-2007-4>
- Albright, M. B. N., Gallegos-Graves, L. V., Feeser, K. L., Montoya, K., Emerson, J. B., Shakya, M., & Dunbar, J. (2022). Experimental evidence for the impact of soil viruses on carbon cycling during surface plant litter decomposition. *ISME Communications*, 2(1), 24. <https://doi.org/10.1038/s43705-022-00109-4>
- Allen, R. H., & Stabler, S. P. (2008). Identification and quantitation of cobalamin and cobalamin analogues in human feces. *American Journal of Clinical Nutrition*, 87(5), 1324–1335. <https://doi.org/10.1016/j.biotechadv.2011.08.021>. Secreted
- Alneberg, J., Bjarnason, B. S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., Lahti, L., Loman, N. J., Andersson, A. F., & Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, 11(11), 1144–1146. <https://doi.org/10.1038/nmeth.3103>
- Alteio, L. V., Schulz, F., Seshadri, R., Varghese, N., Rodriguez-Reillo, W., Ryan, E., Goudeau, D., Eichorst, S. A., Malmstrom, R. R., Bowers, R. M., Katz, L. A., Blanchard, J. L., & Woyke, T. (2020). Complementary Metagenomic Approaches Improve Reconstruction of Microbial Diversity in a Forest Soil. *MSystems*, 5(2), 1–18. <https://doi.org/10.1128/msystems.00768-19>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Andersson, A. F., & Banfield, J. F. (2008). Virus population dynamics and acquired virus resistance in natural microbial communities. *Science*, 320(5879), 1047–1050. <https://doi.org/10.1126/science.1157358>
- Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., Ogata, H., & Valencia, A. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM

- and adaptive score threshold. *Bioinformatics*, 36(7), 2251–2252.
<https://doi.org/10.1093/bioinformatics/btz859>
- Auslander, N., Gussow, A. B., Benler, S., Wolf, Y. I., & Koonin, E. V. (2020). Seeker: Alignment-free identification of bacteriophage genomes by deep learning. *Nucleic Acids Research*, 48(21), E121. <https://doi.org/10.1093/nar/gkaa856>
- Averill, C., Anthony, M. A., Baldrian, P., Finkbeiner, F., van den Hoogen, J., Kiers, T., Kohout, P., Hirt, E., Smith, G. R., & Crowther, T. W. (2022). Defending Earth’s terrestrial microbiome. *Nature Microbiology*, 7(11), 1717–1725. <https://doi.org/10.1038/s41564-022-01228-3>
- Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., Bengtsson-Palme, J., Anslan, S., Coelho, L. P., Harend, H., Huerta-Cepas, J., Medema, M. H., Maltz, M. R., Munda, S., Olsson, P. A., Pent, M., Pölme, S., Sunagawa, S., Ryberg, M., ... Bork, P. (2018). Structure and function of the global topsoil microbiome. *Nature*, 560(7717), 233–237. <https://doi.org/10.1038/s41586-018-0386-6>
- Baker, N. R., Zhalnina, K., Yuan, M., Herman, D., Ceja-Navarro, J. A., Sasse, J., Jordan, J. S., Bowen, B. P., Wu, L., Fossum, C., Chew, A., Fu, Y., Saha, M., Zhou, J., Pett-Ridge, J., Northen, T. R., & Firestone, M. (2022). Nutrient and moisture limitation reveal keystone metabolites that link switchgrass rhizosphere metabolome and microbiome dynamics. *BioRxiv*, 2022.06.20.496911.
<https://www.biorxiv.org/content/10.1101/2022.06.20.496911v1%0Ahttps://www.biorxiv.org/content/10.1101/2022.06.20.496911v1.abstract>
- Banerjee, S., & van der Heijden, M. G. A. (2022). Soil microbiomes and one health. In *Nature Reviews Microbiology* (Vol. 21, Issue 1, pp. 6–20). <https://doi.org/10.1038/s41579-022-00779-w>
- Baran, R., Brodie, E. L., Mayberry-Lewis, J., Hummel, E., Da Rocha, U. N., Chakraborty, R., Bowen, B. P., Karaoz, U., Cadillo-Quiroz, H., Garcia-Pichel, F., & Northen, T. R. (2015). Exometabolite niche partitioning among sympatric soil bacteria. *Nature Communications*, 6, 8289. <https://doi.org/10.1038/ncomms9289>
- Barnard, R. L., Blazewicz, S. J., & Firestone, M. K. (2020). Rewetting of soil: Revisiting the origin of soil CO₂ emissions. *Soil Biology and Biochemistry*, 147(April), 107819. <https://doi.org/10.1016/j.soilbio.2020.107819>
- Barnard, R. L., Osborne, C. A., & Firestone, M. K. (2013). Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *The ISME Journal*, 7(11), 2229–2241. <https://doi.org/10.1038/ismej.2013.104>
- Barnard, R. L., Osborne, C. A., & Firestone, M. K. (2015). Changing precipitation pattern alters soil microbial community response to wet-up under a Mediterranean-type climate. *The ISME Journal*, 9(4), 946–957. <https://doi.org/10.1038/ismej.2014.192>
- 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(1). <https://doi.org/10.1128/msphere.00024-20>
- Beltrame, C. O., Côrtes, M. F., Bandeira, P. T., & Figueiredo, A. M. S. (2015). Optimization of the RNeasy Mini Kit to obtain high-quality total RNA from sessile cells of *Staphylococcus aureus*. *Brazilian Journal of Medical and Biological Research*, 48(12), 1071. <https://doi.org/10.1590/1414-431X20154734>
- Bertrand, E. M., Saito, M. A., Rose, J. M., Riesselman, C. R., Lohan, M. C., Noble, A. E., Lee,

- P. A., & DiTullio, G. R. (2007). Vitamin B12 and iron colimitation of phytoplankton growth in the Ross Sea. *Limnology and Oceanography*, 52(3), 1079–1093. <https://doi.org/10.4319/lo.2007.52.3.1079>
- Bi, L., Yu, D., Du, S., Zhang, L., Zhang, L., Wu, C., Xiong, C., Han, L., & He, J. (2020). Diversity and potential biogeochemical impacts of viruses in bulk and rhizosphere soils. *Environmental Microbiology*, 00, 1462–2920.15010. <https://doi.org/10.1111/1462-2920.15010>
- Birch, H. F. (1958). The effect of soil drying on humus decomposition and nitrogen availability. *Plant and Soil*, 10(1), 9–31. <https://doi.org/10.1007/BF01343734>
- Bland, C., Ramsey, T. L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N. C., & Hugenholtz, P. (2007). CRISPR Recognition Tool (CRT): A tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics*, 8(1), 209. <https://doi.org/10.1186/1471-2105-8-209>
- Blazewicz, S. J., Hungate, B. A., Koch, B. J., Nuccio, E. E., Morrissey, E., Brodie, E. L., Schwartz, E., Pett-Ridge, J., & Firestone, M. K. (2020). Taxon-specific microbial growth and mortality patterns reveal distinct temporal population responses to rewetting in a California grassland soil. *ISME Journal*, 14(6), 1520–1532. <https://doi.org/10.1038/s41396-020-0617-3>
- Blazewicz, S. J., Schwartz, E., & Firestone, M. K. (2014). Growth and death of bacteria and fungi underlie rainfall-induced carbon dioxide pulses from seasonally dried soil. *Ecology*, 95(5), 1162–1172. <https://doi.org/10.1890/13-1031.1>
- Boopathi, S., Liu, D., & Jia, A. Q. (2021). Molecular trafficking between bacteria determines the shape of gut microbial community. In *Gut Microbes* (Vol. 13, Issue 1). Taylor and Francis Ltd. <https://doi.org/10.1080/19490976.2021.1959841>
- Braga, R. M., Dourado, M. N., & Araújo, W. L. (2016). Microbial interactions: ecology in a molecular perspective. In *Brazilian Journal of Microbiology* (Vol. 47, pp. 86–98). Elsevier. <https://doi.org/10.1016/j.bjm.2016.10.005>
- Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A., & Fierer, N. (2016). Genome reduction in an abundant and ubiquitous soil bacterium “Candidatus Udaeobacter copiosus.” *Nature Microbiology*, 2(2), 16198. <https://doi.org/10.1038/nmicrobiol.2016.198>
- Brown, C. T., Hug, L. A., Thomas, B. C., Sharon, I., Castelle, C. J., Singh, A., Wilkins, M. J., Wrighton, K. C., Williams, K. H., & Banfield, J. F. (2015). Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature*, 523(7559), 208–211. <https://doi.org/10.1038/nature14486>
- Brum, J. R., Hurwitz, B. L., Schofield, O., Ducklow, H. W., & Sullivan, M. B. (2016). Seasonal time bombs: dominant temperate viruses affect Southern Ocean microbial dynamics. *The ISME Journal*, 10(2), 437–449. <https://doi.org/10.1038/ismej.2015.125>
- Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doulier, G., Acinas, S. G., Alberti, A., Chaffron, S., Cruaud, C., de Vargas, C., Gasol, J. M., Gorsky, G., Gregory, A. C., Guidi, L., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Poulos, B. T., ... Sullivan, M. B. (2015). Patterns and ecological drivers of ocean viral communities. *Science*, 348(6237), 1261498. <https://doi.org/10.1126/science.1261498>
- Burstein, D., Sun, C. L., Brown, C. T., Sharon, I., Anantharaman, K., Probst, A. J., Thomas, B. C., & Banfield, J. F. (2016). Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. *Nature Communications*, 7, 10613. <https://doi.org/10.1038/ncomms10613>

- Butaite, E., Baumgartner, M., Wyder, S., & Kümmerli, R. (2017). Siderophore cheating and cheating resistance shape competition for iron in soil and freshwater *Pseudomonas* communities. *Nature Communications*, 8(1), 1–12. <https://doi.org/10.1038/s41467-017-00509-4>
- Campbell, K., Herrera-Dominguez, L., Correia-Melo, C., Zelezniak, A., & Ralser, M. (2018). Biochemical principles enabling metabolic cooperativity and phenotypic heterogeneity at the single cell level. *Current Opinion in Systems Biology*, 8, 97–108. <https://doi.org/10.1016/J.COISB.2017.12.001>
- Carini, P., Marsden, P. J., Leff, J. W., Morgan, E. E., Strickland, M. S., & Fierer, N. (2016). Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nature Microbiology*, 2, 16242. <https://doi.org/10.1038/nmicrobiol.2016.242>
- Carlson, H. K., Lui, L. M., Price, M. N., Kazakov, A. E., Carr, A. V., Kuehl, J. V., Owens, T. K., Nielsen, T., Arkin, A. P., & Deutschbauer, A. M. (2020). Selective carbon sources influence the end products of microbial nitrate respiration. *ISME Journal*, 14(8), 2034–2045. <https://doi.org/10.1038/s41396-020-0666-7>
- Castelle, C. J., & Banfield, J. F. (2018). Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. In *Cell* (Vol. 172, Issue 6, pp. 1181–1197). Cell Press. <https://doi.org/10.1016/j.cell.2018.02.016>
- Castelle, C. J., Brown, C. T., Anantharaman, K., Probst, A. J., Huang, R. H., & Banfield, J. F. (2018). Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nature Reviews Microbiology*, 16(10), 629–645. <https://doi.org/10.1038/s41579-018-0076-2>
- Chamkhi, I., El Omari, N., Benali, T., & Bouyahya, A. (2020). Quorum Sensing and Plant-Bacteria Interaction: Role of Quorum Sensing in the Rhizobacterial Community Colonization in the Rhizosphere [Chapter]. In *ACS Symposium Series* (Vol. 1374, pp. 139–153). American Chemical Society. <https://doi.org/10.1021/bk-2020-1374.ch008>
- Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2020). GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*, 36(6), 1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>
- Chow, C.-E. T., Kim, D. Y., Sachdeva, R., Caron, D. A., & Fuhrman, J. A. (2014). Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *The ISME Journal*, 8(4), 816–829. <https://doi.org/10.1038/ismej.2013.199>
- Chuckran, P. F., Fofanov, V., Hungate, B. A., Morrissey, E. M., Schwartz, E., Walkup, J., & Dijkstra, P. (2021). Rapid Response of Nitrogen Cycling Gene Transcription to Labile Carbon Amendments in a Soil Microbial Community. *MSystems*, 6(3). <https://doi.org/10.1128/msystems.00161-21>
- Chuckran, P. F., Hungate, B. A., Schwartz, E., & Dijkstra, P. (2022). Variation in genomic traits of microbial communities among ecosystems. *FEMS Microbes*, 2(August), 1–13. <https://doi.org/10.1093/femsmc/xtab020>
- Clarholm, M. (1985). Interactions of bacteria, protozoa and plants leading to mineralization of soil nitrogen. *Soil Biology and Biochemistry*, 17(2), 181–187. [https://doi.org/10.1016/0038-0717\(85\)90113-0](https://doi.org/10.1016/0038-0717(85)90113-0)
- Coker, J., Zhalnina, K., Marotz, C., Thiruppathy, D., Tjuanta, M., D’Elia, G., Hailu, R., Mahosky, T., Rowan, M., Northen, T. R., & Zengler, K. (2022). A Reproducible and Tunable Synthetic Soil Microbial Community Provides New Insights into Microbial

- Ecology. *MSystems*, 2022.05.19.492707. <https://doi.org/10.1128/msystems.00951-22>
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R., & Tiedje, J. M. (2014). Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, 42(D1), 633–642. <https://doi.org/10.1093/nar/gkt1244>
- Coleman-Derr, D., & Tringe, S. G. (2014). Building the crops of tomorrow: advantages of symbiont-based approaches to improving abiotic stress tolerance. *Frontiers in Microbiology*, 5, 283. <https://doi.org/10.3389/fmicb.2014.00283>
- Cook, B. I., Mankin, J. S., & Anchukaitis, K. J. (2018). Climate Change and Drought: From Past to Future. *Current Climate Change Reports*, 4(2), 164–179. <https://doi.org/10.1007/s40641-018-0093-2>
- Correa, A. M. S., Howard-Varona, C., Coy, S. R., Buchan, A., Sullivan, M. B., & Weitz, J. S. (2021). Revisiting the rules of life for viruses of microorganisms. *Nature Reviews Microbiology*, 0123456789. <https://doi.org/10.1038/s41579-021-00530-x>
- Couradeau, E., Sasse, J., Goudeau, D., Nath, N., Hazen, T. C., Bowen, B. P., Chakraborty, R., Malmstrom, R. R., & Northen, T. R. (2019). Probing the active fraction of soil microbiomes using BONCAT-FACS. *Nature Communications*, 10(1), 2770. <https://doi.org/10.1038/s41467-019-10542-0>
- 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>
- Crits-Christoph, A., Gelsinger, D. R., Ma, B., Wierzchos, J., Ravel, J., Davila, A., Casero, M. C., & DiRuggiero, J. (2016). Functional interactions of archaea, bacteria and viruses in a hypersaline endolithic community. *Environmental Microbiology*, 18(6), 2064–2077. <https://doi.org/10.1111/1462-2920.13259>
- Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J., & Smith, A. G. (2005). Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature*, 438(7064), 90–93. <https://doi.org/10.1038/nature04056>
- Crofts, T. S., Hazra, A. B., Tran, J. La, Sokolovskaya, O. M., Osadchiy, V., Ad, O., Pelton, J., Bauer, S., & Taga, M. E. (2014). Regiospecific formation of cobamide isomers is directed by CobT. *Biochemistry*, 53(49), 7805–7815. <https://doi.org/10.1021/bi501147d>
- Crowther, T. W., Thomas, S. M., Maynard, D. S., Baldrian, P., Covey, K., Frey, S. D., van Diepen, L. T. A., & Bradford, M. A. (2015). Biotic interactions mediate soil microbial feedbacks to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 112(22), 7033–7038. <https://doi.org/10.1073/pnas.1502956112>
- Crowther, T. W., Todd-Brown, K. E. O., Rowe, C. W., Wieder, W. R., Carey, J. C., Machmuller, M. B., Snoek, L., Sokol, W., Templer, P. H., Treseder, K. K., Welker, J. M., & Bradford, M. A. (2016). Quantifying global soil carbon losses in response to warming. *Nature*, 540. <https://doi.org/10.1038/nature20150>
- D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., & Kost, C. (2018). Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports*, 35(5), 455–488. <https://doi.org/10.1039/C8NP00009C>
- da Rocha, U. N., Cadillo-Quiroz, H., Karaoz, U., Rajeev, L., Klitgord, N., Dunn, S., Truong, V., Buenrostro, M., Bowen, B. P., Garcia-Pichel, F., Mukhopadhyay, A., Northen, T. R., & Brodie, E. L. (2015). Isolation of a significant fraction of non-phototroph diversity from a desert biological soil crust. *Frontiers in Microbiology*, 6(MAR), 277.

- <https://doi.org/10.3389/fmicb.2015.00277>
- Dal Co, A., van Vliet, S., Kiviet, D. J., Schlegel, S., & Ackermann, M. (2020). Short-range interactions govern the dynamics and functions of microbial communities. *Nature Ecology & Evolution*, 4(3), 366–375. <https://doi.org/10.1038/s41559-019-1080-2>
- DeAngelis, K. M., Lindow, S. E., & Firestone, M. K. (2008). Bacterial quorum sensing and nitrogen cycling in rhizosphere soil. *FEMS Microbiology Ecology*, 66(2), 197–207. <https://doi.org/10.1111/j.1574-6941.2008.00550.x>
- Degnan, P. H., Barry, N. A., Mok, K. C., Taga, M. E., & Goodman, A. L. (2014). Human Gut Microbes Use Multiple Transporters to Distinguish Vitamin B12 Analogs and Compete in the Gut. *Cell Host & Microbe*, 15(1), 47–57. <https://doi.org/10.1016/j.chom.2013.12.007>
- Degnan, P. H., Taga, M. E., & Goodman, A. L. (2014). Vitamin B12 as a modulator of gut microbial ecology. In *Cell Metabolism* (Vol. 20, Issue 5, pp. 769–778). NIH Public Access. <https://doi.org/10.1016/j.cmet.2014.10.002>
- Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K., & Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. *Science*, 359(6373), 320–325. <https://doi.org/10.1126/science.aap9516>
- Delmont, T. O., Robe, P., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., Hirsch, P. R., & Vogel, T. M. (2011). Accessing the soil metagenome for studies of microbial diversity. *Applied and Environmental Microbiology*, 77(4), 1315–1324. <https://doi.org/10.1128/AEM.01526-10>
- Diamond, S., Andeer, P. F., Li, Z., Crits-Christoph, A., Burstein, D., Anantharaman, K., Lane, K. R., Thomas, B. C., Pan, C., Northen, T. R., & Banfield, J. F. (2019). Mediterranean grassland soil C–N compound turnover is dependent on rainfall and depth, and is mediated by genomically divergent microorganisms. *Nature Microbiology*, 4(8), 1356–1367. <https://doi.org/10.1038/s41564-019-0449-y>
- Diffenbaugh, N. S., Swain, D. L., Touma, D., & Lubchenco, J. (2015). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13), 3931–3936. <https://doi.org/10.1073/pnas.1422385112>
- Downie, H., Holden, N., Otten, W., Spiers, A. J., Valentine, T. A., & Dupuy, L. X. (2012). Transparent Soil for Imaging the Rhizosphere. *PLoS ONE*, 7(9), e44276. <https://doi.org/10.1371/journal.pone.0044276>
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>
- Edgar, R. C. (2016). UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *BioRxiv*, 081257. <https://doi.org/10.1101/081257>
- Embree, M., Liu, J. K., Al-Bassam, M. M., & Zengler, K. (2015). Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 112(50), 15450–15455. <https://doi.org/10.1073/pnas.1506034112>
- Emerson, J. B. (2019). Soil Viruses: A New Hope. *MSystems*, 4(3), e00120-19. <https://doi.org/10.1128/mSystems.00120-19>
- Emerson, J. B., Roux, S., Brum, J. R., Bolduc, B., Woodcroft, B. J., Jang, H. Bin, Singleton, C. M., Solden, L. M., Naas, A. E., Boyd, J. A., Hodgkins, S. B., Wilson, R. M., Trubl, G., Li, C., Frolking, S., Pope, P. B., Wrighton, K. C., Crill, P. M., Chanton, J. P., ... Sullivan, M.

- B. (2018). Host-linked soil viral ecology along a permafrost thaw gradient. In *Nature Microbiology* (Vol. 3, Issue 8, pp. 870–880). <https://doi.org/10.1038/s41564-018-0190-y>
- Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, 15(10), 579–590. <https://doi.org/10.1038/nrmicro.2017.87>
- Fierer, N., & Lennon, J. T. (2011). The generation and maintenance of diversity in microbial communities. *American Journal of Botany*, 98(3), 439–448. <https://doi.org/10.3732/ajb.1000498>
- Fraser, F. C., Corstanje, R., Deeks, L. K., Harris, J. A., Pawlett, M., Todman, L. C., Whitmore, A. P., & Ritz, K. (2016). On the origin of carbon dioxide released from rewetted soils. *Soil Biology and Biochemistry*, 101, 1–5. <https://doi.org/10.1016/j.soilbio.2016.06.032>
- Gallegos-Monterrosa, R., & Coulthurst, S. J. (2021). The ecological impact of a bacterial weapon: Microbial interactions and the Type VI secretion system. In *FEMS Microbiology Reviews* (Vol. 45, Issue 6). Oxford University Press. <https://doi.org/10.1093/femsre/fuab033>
- Gao, Z., Karlsson, I., Geisen, S., Kowalchuk, G., & Jousset, A. (2019). Protists: Puppet Masters of the Rhizosphere Microbiome. *Trends in Plant Science*, 24(2), 165–176. <https://doi.org/10.1016/j.tplants.2018.10.011>
- Garcia, E. C., Perault, A. I., Marlatt, S. A., & Cotter, P. A. (2016). Interbacterial signaling via Burkholderia contact-dependent growth inhibition system proteins. *Proceedings of the National Academy of Sciences*, 113(29), 8296–8301. <https://doi.org/10.1073/pnas.1606323113>
- Ghosh, D., Roy, K., Williamson, K. E., Srinivasiah, S., Wommack, K. E., & Radosevich, M. (2009). Acyl-homoserine lactones can induce virus production in lysogenic bacteria: An alternative paradigm for prophage induction. *Applied and Environmental Microbiology*, 75(22), 7142–7152. <https://doi.org/10.1128/AEM.00950-09>
- Ghosh, D., Roy, K., Williamson, K. E., White, D. C., Wommack, K. E., Sublette, K. L., & Radosevich, M. (2008). Prevalence of lysogeny among soil bacteria and presence of 16S rRNA and trzN genes in viral-community DNA. *Applied and Environmental Microbiology*, 74(2), 495–502. <https://doi.org/10.1128/AEM.01435-07>
- Giovannoni, S. J., Cameron Thrash, J., & Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *The ISME Journal*, 8(8), 1553–1565. <https://doi.org/10.1038/ismej.2014.60>
- Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P., & Sanchez, A. (2018). Emergent simplicity in microbial community assembly. *Science*, 361(6401), 469–474. <https://doi.org/10.1126/science.aat1168>
- Göller, P. C., Haro-Moreno, J. M., Rodriguez-Valera, F., Loessner, M. J., & Gómez-Sanz, E. (2020). Uncovering a hidden diversity: optimized protocols for the extraction of dsDNA bacteriophages from soil. *Microbiome*, 8(1), 17. <https://doi.org/10.1186/s40168-020-0795-2>
- Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiology Spectrum*, 4(1). <https://doi.org/10.1128/microbiolspec.VMBF-0012-2015>
- Greenlon, A., Sieradzki, E. T., Zablocki, O., Koch, B. J., & ... (2022). Quantitative stable-isotope probing (qSIP) with metagenomics links microbial physiology and activity to soil moisture in Mediterranean-climate grassland ecosystems. *BioRxiv*. <https://doi.org/10.1101/2022.05.02.490339>
- Guo, J., Bolduc, B., Zayed, A. A., Varsani, A., Dominguez-Huerta, G., Delmont, T. O., Pratama, A. A., Gazitúa, M. C., Vik, D., Sullivan, M. B., & Roux, S. (2021). VirSorter2: a multi-

- classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome*, 9(1), 37. <https://doi.org/10.1186/s40168-020-00990-y>
- Hallberg, Z. F., Seth, E. C., Thevasundaram, K., & Taga, M. E. (2022). Comparative Analysis of Corrinoid Profiles across Host-Associated and Environmental Samples. *Biochemistry*. <https://doi.org/10.1021/acs.biochem.2c00367>
- Halverson, L. J., Jones, T. M., & Firestone, M. K. (2000). Release of Intracellular Solutes by Four Soil Bacteria Exposed to Dilution Stress. *Soil Science Society of America Journal*, 64(5), 1630–1637. <https://doi.org/10.2136/sssaj2000.6451630x>
- Hampton, H. G., Watson, B. N. J., & Fineran, P. C. (2020). The arms race between bacteria and their phage foes. In *Nature* (Vol. 577, Issue 7790, pp. 327–336). Nature Research. <https://doi.org/10.1038/s41586-019-1894-8>
- Han, L.-L., Yu, D.-T., Zhang, L.-M., Shen, J.-P., & He, J.-Z. (2017). Genetic and functional diversity of ubiquitous DNA viruses in selected Chinese agricultural soils. *Scientific Reports*, 7, 45142. <https://doi.org/10.1038/srep45142>
- Hauser, M., Steinegger, M., & Söding, J. (2016). MMseqs software suite for fast and deep clustering and searching of large protein sequence sets. *Bioinformatics*, 32(9), 1323–1330. <https://doi.org/10.1093/bioinformatics/btw006>
- Hazra, A. B., Tran, J. L. A., Crofts, T. S., & Taga, M. E. (2013). Analysis of substrate specificity in CobT homologs reveals widespread preference for DMB, the lower axial ligand of vitamin B12. *Chemistry and Biology*, 20(10), 1275–1285. <https://doi.org/10.1016/j.chembiol.2013.08.007>
- He, C., Keren, R., Whittaker, M. L., Farag, I. F., Doudna, J. A., Cate, J. H. D., & Banfield, J. F. (2021). Genome-resolved metagenomics reveals site-specific diversity of episymbiotic CPR bacteria and DPANN archaea in groundwater ecosystems. *Nature Microbiology*, 6(3), 354–365. <https://doi.org/10.1038/s41564-020-00840-5>
- He, X., McLean, J. S., Edlund, A., Yooseph, S., Hall, A. P., Liu, S.-Y., Dorrestein, P. C., Esquenazi, E., Hunter, R. C., Cheng, G., Nelson, K. E., Lux, R., & Shi, W. (2015). Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. *Proceedings of the National Academy of Sciences*, 112(1), 244–249. <https://doi.org/10.1073/pnas.1419038112>
- Heal, K. R., Qin, W., Ribalet, F., Bertagnolli, A. D., Coyote-Maestas, W., Hmelo, L. R., Moffett, J. W., Devol, A. H., Armbrust, E. V., Stahl, D. A., & Ingalls, A. E. (2017). Two distinct pools of B 12 analogs reveal community interdependencies in the ocean. *Proceedings of the National Academy of Sciences*, 114(2), 364–369. <https://doi.org/10.1073/pnas.1608462114>
- Helliwell, K. E. (2017). The roles of B vitamins in phytoplankton nutrition: new perspectives and prospects. *New Phytologist*, 216(1), 62–68. <https://doi.org/10.1111/nph.14669>
- Herrera Paredes, S., & Lebeis, S. L. (2016). Giving back to the community: microbial mechanisms of plant–soil interactions. *Functional Ecology*, 30(7), 1043–1052. <https://doi.org/10.1111/1365-2435.12684>
- Howard-Varona, C., Hargreaves, K. R., Abedon, S. T., & Sullivan, M. B. (2017). Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *The ISME Journal*, 11(7), 1511–1520. <https://doi.org/10.1038/ismej.2017.16>
- Hug, L. A., & Co, R. (2018). It Takes a Village: Microbial Communities Thrive through Interactions and Metabolic Handoffs. *MSystems*, 3(2). <https://doi.org/10.1128/msystems.00152-17>
- Hungate, B. A., Marks, J. C., Power, M. E., Schwartz, E., Groenigen, K. J. van, Blazewicz, S. J.,

- Chuckran, P., Dijkstra, P., Finley, B. K., Firestone, M. K., Foley, M., Greenlon, A., Hayer, M., Hofmockel, K. S., Koch, B. J., Mack, M. C., Mau, R. L., Miller, S. N., Morrissey, E. M., ... Pett-Ridge, J. (2021). The functional significance of bacterial predators. *MBio*, 12(2). <https://doi.org/10.1128/mBio.00466-21>
- Hungate, B. A., Mau, R. L., Schwartz, E., Caporaso, J. G., Dijkstra, P., van Gestel, N., Koch, B. J., Liu, C. M., McHugh, T. A., Marks, J. C., Morrissey, E. M., & Price, L. B. (2015). Quantitative microbial ecology through stable isotope probing. *Applied and Environmental Microbiology*, 81(21), 7570–7581. <https://doi.org/10.1128/AEM.02280-15>
- Hurwitz, B. L., Hallam, S. J., & Sullivan, M. B. (2013). Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biology*, 14(11), R123. <https://doi.org/10.1186/gb-2013-14-11-r123>
- Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11, 119. <https://doi.org/10.1186/1471-2105-11-119>
- Ignacio-Espinoza, J. C., Ahlgren, N. A., & Fuhrman, J. A. (2020). Long-term stability and Red Queen-like strain dynamics in marine viruses. *Nature Microbiology*, 5(2), 265–271. <https://doi.org/10.1038/s41564-019-0628-x>
- Iguchi, H., Yurimoto, H., & Sakai, Y. (2011). Stimulation of methanotrophic growth in cocultures by cobalamin excreted by rhizobia. *Applied and Environmental Microbiology*, 77(24), 8509–8515. <https://doi.org/10.1128/AEM.05834-11>
- Jaffe, A. L., Castelle, C. J., Matheus Carnevali, P. B., Gribaldo, S., & Banfield, J. F. (2020). The rise of diversity in metabolic platforms across the Candidate Phyla Radiation. *BMC Biology*, 18(1), 69. <https://doi.org/10.1186/s12915-020-00804-5>
- Jaffe, A. L., Thomas, A. D., He, C., Keren, R., Valentin-Alvarado, L. E., Munk, P., Bouma-Gregson, K., Farag, I. F., Amano, Y., Sachdeva, R., West, P. T., & Banfield, J. F. (2021). Patterns of gene content and co-occurrence constrain the evolutionary path toward animal association in candidate phyla radiation bacteria. *MBio*, 12(4). <https://doi.org/10.1128/mBio.00521-21>
- Jansson, J. K., & Hofmockel, K. S. (2020). Soil microbiomes and climate change. In *Nature Reviews Microbiology* (Vol. 18, Issue 1, pp. 35–46). Nature Research. <https://doi.org/10.1038/s41579-019-0265-7>
- Jiang, X., Zerfaß, C., Feng, S., Eichmann, R., Asally, M., Schäfer, P., & Soyer, O. S. (2018). Impact of spatial organization on a novel auxotrophic interaction among soil microbes. *The ISME Journal*, 12(6), 1443–1456. <https://doi.org/10.1038/s41396-018-0095-z>
- Jiao, S., Yang, Y., Xu, Y., Zhang, J., & Lu, Y. (2020). Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *The ISME Journal*, 14(1), 202–216. <https://doi.org/10.1038/s41396-019-0522-9>
- Johnson, D. R., Nemir, A., Andersen, G. L., Zinder, S. H., & Alvarez-Cohen, L. (2009). Transcriptomic microarray analysis of corrinoid responsive genes in *Dehalococcoides ethenogenes* strain 195. *FEMS Microbiology Letters*, 294(2), 198–206. <https://doi.org/10.1111/j.1574-6968.2009.01569.x>
- Joshi, N., & Fass, J. (2011). *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)*.
- Jover, L. F., Effler, T. C., Buchan, A., Wilhelm, S. W., & Weitz, J. S. (2014). The elemental composition of virus particles: Implications for marine biogeochemical cycles. *Nature Reviews Microbiology*, 12(7), 519–528. <https://doi.org/10.1038/nrmicro3289>

- Jurgensen, S. K., Roux, S., Schwenck, S. M., Stewart, F. J., Sullivan, M. B., & Brum, J. R. (2022). Viral community analysis in a marine oxygen minimum zone indicates increased potential for viral manipulation of microbial physiological state. *The ISME Journal*, 16(4), 972–982. <https://doi.org/10.1038/s41396-021-01143-1>
- Jurkevitch, E., Minz, D., Ramati, B., & Barel, G. (2000). Prey range characterization, ribotyping, and diversity of soil and rhizosphere *Bdellovibrio* spp. isolated on phytopathogenic bacteria. *Applied and Environmental Microbiology*, 66(6), 2365–2371. <https://doi.org/10.1128/aem.66.6.2365-2371.2000>
- 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. *Applied and Environmental Microbiology*, 83(20), e01253-17. <https://doi.org/10.1128/AEM.01253-17>
- Kakouridis, A., Hagen, J. A., Kan, M. P., Mambelli, S., Feldman, L. J., Herman, D. J., Weber, P. K., Pett-Ridge, J., & Firestone, M. K. (2022). Routes to roots: direct evidence of water transport by arbuscular mycorrhizal fungi to host plants. *New Phytologist*, 236(1), 210–221. <https://doi.org/10.1111/nph.18281>
- Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ*, 3(8), e1165. <https://doi.org/10.7717/peerj.1165>
- Kazamia, E., Czesnick, H., Nguyen, T. T. Van, Croft, M. T., Sherwood, E., Sasso, S., Hodson, S. J., Warren, M. J., & Smith, A. G. (2012). Mutualistic interactions between vitamin B12 - dependent algae and heterotrophic bacteria exhibit regulation. *Environmental Microbiology*, 14(6), 1466–1476. <https://doi.org/10.1111/j.1462-2920.2012.02733.x>
- Kehe, J., Ortiz, A., Kulesa, A., Gore, J., Blainey, P. C., & Friedman, J. (2021). Positive interactions are common among culturable bacteria. *Science Advances*, 7(45), 1–11. <https://doi.org/10.1126/sciadv.abi7159>
- Kelly, C. J., Alexeev, E. E., Farb, L., Vickery, T. W., Zheng, L., Eric, C. L., Kitzenberg, D. A., Battista, K. D., Kominsky, D. J., Robertson, C. E., Frank, D. N., Stabler, S. P., & Colgan, S. P. (2019). Gut Microbes Oral vitamin B 12 supplement is delivered to the distal gut, altering the corrinoid profile and selectively depleting *Bacteroides* in C57BL/6 mice Oral vitamin B 12 supplement is delivered to the distal gut, altering the corrinoid profile and selectively depleting *Bacteroides* in C57BL/6 mice. <https://doi.org/10.1080/19490976.2019.1597667>
- Kennedy, K. J., Widner, F. J., Sokolovskaya, O. M., Innocent, L. V., Procknow, R. R., Mok, K. C., & Taga, M. E. (2022). Cobalamin Riboswitches Are Broadly Sensitive to Corrinoid Cofactors to Enable an Efficient Gene Regulatory Strategy. *MBio*, 13(5). <https://doi.org/10.1128/mbio.01121-22>
- Kieft, K., & Anantharaman, K. (2022). Deciphering Active Prophages from Metagenomes. *MSystems*. <https://doi.org/10.1128/msystems.00084-22>
- Kieft, K., Zhou, Z., & Anantharaman, K. (2020). VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome*, 8(1), 90. <https://doi.org/10.1186/s40168-020-00867-0>
- Kimura, M., Jia, Z.-J., Nakayama, N., & Asakawa, S. (2008). Ecology of viruses in soils: Past, present and future perspectives. *Soil Science and Plant Nutrition*, 54(1), 1–32. <https://doi.org/10.1111/j.1747-0765.2007.00197.x>
- Klappenbach, J. A., Dunbar, J. M., & Schmidt, T. M. (2000). rRNA operon copy number reflects

- ecological strategies of bacteria. *Applied and Environmental Microbiology*, 66(4), 1328–1333. <https://doi.org/10.1128/AEM.66.4.1328-1333.2000>
- Knowles, B., Silveira, C. B., Bailey, B. A., Barott, K., Cantu, V. A., Cobian-Guêmes, A. G., Coutinho, F. H., Dinsdale, E. A., Felts, B., Furby, K. A., George, E. E., Green, K. T., Gregoracci, G. B., Haas, A. F., Haggerty, J. M., Hester, E. R., Hisakawa, N., Kelly, L. W., Lim, Y. W., ... Rohwer, F. (2016). Lytic to temperate switching of viral communities. *Nature*, 531(7595), 466–470. <https://doi.org/10.1038/nature17193>
- Koch, B. J., McHugh, T. A., Hayer, M., Schwartz, E., Blazewicz, S. J., Dijkstra, P., Gestel, N., Marks, J. C., Mau, R. L., Morrissey, E. M., Pett-Ridge, J., & Hungate, B. A. (2018). Estimating taxon-specific population dynamics in diverse microbial communities. *Ecosphere*, 9(1). <https://doi.org/10.1002/ecs2.2090>
- Koch, H., van Kessel, M. A. H. J., & Lüscher, S. (2019). Complete nitrification: insights into the ecophysiology of comammox Nitrospira. *Applied Microbiology and Biotechnology*, 103(1), 177–189. <https://doi.org/10.1007/s00253-018-9486-3>
- Kothari, A., Wu, Y.-W., Chandonia, J.-M., Charrier, M., Rajeev, L., Rocha, A. M., Joyner, D. C., Hazen, T. C., Singer, S. W., & Mukhopadhyay, A. (2019). Large Circular Plasmids from Groundwater Plasmidomes Span Multiple Incompatibility Groups and Are Enriched in Multimetal Resistance Genes. *MBio*, 10(1), e02899-18. <https://doi.org/10.1128/MBIO.02899-18>
- Kroeger, M. E., Delmont, T. O., Eren, A. M., Meyer, K. M., Guo, J., Khan, K., Rodrigues, J. L. M., Bohannan, B. J. M., Tringe, S. G., Borges, C. D., Tiedje, J. M., Tsai, S. M., & Nüsslein, K. (2018). New Biological Insights Into How Deforestation in Amazonia Affects Soil Microbial Communities Using Metagenomics and Metagenome-Assembled Genomes. *Frontiers in Microbiology*, 9(JUL), 1635. <https://doi.org/10.3389/fmicb.2018.01635>
- Kuypers, M. M. M., Marchant, H. K., & Kartal, B. (2018). The microbial nitrogen-cycling network. *Nature Reviews Microbiology*, 16(5), 263–276. <https://doi.org/10.1038/nrmicro.2018.9>
- Kuzyakov, Y., & Mason-Jones, K. (2018). Viruses in soil: Nano-scale undead drivers of microbial life, biogeochemical turnover and ecosystem functions. *Soil Biology and Biochemistry*, 127, 305–317. <https://doi.org/10.1016/j.soilbio.2018.09.032>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lavy, A., McGrath, D. G., Matheus Carnevali, P. B., Wan, J., Dong, W., Tokunaga, T. K., Thomas, B. C., Williams, K. H., Hubbard, S. S., & Banfield, J. F. (2019). Microbial communities across a hillslope-riparian transect shaped by proximity to the stream, groundwater table, and weathered bedrock. *Ecology and Evolution*, 9(12), 6869–6900. <https://doi.org/10.1002/ece3.5254>
- Lee, S., Sieradzki, E. T., Nicolas, A. M., Walker, R. L., Firestone, M. K., Hazard, C., & Nicol, G. W. (2021). Methane-derived carbon flows into host-virus networks at different trophic levels in soil. *Proceedings of the National Academy of Sciences of the United States of America*, 118(32), 1–8. <https://doi.org/10.1073/pnas.2105124118>
- Lennon, J. T., & Jones, S. E. (2011). Microbial seed banks: The ecological and evolutionary implications of dormancy. In *Nature Reviews Microbiology* (Vol. 9, Issue 2, pp. 119–130). Nature Publishing Group. <https://doi.org/10.1038/nrmicro2504>
- Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn

- graph. *Bioinformatics*, 31(10), 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liang, X., Wagner, R. E., Li, B., Zhang, N., & Radosevich, M. (2020). Quorum Sensing Signals Alter in vitro Soil Virus Abundance and Bacterial Community Composition. *Frontiers in Microbiology*, 11(June), 1–13. <https://doi.org/10.3389/fmicb.2020.01287>
- Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Ling, N., Wang, T., & Kuzyakov, Y. (2022). Rhizosphere bacteriome structure and functions. *Nature Communications*, 13(1), 836. <https://doi.org/10.1038/s41467-022-28448-9>
- Liu, Z., Wang, K., Nan, X., Yang, L., Wang, Y., Zhang, F., Cai, M., Zhao, Y., & Xiong, B. (2022). Effects of combined addition of 3-nitrooxypropanol and vitamin B12 on methane and propionate production in dairy cows by in vitro-simulated fermentation. *Journal of Dairy Science*. <https://doi.org/10.3168/jds.2022-22207>
- Lochhead, A. G., & Thexton, R. H. (1951). Vitamin B12 as a Growth Factor for Soil Bacteria. *Nature*, 167(4260), 1034–1034. <https://doi.org/10.1038/1671034a0>
- Lombard, N., Prestat, E., van Elsas, J. D., & Simonet, P. (2011). Soil-specific limitations for access and analysis of soil microbial communities by metagenomics. In *FEMS Microbiology Ecology* (Vol. 78, Issue 1, pp. 31–49). Oxford Academic. <https://doi.org/10.1111/j.1574-6941.2011.01140.x>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lu, X., Heal, K. R., Ingalls, A. E., Doxey, A. C., & Neufeld, J. D. (2020). Metagenomic and chemical characterization of soil cobalamin production. *ISME Journal*, 14(1), 53–66. <https://doi.org/10.1038/s41396-019-0502-0>
- Lucas, J. M., McBride, S. G., & Strickland, M. S. (2020). Trophic level mediates soil microbial community composition and function. *Soil Biology and Biochemistry*, 143, 107756. <https://doi.org/10.1016/j.soilbio.2020.107756>
- Luef, B., Frischkorn, K. R., Wrighton, K. C., Holman, H.-Y. N., Birarda, G., Thomas, B. C., Singh, A., Williams, K. H., Siegerist, C. E., Tringe, S. G., Downing, K. H., Comolli, L. R., & Banfield, J. F. (2015). Diverse uncultivated ultra-small bacterial cells in groundwater. *Nature Communications*, 6(1), 6372. <https://doi.org/10.1038/ncomms7372>
- Luk, A. W. S., Williams, T. J., Erdmann, S., Thane Papke, R., & Cavicchioli, R. (2014). Viruses of haloarchaea. In *Life* (Vol. 4, Issue 4, pp. 681–715). MDPI AG. <https://doi.org/10.3390/life4040681>
- Lynch, M. D. J., & Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nature Reviews Microbiology*, 13(4), 217–229. <https://doi.org/10.1038/nrmicro3400>
- Ma, L.-S., Hachani, A., Lin, J.-S., Filloux, A., & Lai, E.-M. (2014). Agrobacterium tumefaciens Deploys a Superfamily of Type VI Secretion DNase Effectors as Weapons for Interbacterial Competition In Planta. *Cell Host & Microbe*, 16(1), 94–104. <https://doi.org/10.1016/j.chom.2014.06.002>
- Magrane, M., & UniProt Consortium. (2011). UniProt Knowledgebase: A hub of integrated protein data. *Database*, 2011. <https://doi.org/10.1093/database/bar009>

- Martins, S. J., Taerum, S. J., Triplett, L., Emerson, J. B., Zasada, I., de Toledo, B. F., Kovac, J., Martin, K., & Bull, C. T. (2022). Predators of Soil Bacteria in Plant and Human Health. *Phytobiomes Journal*, 6(3), 184–200. <https://doi.org/10.1094/PBIOMES-11-21-0073-RVW>
- Matheus Carnevali, P. B., Lavy, A., Thomas, A. D., Crits-Christoph, A., Diamond, S., Méheust, R., Olm, M. R., Sharrar, A., Lei, S., Dong, W., Falco, N., Bouskill, N., Newcomer, M. E., Nico, P., Wainwright, H., Dwivedi, D., Williams, K. H., Hubbard, S., & Banfield, J. F. (2021). Meanders as a scaling motif for understanding of floodplain soil microbiome and biogeochemical potential at the watershed scale. *Microbiome*, 9(1), 121. <https://doi.org/10.1186/s40168-020-00957-z>
- Matthews, R. G. (2009). Cobalamin- and corrinoid-dependent enzymes. *Metal Ions in Life Sciences*, 6(Ii), 53–114. <https://doi.org/10.1039/BK9781847559159-00053>
- McDaniel, L., & Paul, J. H. (2005). Effect of nutrient addition and environmental factors on prophage induction in natural populations of marine synechococcus species. *Applied and Environmental Microbiology*, 71(2), 842–850. <https://doi.org/10.1128/AEM.71.2.842-850.2005>
- Meaden, S., Biswas, A., Arkhipova, K., Morales, S. E., Dutilh, B. E., Westra, E. R., & Fineran, P. C. (2022). High viral abundance and low diversity are associated with increased CRISPR-Cas prevalence across microbial ecosystems. *Current Biology*, 32(1), 220–227.e5. <https://doi.org/10.1016/j.cub.2021.10.038>
- Mellbye, B. L., Giguere, A. T., Bottomley, P. J., & Sayavedra-Soto, L. A. (2016). Quorum Quenching of *Nitrobacter winogradskyi* Suggests that Quorum Sensing Regulates Fluxes of Nitrogen Oxide(s) during Nitrification. *MBio*, 7(5). <https://doi.org/10.1128/mBio.01753-16>
- Men, Y., Seth, E. C., Yi, S., Allen, R. H., Taga, M. E., & Alvarez-Cohen, L. (2014). Sustainable growth of *Dehalococcoides mccartyi* 195 by corrinoid salvaging and remodeling in defined lactate-fermenting consortia. *Applied and Environmental Microbiology*, 80(7), 2133–2141. <https://doi.org/10.1128/AEM.03477-13>
- Men, Y., Seth, E. C., Yi, S., Crofts, T. S., Allen, R. H., Taga, M. E., & Alvarez-Cohen, L. (2015). Identification of specific corrinoids reveals corrinoid modification in dechlorinating microbial communities. *Environmental Microbiology*, 17(12), 4873–4884. <https://doi.org/10.1111/1462-2920.12500>
- Men, Y., Yu, K., Bælum, J., Gao, Y., Tremblay, J., Prestat, E., Stenuit, B., Tringe, S. G., Jansson, J., Zhang, T., & Alvarez-Cohen, L. (2017). Metagenomic and metatranscriptomic analyses reveal the structure and dynamics of a dechlorinating community containing *Dehalococcoides mccartyi* and corrinoidproviding microorganisms under cobalaminlimited conditions. *Applied and Environmental Microbiology*, 83(8). <https://doi.org/10.1128/AEM.03508-16>
- Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, 7. <https://doi.org/10.1038/ncomms11257>
- Miao, J., Zhang, N., Liu, H., Wang, H., Zhong, Z., & Zhu, J. (2018). Soil commensal rhizobia promote *Rhizobium etli* nodulation efficiency through CinR-mediated quorum sensing. *Archives of Microbiology*, 200(5), 685–694. <https://doi.org/10.1007/s00203-018-1478-2>
- Mihara, T., Nishimura, Y., Shimizu, Y., Nishiyama, H., Yoshikawa, G., Uehara, H., Hingamp, P., Goto, S., & Ogata, H. (2016). Linking Virus Genomes with Host Taxonomy. *Viruses*, 8(3), 66. <https://doi.org/10.3390/v8030066>
- Miller, R. V., & Day, M. J. (2009). Contribution of lysogeny, pseudolysogeny, and starvation to

- phage ecology. In S. T. Abedon (Ed.), *Bacteriophage Ecology* (pp. 114–144). Cambridge University Press. <https://doi.org/10.1017/CBO9780511541483.008>
- Mireku, S. A., Ruetz, M., Zhou, T., Korkhov, V. M., Kräutler, B., & Locher, K. P. (2017). Conformational Change of a Tryptophan Residue in BtuF Facilitates Binding and Transport of Cobinamide by the Vitamin B12 Transporter BtuCD-F. *Scientific Reports*, 7, 41575. <https://doi.org/10.1038/srep41575>
- Mok, K. C., Sokolovskaya, O. M., Nicolas, A. M., Hallberg, Z. F., Deutschbauer, A., Carlson, H. K., & Taga, M. E. (2020). Identification of a Novel Cobamide Remodeling Enzyme in the Beneficial Human Gut Bacterium *Akkermansia muciniphila*. *MBio*. <https://doi.org/10.1128/MBIO.02507-20>
- Mok, K. C., & Taga, M. E. (2013). Growth inhibition of *Sporomusa ovata* by incorporation of benzimidazole bases into cobamides. *Journal of Bacteriology*, 195(9), 1902–1911. <https://doi.org/10.1128/JB.01282-12>
- Musat, N., Musat, F., Weber, P. K., & Pett-Ridge, J. (2016). Tracking microbial interactions with NanoSIMS. *Current Opinion in Biotechnology*, 41(May 2018), 114–121. <https://doi.org/10.1016/j.copbio.2016.06.007>
- Nafissi, N., & Slavcev, R. (2014). Bacteriophage recombination systems and biotechnical applications. *Applied Microbiology and Biotechnology*, 98(7), 2841–2851. <https://doi.org/10.1007/s00253-014-5512-2>
- Naylor, D., Sadler, N., Bhattacharjee, A., Graham, E. B., Anderton, C. R., McClure, R., Lipton, M., Hofmockel, K. S., & Jansson, J. K. (2020). Soil microbiomes under climate change and implications for carbon cycling. In *Annual Review of Environment and Resources* (Vol. 45, Issue 1, pp. 29–59). <https://doi.org/10.1146/annurev-environ-012320-082720>
- Nazir, R., Mazurier, S., Yang, P., Lemanceau, P., & van Elsas, J. D. (2017). The Ecological Role of Type Three Secretion Systems in the Interaction of Bacteria with Fungi in Soil and Related Habitats Is Diverse and Context-Dependent. *Frontiers in Microbiology*, 8(JAN), 38. <https://doi.org/10.3389/fmicb.2017.00038>
- 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>
- Neufeld, J. D., Vohra, J., Dumont, M. G., Lueders, T., Manefield, M., Friedrich, M. W., & Murrell, C. J. (2007). DNA stable-isotope probing. *Nature Protocols*, 2(4), 860–866. <https://doi.org/10.1038/nprot.2007.109>
- Ng, W. L., & Bassler, B. L. (2009). Bacterial quorum-sensing network architectures. In *Annual Review of Genetics* (Vol. 43, Issue 1, pp. 197–222). <https://doi.org/10.1146/annurev-genet-102108-134304>
- Nicolas, A. M., Jaffe, A. L., Nuccio, E. E., Taga, M. E., Firestone, M. K., & Banfield, J. F. (2021). Soil Candidate Phyla Radiation Bacteria Encode Components of Aerobic Metabolism and Co-occur with Nanoarchaea in the Rare Biosphere of Rhizosphere Grassland Communities. *MSystems*, 6(4). <https://doi.org/10.1128/mSystems.01205-20>
- Nicolas, A. M., Sieradzki, E. T., Pett-Ridge, J., Banfield, J. F., Taga, M. E., Firestone, M. K., & Blazewicz, S. J. (2022). Isotope-enrichment reveals active viruses follow microbial host dynamics during rewetting of a California grassland soil. *BioRxiv*, 2022.09.30.510406. <https://doi.org/10.1101/2022.09.30.510406>
- Nielsen, D. A., Fierer, N., Geoghegan, J. L., Gillings, M. R., Gumerov, V., Madin, J. S., Moore,

- L., Paulsen, I. T., Reddy, T. B. K., Tetu, S. G., & Westoby, M. (2021). Aerobic bacteria and archaea tend to have larger and more versatile genomes. *Oikos*, 130(4), 501–511. <https://doi.org/10.1111/oik.07912>
- Nuccio, E. E., Blazewicz, S. J., Lafler, M., Campbell, A. N., Kakouridis, A., Kimbrel, J. A., Wollard, J., Vyshenska, D., Riley, R., Tomatsu, A., Hestrin, R., Malmstrom, R., Mary K Firestone, J., & Pett-Ridge, E. (2022). HT-SIP: A semi-automated Stable Isotope Probing pipeline identifies interactions in the hyphosphere of arbuscular mycorrhizal fungi. *BioRxiv*. <https://doi.org/10.1101/2022.07.01.498377>
- Nuccio, E. E., Hodge, A., Pett-Ridge, J., Herman, D. J., Weber, P. K., & Firestone, M. K. (2013). An arbuscular mycorrhizal fungus significantly modifies the soil bacterial community and nitrogen cycling during litter decomposition. *Environmental Microbiology*, 15(6), 1870–1881. <https://doi.org/10.1111/1462-2920.12081>
- Ofir, G., & Sorek, R. (2018). Contemporary Phage Biology: From Classic Models to New Insights. *Cell*, 172(6), 1260–1270. <https://doi.org/10.1016/J.CELL.2017.10.045>
- Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., & Kanehisa, M. (1999). KEGG: Kyoto encyclopedia of genes and genomes. In *Nucleic Acids Research* (Vol. 27, Issue 1, pp. 29–34). Oxford University Press. <https://doi.org/10.1093/nar/27.1.29>
- Olm, M. R., Brown, C. T., Brooks, B., & Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME Journal*, 11(12), 2864–2868. <https://doi.org/10.1038/ismej.2017.126>
- Osler, G. H. R., & Sommerkorn, M. (2007). Toward a complete soil C and N cycle: incorporating the soil fauna. *Ecology*, 88(7), 1611–1621. <https://doi.org/10.1890/06-1357.1>
- Ossowicki, A., Raaijmakers, J. M., & Garbeva, P. (2021). Disentangling soil microbiome functions by perturbation. *Environmental Microbiology Reports*, 13(5), 582–590. <https://doi.org/10.1111/1758-2229.12989>
- Paez-Espino, D., Elie-Fadrosh, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N. N., & Kyrpides, N. C. (2016). Uncovering Earth's virome. *Nature*. <https://doi.org/10.1038/nature19094>
- Pande, S., Kaftan, F., Lang, S., Svato, A., Germerodt, S., & Kost, C. (2016). Privatization of cooperative benefits stabilizes mutualistic cross-feeding interactions in spatially structured environments. *ISME Journal*, 10(6), 1413–1423. <https://doi.org/10.1038/ismej.2015.212>
- Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). 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. <https://doi.org/10.1111/1462-2920.13023>
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, 25(7), 1043–1055. <https://doi.org/10.1101/gr.186072.114>
- Pett-Ridge, J., & Firestone, M. K. (2017). Using stable isotopes to explore root-microbe-mineral interactions in soil. *Rhizosphere*. <https://doi.org/10.1016/j.rhisph.2017.04.016>
- Pett-Ridge, J., Petersen, D. G., Nuccio, E., & Firestone, M. K. (2013). Influence of oxic/anoxic fluctuations on ammonia oxidizers and nitrification potential in a wet tropical soil. *FEMS Microbiology Ecology*, 85(1), 179–194. <https://doi.org/10.1111/1574-6941.12111>
- Pina, M., Bize, A., Forterre, P., & Prangishvili, D. (2011). The archeoviruses. In *FEMS Microbiology Reviews* (Vol. 35, Issue 6, pp. 1035–1054). <https://doi.org/10.1111/j.1574->

- Pinilla-Redondo, R., Russel, J., Mayo-Muñoz, D., Shah, S. A., Garrett, R. A., Nesme, J., Madsen, J. S., Fineran, P. C., & Sørensen, S. J. (2022). CRISPR-Cas systems are widespread accessory elements across bacterial and archaeal plasmids. *Nucleic Acids Research*, 50(8), 4315–4328. <https://doi.org/10.1093/nar/gkab859>
- Placella, S. A., Brodie, E. L., & Firestone, M. K. (2012). Rainfall-induced carbon dioxide pulses result from sequential resuscitation of phylogenetically clustered microbial groups. *Proceedings of the National Academy of Sciences*, 109(27), 10931–10936. <https://doi.org/10.1073/pnas.1204306109>
- Placella, S. A., & Firestone, M. K. (2013). Transcriptional response of nitrifying communities to wetting of dry soil. *Applied and Environmental Microbiology*, 79(10), 3294–3302. <https://doi.org/10.1128/AEM.00404-13>
- Pomeroy, L. R., Williams, P. J. leB., Azam, F., & Hobbie, J. E. (2007). Microbial Loop. In *Oceanography* (Vol. 20, Issue 2). <https://doi.org/10.5670/oceanog.2007.45>
- Ponomarova, O., & Patil, K. R. (2015). Metabolic interactions in microbial communities: Untangling the Gordian knot. In *Current Opinion in Microbiology* (Vol. 27, pp. 37–44). <https://doi.org/10.1016/j.mib.2015.06.014>
- Portillo, M. C., Leff, J. W., Lauber, C. L., & Fierer, N. (2013). Cell Size Distributions of Soil Bacterial and Archaeal Taxa. *Applied and Environmental Microbiology*, 79(24), 7610–7617. <https://doi.org/10.1128/AEM.02710-13>
- Pratama, A. A., & van Elsas, J. D. (2018). The ‘Neglected’ Soil Virome – Potential Role and Impact. *Trends in Microbiology*, 26(8), 649–662. <https://doi.org/10.1016/j.tim.2017.12.004>
- Price, M. N., Wetmore, K. M., Waters, R. J., Callaghan, M., Ray, J., Liu, H., Kuehl, J. V., Melnyk, R. A., Lamson, J. S., Suh, Y., Carlson, H. K., Esquivel, Z., Sadeeshkumar, H., Chakraborty, R., Zane, G. M., Rubin, B. E., Wall, J. D., Visel, A., Bristow, J., ... Deutschbauer, A. M. (2018). Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature*, 557(7706), 503–509. <https://doi.org/10.1038/s41586-018-0124-0>
- Probst, A. J., Ladd, B., Jarett, J. K., Geller-Mcgrath, D. E., Sieber, C. M. K., Emerson, J. B., Anantharaman, K., Thomas, B. C., Malmstrom, R. R., Stieglmeier, M., Klingl, A., Woyke, T., Ryan, M. C., & Banfield, J. F. (2018). Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep terrestrial subsurface. *Nature Microbiology*, 3(3), 328–336. <https://doi.org/10.1038/s41564-017-0098-y>
- Quince, C., Lanzen, A., Davenport, R. J., & Turnbaugh, P. J. (2011). Removing Noise From Pyrosequenced Amplicons. *BMC Bioinformatics*, 12(1), 38. <https://doi.org/10.1186/1471-2105-12-38>
- Radajewski, S., Ineson, P., Parekh, N. R., & Murrell, J. C. (2000). Stable-isotope probing as a tool in microbial ecology. *Nature*, 403(6770), 646–649. <https://doi.org/10.1038/35001054>
- Rakoff-Nahoum, S., Foster, K. R., & Comstock, L. E. (2016). The evolution of cooperation within the gut microbiota. *Nature*, 533(7602), 255–259. <https://doi.org/10.1038/nature17626>
- Raynaud, X., & Nunan, N. (2014). Spatial Ecology of Bacteria at the Microscale in Soil. *PLoS ONE*, 9(1), e87217. <https://doi.org/10.1371/journal.pone.0087217>
- Reavy, B., Swanson, M. M., & Tiliansky, M. (2014). *Viruses in Soil* (pp. 163–180). https://doi.org/10.1007/978-94-017-8890-8_8

- Ren, J., Song, K., Deng, C., Ahlgren, N. A., Fuhrman, J. A., Li, Y., Xie, X., Poplin, R., & Sun, F. (2020). Identifying viruses from metagenomic data using deep learning. *Quantitative Biology*, 8(1), 64–77. <https://doi.org/10.1007/s40484-019-0187-4>
- Romdhane, S., Spor, A., Aubert, J., Bru, D., Breuil, M. C., Hallin, S., Mounier, A., Ouadah, S., Tsiknia, M., & Philippot, L. (2022). Unraveling negative biotic interactions determining soil microbial community assembly and functioning. *ISME Journal*, 16(1), 296–306. <https://doi.org/10.1038/s41396-021-01076-9>
- Romine, M. F., Rodionov, D. A., Maezato, Y., Anderson, L. N., Nandhikonda, P., Rodionova, I. A., Carre, A., Li, X., Xu, C., Clauss, T. R. W., Kim, Y.-M., Metz, T. O., & Wright, A. T. (2017). Elucidation of roles for vitamin B 12 in regulation of folate, ubiquinone, and methionine metabolism. *Proceedings of the National Academy of Sciences*, 114(7), E1205–E1214. <https://doi.org/10.1073/pnas.1612360114>
- Romine, M. F., Rodionov, D. A., Maezato, Y., Osterman, A. L., & Nelson, W. C. (2017). Underlying mechanisms for syntrophic metabolism of essential enzyme cofactors in microbial communities. *The ISME Journal*, 11(6), 1434–1446. <https://doi.org/10.1038/ismej.2017.2>
- Rønn, R., Vestergård, M., & Ekelund, F. (2012). Interactions between bacteria, protozoa and nematodes in soil. *Acta Protozoologica*, 51(3), 223–235. <https://doi.org/10.4467/16890027AP.12.018.0764>
- Rosnow, J. J., Hwang, S., Killinger, B. J., Kim, Y. M., Moore, R. J., Lindemann, S. R., Maupin-Furlow, J. A., & Wright, A. T. (2018). A cobalamin activity-based probe enables microbial cell growth and finds new cobalamin-protein interactions across domains. *Applied and Environmental Microbiology*, 84(18). <https://doi.org/10.1128/AEM.00955-18>
- Ross, B. D., Verster, A. J., Radey, M. C., Schmidtke, D. T., Pope, C. E., Hoffman, L. R., Hajjar, A. M., Peterson, S. B., Borenstein, E., & Mougous, J. D. (2019). Human gut bacteria contain acquired interbacterial defence systems. *Nature*, 575(7781), 224–228. <https://doi.org/10.1038/s41586-019-1708-z>
- Roth, J. R., Lawrence, J. G., & Bobik, T. A. (1996). Cobalamin (coenzyme B12): Synthesis and biological significance. *Annual Review of Microbiology*, 50(1), 137–181. <https://doi.org/10.1146/annurev.micro.50.1.137>
- Rousk, J., & Bååth, E. (2011). Growth of saprotrophic fungi and bacteria in soil. *FEMS Microbiology Ecology*, 78(1), 17–30. <https://doi.org/10.1111/j.1574-6941.2011.01106.x>
- Roux, S., Brum, J. R., Dutilh, B. E., Sunagawa, S., Duhaime, M. B., Loy, A., Poulos, B. T., Solonenko, N., Lara, E., Poulain, J., Pesant, S., Kandels-Lewis, S., Dimier, C., Picheral, M., Searson, S., Cruaud, C., Alberti, A., Duarte, C. M., Gasol, J. M., ... Sullivan, M. B. (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature*, 537(7622), 689–693. <https://doi.org/10.1038/nature19366>
- Roux, S., Enault, F., Hurwitz, B. L., & Sullivan, M. B. (2015). VirSorter: mining viral signal from microbial genomic data. *PeerJ*, 3, e985. <https://doi.org/10.7717/peerj.985>
- Roux, S., Krupovic, M., Daly, R. A., Borges, A. L., Nayfach, S., Schulz, F., Sharrar, A., Matheus Carnevali, P. B., Cheng, J.-F., Ivanova, N. N., Bondy-Denomy, J., Wrigton, K. C., Woyke, T., Visel, A., Kyrpides, N. C., & Elie-Fadrosh, E. A. (2019). Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth's biomes. *Nature Microbiology*, 4(11), 1895–1906. <https://doi.org/10.1038/s41564-019-0510-x>
- Rubin, B. E., Diamond, S., Cress, B. F., Crits-Christoph, A., Lou, Y. C., Borges, A. L., Shivram, H., He, C., Xu, M., Zhou, Z., Smith, S. J., Rovinsky, R., Smock, D. C. J., Tang, K., Owens,

- T. K., Krishnappa, N., Sachdeva, R., Barrangou, R., Deutschbauer, A. M., ... Doudna, J. A. (2021). Species- and site-specific genome editing in complex bacterial communities. *Nature Microbiology*, 7(1), 34–47. <https://doi.org/10.1038/s41564-021-01014-7>
- Santos-Medellín, C., Estera-Molina, K., Yuan, M., Pett-Ridge, J., Firestone, M. K., & Emerson, J. B. (2022). Spatial turnover of soil viral populations and genotypes overlain by cohesive responses to moisture in grasslands. *BioRxiv*. <https://doi.org/2022.03.24.485562>
- Sañudo-Wilhelmy, S. A., Gómez-Consarnau, L., Suffridge, C., & Webb, E. A. (2014). The Role of B Vitamins in Marine Biogeochemistry. *Annual Review of Marine Science*, 6(1), 339–367. <https://doi.org/10.1146/annurev-marine-120710-100912>
- Sasse, J., Kant, J., Cole, B. J., Klein, A. P., Arsova, B., Schlaepfer, P., Gao, J., Lewald, K., Zhakhina, K., Kosina, S., Bowen, B. P., Treen, D., Vogel, J., Visel, A., Watt, M., Dangel, J. L., & Northen, T. R. (2019). Multilab EcoFAB study shows highly reproducible physiology and depletion of soil metabolites by a model grass. *New Phytologist*, 222(2), 1149–1160. <https://doi.org/10.1111/nph.15662>
- Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R., & Kapos, V. (2014). Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, 5(1), 81–91. <https://doi.org/10.4155/cmt.13.77>
- Schimel, J. P. (2018). Life in dry soils: Effects of drought on soil microbial communities and processes. In *Annual Review of Ecology, Evolution, and Systematics* (Vol. 49, pp. 409–432). Annual Reviews Inc. <https://doi.org/10.1146/annurev-ecolsys-110617-062614>
- Schwartz, E. (2007). Characterization of growing microorganisms in soil by stable isotope probing with H₂¹⁸O. *Applied and Environmental Microbiology*, 73(8), 2541–2546. <https://doi.org/10.1128/AEM.02021-06>
- Scott, S. R., & Hasty, J. (2016). Quorum Sensing Communication Modules for Microbial Consortia. *ACS Synthetic Biology*, 5(9), 969–977. <https://doi.org/10.1021/acssynbio.5b00286>
- Seed, K. D., Lazinski, D. W., Calderwood, S. B., & Camilli, A. (2013). A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature*, 494(7438), 489–491. <https://doi.org/10.1038/nature11927>
- Segobola, J., Adriaenssens, E., Tsekoa, T., Rashamuse, K., & Cowan, D. (2018). Exploring Viral Diversity in a Unique South African Soil Habitat. *Scientific Reports*, 8(1), 111. <https://doi.org/10.1038/s41598-017-18461-0>
- Seth, E. C., & Taga, M. E. (2014). Nutrient cross-feeding in the microbial world. *Frontiers in Microbiology*, 5(JULY), 350. <https://doi.org/10.3389/fmicb.2014.00350>
- Shaffer, M., Borton, M. A., McGivern, B. B., Zayed, A. A., La Rosa, S. L. 0003 3527 8101, Solden, L. M., Liu, P., Narrowe, A. B., Rodríguez-Ramos, J., Bolduc, B., Gazitúa, M. C., Daly, R. A., Smith, G. J., Vik, D. R., Pope, P. B., Sullivan, M. B., Roux, S., & Wrighton, K. C. (2020). DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Research*, 48(16), 8883–8900. <https://doi.org/10.1093/nar/gkaa621>
- Shang, Y., Sikorski, J., Bonkowski, M., Fiore-Donno, A.-M., Kandeler, E., Marhan, S., Boeddinghaus, R. S., Solly, E. F., Schruppf, M., Schöning, I., Wubet, T., Buscot, F., & Overmann, J. (2017). Inferring interactions in complex microbial communities from nucleotide sequence data and environmental parameters. *PLOS ONE*, 12(3), e0173765. <https://doi.org/10.1371/journal.pone.0173765>
- Sharma, K., Palatinszky, M., Nikolov, G., Berry, D., & Shank, E. A. (2020). Transparent soil

- microcosms for live-cell imaging and non-destructive stable isotope probing of soil microorganisms. *ELife*, 9, 1–28. <https://doi.org/10.7554/eLife.56275>
- Sharma, V., Rodionov, D. A., Leyn, S. A., Tran, D., Iablokov, S. N., Ding, H., Peterson, D. A., Osterman, A. L., & Peterson, S. N. (2019). B-Vitamin sharing promotes stability of gut microbial communities. *Frontiers in Microbiology*, 10(JUL), 1–15. <https://doi.org/10.3389/fmicb.2019.01485>
- Sharpless, W. A., Sander, K. B., Song, F., Kuehl, J. V., & Arkin, A. P. (2022). Towards Environmental Control of Microbiomes. *BioRxiv*, 2022.11.04.515211. <https://doi.org/10.1101/2022.11.04.515211>
- Sharrar, A. M., Crits-Christoph, A., Méheust, R., Diamond, S., Starr, E. P., & Banfield, J. F. (2020). Bacterial Secondary Metabolite Biosynthetic Potential in Soil Varies with Phylum, Depth, and Vegetation Type. *MBio*, 11(3). <https://doi.org/10.1128/mBio.00416-20>
- Shelton, A. N., Seth, E. C., Mok, K. C., Han, A. W., Jackson, S. N., Haft, D. R., & Taga, M. E. (2019). Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *The ISME Journal*, 13(3), 789–804. <https://doi.org/10.1038/s41396-018-0304-9>
- Sherlock, M. E., & Breaker, R. R. (2020). Former orphan riboswitches reveal unexplored areas of bacterial metabolism, signaling, and gene control processes. *RNA*, 26(6), 675–693. <https://doi.org/10.1261/rna.074997.120>
- Shmakov, S. A., Sitnik, V., Makarova, K. S., Wolf, Y. I., Severinov, K. V., & Koonin, E. V. (2017). The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. *MBio*, 8(5), e01397-17. <https://doi.org/10.1128/mBio.01397-17>
- Shu, L., He, Z., Guan, X., Yang, X., Tian, Y., Zhang, S., Wu, C., He, Z., Yan, Q., Wang, C., & Shi, Y. (2021). A dormant amoeba species can selectively sense and predate on different soil bacteria. *Functional Ecology*, 35(8), 1708–1721. <https://doi.org/10.1111/1365-2435.13824>
- Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., & Banfield, J. F. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nature Microbiology*, 3(7), 836–843. <https://doi.org/10.1038/s41564-018-0171-1>
- Sieradzki, E. T., Greenlon, A., Nicolas, A. M., Firestone, M. K., Pett-Ridge, J., Blazewicz, S. J., & Banfield, J. F. (2022). Functional succession of actively growing soil microorganisms during rewetting is shaped by precipitation history. *BioRxiv*, 2022.06.28.498032. <https://doi.org/10.1101/2022.06.28.498032>
- Silpe, J. E., & Bassler, B. L. (2019). A Host-Produced Quorum-Sensing Autoinducer Controls a Phage Lysis-Lysogeny Decision. *Cell*, 176(1–2), 268–280.e13. <https://doi.org/10.1016/j.cell.2018.10.059>
- Sistla, S. A., Moore, J. C., Simpson, R. T., Gough, L., Shaver, G. R., & Schimel, J. P. (2013). Long-term warming restructures Arctic tundra without changing net soil carbon storage. *Nature*, 497(7451), 615–618. <https://doi.org/10.1038/nature12129>
- Sokol, N. W., Slessarev, E., Marschmann, G. L., Nicolas, A., Blazewicz, S. J., Brodie, E. L., Firestone, M. K., Foley, M. M., Hestrin, R., Hungate, B. A., Koch, B. J., Stone, B. W., Sullivan, M. B., Zablocki, O., LLNL Soil Microbiome Consortium, & Pett-Ridge, J. (2022). Life and death in the soil microbiome: how ecological processes influence biogeochemistry. *Nature Reviews Microbiology*, 0123456789. <https://doi.org/10.1038/s41579-022-00695-z>
- Sokolovskaya, O. M., Mok, K. C., Park, J. D., Tran, J. L. A., Quanstrom, K. A., & Taga, M. E.

- (2019). Cofactor selectivity in methylmalonyl coenzyme a mutase, a model cobamide-dependent enzyme. *MBio*, 10(5). <https://doi.org/10.1128/mBio.01303-19>
- Sokolovskaya, O. M., Shelton, A. N., & Taga, M. E. (2020). Sharing vitamins: Cobamides unveil microbial interactions. *Science*, 369(6499). <https://doi.org/10.1126/science.aba0165>
- Soong, J. L., & Nielsen, U. N. (2016). The role of microarthropods in emerging models of soil organic matter. *Soil Biology and Biochemistry*, 102, 37–39. <https://doi.org/10.1016/j.soilbio.2016.06.020>
- Stark, J. M., & Firestone, M. K. (1995). Mechanisms for soil moisture effects on activity of nitrifying bacteria. *Applied and Environmental Microbiology*, 61(1), 218–221. <https://doi.org/10.1128/aem.61.1.218-221.1995>
- Starr, E. P., Nuccio, E. E., Pett-Ridge, J., Banfield, J. F., & Firestone, M. K. (2019). Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proceedings of the National Academy of Sciences of the United States of America*, 116(51), 25900–25908. <https://doi.org/10.1073/pnas.1908291116>
- Starr, E. P., Shi, S., Blazewicz, S. J., Koch, B. J., Probst, A. J., Hungate, B. A., Pett-Ridge, J., Firestone, M. K., & Banfield, J. F. (2021). Stable-Isotope-Informed, Genome-Resolved Metagenomics Uncovers Potential Cross-Kingdom Interactions in Rhizosphere Soil. *MSphere*, 6(5). <https://doi.org/10.1128/msphere.00085-21>
- Starr, E. P., Shi, S., Blazewicz, S. J., Probst, A. J., Herman, D. J., Firestone, M. K., & Banfield, J. F. (2018). Stable isotope informed genome-resolved metagenomics reveals that Saccharibacteria utilize microbially-processed plant-derived carbon. *Microbiome*, 6(1), 122. <https://doi.org/10.1186/s40168-018-0499-z>
- Steen, A. D., Crits-Christoph, A., Carini, P., DeAngelis, K. M., Fierer, N., Lloyd, K. G., & Cameron Thrash, J. (2019). High proportions of bacteria and archaea across most biomes remain uncultured. *ISME Journal*, 13(12), 3126–3130. <https://doi.org/10.1038/s41396-019-0484-y>
- Stopnisek, N., & Shade, A. (2021). Persistent microbiome members in the common bean rhizosphere: an integrated analysis of space, time, and plant genotype. *The ISME Journal*, 15(9), 2708–2722. <https://doi.org/10.1038/s41396-021-00955-5>
- Sui, B., Qi, X., Wang, X., Ren, H., Liu, W., & Zhang, C. (2021). Characterization of a Novel Bacteriophage swi2 Harboring Two Lysins Can Naturally Lyse Escherichia coli. *Frontiers in Microbiology*, 12, 1201. <https://doi.org/10.3389/fmicb.2021.670799>
- Suttle, C. A. (2005). Viruses in the sea. In *Nature* (Vol. 437, Issue 7057, pp. 356–361). Nature Publishing Group. <https://doi.org/10.1038/nature04160>
- Suttle, C. A. (2007). Marine viruses - Major players in the global ecosystem. *Nature Reviews Microbiology*, 5(10), 801–812. <https://doi.org/10.1038/nrmicro1750>
- Suttle, C. A., & Culley, A. I. (2007). Viral Community Structure. In *Manual of Environmental Microbiology, Third Edition* (pp. 445–453). American Society of Microbiology. <https://doi.org/10.1128/9781555815882.ch36>
- Suzek, B. E., Huang, H., McGarvey, P., Mazumder, R., & Wu, C. H. (2007). UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, 23(10), 1282–1288. <https://doi.org/10.1093/bioinformatics/btm098>
- Tecon, R., Ebrahimi, A., Kleyer, H., Erev Levi, S., & Or, D. (2018). Cell-to-cell bacterial interactions promoted by drier conditions on soil surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, 115(39), 9791–9796. <https://doi.org/10.1073/pnas.1808274115>

- Thakur, M. P., & Geisen, S. (2019). Trophic Regulations of the Soil Microbiome. *Trends in Microbiology*, 27(9), 771–780. <https://doi.org/10.1016/j.tim.2019.04.008>
- The Pandas Development Team. (2022). pandas-dev/pandas: Pandas (v1.5.1). *Zenodo*. <https://doi.org/10.5281/zenodo.7223478>
- Touchon, M., Bernheim, A., & Rocha, E. P. C. (2016). Genetic and life-history traits associated with the distribution of prophages in bacteria. *ISME Journal*, 10(11), 2744–2754. <https://doi.org/10.1038/ismej.2016.47>
- Traxler, M. F., Seyedsayamdost, M. R., Clardy, J., & Kolter, R. (2012). Interspecies modulation of bacterial development through iron competition and siderophore piracy. *Molecular Microbiology*, 86(3), 628–644. <https://doi.org/10.1111/mmi.12008>
- Trubl, G., Kimbrel, J. A., Lique-Gonzalez, J., Nuccio, E. E., Weber, P. K., Pett-Ridge, J., Jansson, J. K., Waldrop, M. P., & Blazewicz, S. J. (2021). Active virus-host interactions at sub-freezing temperatures in Arctic peat soil. *Microbiome*, 9(1), 208. <https://doi.org/10.1186/s40168-021-01154-2>
- Trubl, G., Solonenko, N., Chittick, L., Solonenko, S. A., Rich, V. I., & Sullivan, M. B. (2016). Optimization of viral resuspension methods for carbon-rich soils along a permafrost thaw gradient. *PeerJ*, 4(C), e1999. <https://doi.org/10.7717/peerj.1999>
- Turnbull, L., Toyofuku, M., Hynen, A. L., Kurosawa, M., Pessi, G., Petty, N. K., Osvath, S. R., Cárcamo-Oyarce, G., Gloag, E. S., Shimoni, R., Omasits, U., Ito, S., Yap, X., Monahan, L. G., Cavaliere, R., Ahrens, C. H., Charles, I. G., Nomura, N., Eberl, L., & Whitchurch, C. B. (2016). Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nature Communications*, 7(1), 11220. <https://doi.org/10.1038/ncomms11220>
- Uritskiy, G. V., Diruggiero, J., & Taylor, J. (2018). MetaWRAP - A flexible pipeline for genome-resolved metagenomic data analysis 08 Information and Computing Sciences 0803 Computer Software 08 Information and Computing Sciences 0806 Information Systems. *Microbiome*, 6(1). <https://doi.org/10.1186/s40168-018-0541-1>
- Utter, D. R., He, X., Cavanaugh, C. M., McLean, J. S., & Bor, B. (2020). The saccharibacterium TM7x elicits differential responses across its host range. *ISME Journal*, 14(12), 3054–3067. <https://doi.org/10.1038/s41396-020-00736-6>
- Van Goethem, M. W., Swenson, T. L., Trubl, G., Roux, S., & Northen, T. R. (2019). Characteristics of Wetting-Induced Bacteriophage Blooms in Biological Soil Crust. *MBio*, 10(6). <https://doi.org/10.1128/mBio.02287-19>
- Verbruggen, E., El Mouden, C., Jansa, J., Akkermans, G., Bücking, H., West, S. A., & Kiers, E. T. (2012). Spatial structure and interspecific cooperation: theory and an empirical test using the mycorrhizal mutualism. *The American Naturalist*, 179(5), E133–46. <https://doi.org/10.1086/665032>
- Větrovský, T., & Baldrian, P. (2013). The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLoS ONE*, 8(2), 1–10. <https://doi.org/10.1371/journal.pone.0057923>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wang, C., Michalet, R., Liu, Z., Jiang, X., Wang, X., Zhang, G., An, L., Chen, S., & Xiao, S.

- (2020). Disentangling Large- and Small-Scale Abiotic and Biotic Factors Shaping Soil Microbial Communities in an Alpine Cushion Plant System. *Frontiers in Microbiology*, 11, 925. <https://doi.org/10.3389/fmicb.2020.00925>
- Warren, M. J., Raux, E., Schubert, H. L., & Escalante-Semerena, J. C. (2002). The biosynthesis of adenosylcobalamin (vitamin B12). *Natural Product Reports*, 19(4), 390–412. <https://doi.org/10.1039/b108967f>
- Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>
- Weitz, J. S., Beckett, S. J., Brum, J. R., Cael, B. B., & Dushoff, J. (2017). Lysis, lysogeny and virus–microbe ratios. *Nature*, 549(7672), E1–E3. <https://doi.org/10.1038/nature23295>
- Weitz, J. S., Stock, C. A., Wilhelm, S. W., Bourouiba, L., Coleman, M. L., Buchan, A., Follows, M. J., Fuhrman, J. A., Jover, L. F., Lennon, J. T., Middelboe, M., Sonderegger, D. L., Suttle, C. A., Taylor, B. P., Frede Thingstad, T., Wilson, W. H., & Eric Wommack, K. (2015). A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *The ISME Journal*, 9(6), 1352–1364. <https://doi.org/10.1038/ismej.2014.220>
- Wellington, S., & Greenberg, E. P. (2019). Quorum Sensing Signal Selectivity and the Potential for Interspecies Cross Talk. *MBio*, 10(2), e00146-19. <https://doi.org/10.1128/mBio.00146-19>
- Wilhelm, S. W., & Suttle, C. A. (1999). Viruses and Nutrient Cycles in the Sea. *BioScience*, 49(10), 781–788. <https://doi.org/10.2307/1313569>
- Williamson, K. E., Corzo, K. A., Drissi, C. L., Buckingham, J. M., Thompson, C. P., & Helton, R. R. (2013). Estimates of viral abundance in soils are strongly influenced by extraction and enumeration methods. *Biology and Fertility of Soils*, 49(7), 857–869. <https://doi.org/10.1007/s00374-013-0780-z>
- Williamson, K. E., Fuhrmann, J. J., Wommack, K. E., & Radosevich, M. (2017). Viruses in Soil Ecosystems: An Unknown Quantity Within an Unexplored Territory. *Annual Review of Virology*, 4(1), 201–219. <https://doi.org/10.1146/annurev-virology-101416-041639>
- Williamson, K. E., Radosevich, M., Smith, D. W., & Wommack, K. E. (2007). Incidence of lysogeny within temperate and extreme soil environments. *Environmental Microbiology*, 9(10), 2563–2574. <https://doi.org/10.1111/j.1462-2920.2007.01374.x>
- Williamson, K. E., Radosevich, M., & Wommack, K. E. (2005). Abundance and diversity of viruses in six Delaware soils. *Applied and Environmental Microbiology*, 71(6), 3119–3125. <https://doi.org/10.1128/AEM.71.6.3119-3125.2005>
- Williamson, K. E., Schnitker, J. B., Radosevich, M., Smith, D. W., & Wommack, K. E. (2008). Cultivation-based assessment of lysogeny among soil bacteria. *Microbial Ecology*, 56(3), 437–447. <https://doi.org/10.1007/s00248-008-9362-2>
- Williamson, S. J., Allen, L. Z., Lorenzi, H. A., Fadrosch, D. W., Bami, D., Thiagarajan, M., McCrow, J. P., Tovchigrechko, A., Yooseph, S., & Venter, J. C. (2012). Metagenomic Exploration of Viruses throughout the Indian Ocean. *PLoS ONE*, 7(10), e42047. <https://doi.org/10.1371/journal.pone.0042047>
- Williamson, S. J., Houchin, L. A., McDaniel, L., & Paul, J. H. (2002). Seasonal Variation in Lysogeny as Depicted by Prophage Induction in Tampa Bay, Florida. *Applied and Environmental Microbiology*, 68(9), 4307–4314. <https://doi.org/10.1128/AEM.68.9.4307-4314.2002>
- Wilpieszski, R. L., Aufrecht, J. A., Retterer, S. T., Sullivan, M. B., Graham, D. E., Pierce, E. M.,

- Zablocki, O. D., Palumbo, A. V., & Elias, D. A. (2019). Soil Aggregate Microbial Communities: Towards Understanding Microbiome Interactions at Biologically Relevant Scales. In V. Müller (Ed.), *Applied and environmental microbiology* (Vol. 85, Issue 14). NLM (Medline). <https://doi.org/10.1128/AEM.00324-19>
- Woodcroft, B. J., Singleton, C. M., Boyd, J. A., Evans, P. N., Emerson, J. B., Zayed, A. A. F., Hoelzle, R. D., Lamberton, T. O., McCalley, C. K., Hodgkins, S. B., Wilson, R. M., Purvine, S. O., Nicora, C. D., Li, C., Frolking, S., Chanton, J. P., Crill, P. M., Saleska, S. R., Rich, V. I., & Tyson, G. W. (2018). Genome-centric view of carbon processing in thawing permafrost. *Nature*, 560(7716), 49–54. <https://doi.org/10.1038/s41586-018-0338-1>
- Woodson, J. D., & Escalante-Semerena, J. C. (2004). CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B12 precursor cobinamide in archaea. *Proceedings of the National Academy of Sciences of the United States of America*, 101(10), 3591–3596. <https://doi.org/10.1073/pnas.0305939101>
- Wu, R., Davison, M. R., Gao, Y., Nicora, C. D., Mcdermott, J. E., Burnum-Johnson, K. E., Hofmockel, K. S., & Jansson, J. K. (2021). Moisture modulates soil reservoirs of active DNA and RNA viruses. *Communications Biology*, 4(1), 1–11. <https://doi.org/10.1038/s42003-021-02514-2>
- Wu, Y.-W., Tang, Y.-H., Tringe, S. G., Simmons, B. A., & Singer, S. W. (2014). MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome*, 2(1), 26. <https://doi.org/10.1186/2049-2618-2-26>
- Xu, L., Naylor, D., Dong, Z., Simmons, T., Pierroz, G., Hixson, K. K., Kim, Y.-M., Zink, E. M., Engbrecht, K. M., Wang, Y., Gao, C., DeGraaf, S., Madera, M. A., Sievert, J. A., Hollingsworth, J., Birdseye, D., Scheller, H. V., Hutmacher, R., Dahlberg, J., ... Coleman-Derr, D. (2018). Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. *Proceedings of the National Academy of Sciences*, 115(18), 201717308. <https://doi.org/10.1073/pnas.1717308115>
- Yan, J., Şimşir, B., Farmer, A. T., Bi, M., Yang, Y., Campagna, S. R., & Löffler, F. E. (2016). The corrinoid cofactor of reductive dehalogenases affects dechlorination rates and extents in organohalide-respiring *Dehalococcoides mccartyi*. *ISME Journal*, 10(5), 1092–1101. <https://doi.org/10.1038/ismej.2015.197>
- 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, 107531. <https://doi.org/10.1016/j.soilbio.2019.107531>
- Ye, X., Li, Z., Luo, X., Wang, W., Li, Y., Li, R., Zhang, B., Qiao, Y., Zhou, J., Fan, J., Wang, H., Huang, Y., Cao, H., Cui, Z., & Zhang, R. (2020). A predatory myxobacterium controls cucumber *Fusarium* wilt by regulating the soil microbial community. *Microbiome*, 8(1), 1–17. <https://doi.org/10.1186/s40168-020-00824-x>
- Yi, S., Seth, E. C., Men, Y.-J., Stabler, S. P., Allen, R. H., Alvarez-Cohen, L., & Taga, M. E. (2012). Versatility in corrinoid salvaging and remodeling pathways supports corrinoid-dependent metabolism in *Dehalococcoides mccartyi*. *Applied and Environmental Microbiology*, 78(21), 7745–7752. <https://doi.org/10.1128/AEM.02150-12>
- Young, I. M., & Crawford, J. W. (2004). Interactions and self-organization in the soil-microbe complex. In *Science* (Vol. 304, Issue 5677, pp. 1634–1637). <https://doi.org/10.1126/science.1097394>
- Yuan, M. M., Guo, X., Wu, L., Zhang, Y., Xiao, N., Ning, D., Shi, Z., Zhou, X., Wu, L., Yang,

- Y., Tiedje, J. M., & Zhou, J. (2021). Climate warming enhances microbial network complexity and stability. *Nature Climate Change*, 11(4), 343–348. <https://doi.org/10.1038/s41558-021-00989-9>
- Zegeye, E. K., Brislawn, C. J., Farris, Y., Fansler, S. J., Hofmockel, K. S., Jansson, J. K., Wright, A. T., Graham, E. B., Naylor, D., McClure, R. S., & Bernstein, H. C. (2019). Selection, Succession, and Stabilization of Soil Microbial Consortia. *MSystems*, 4(4), 55–74. <https://doi.org/10.1128/msystems.00055-19>
- Zengler, K., & Zaramela, L. S. (2018). The social network of microorganisms — how auxotrophies shape complex communities. *Nature Reviews Microbiology*, 1. <https://doi.org/10.1038/s41579-018-0004-5>
- Zepeda-Rivera, M. A., Saak, C. C., & Gibbs, K. A. (2018). A proposed chaperone of the bacterial type VI secretion system functions to constrain a self-identity protein. *Journal of Bacteriology*, 200(14). <https://doi.org/10.1128/JB.00688-17>
- Zhalnina, K., Louie, K. B., Hao, Z., Mansoori, N., Da Rocha, U. N., Shi, S., Cho, H., Karaoz, U., Loqué, D., Bowen, B. P., Firestone, M. K., Northen, T. R., & Brodie, E. L. (2018). Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nature Microbiology*, 3(4), 470–480. <https://doi.org/10.1038/s41564-018-0129-3>
- Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 30(5), 614–620. <https://doi.org/10.1093/bioinformatics/btt593>
- Zhang, Z., Qu, Y., Li, S., Feng, K., Wang, S., Cai, W., Liang, Y., Li, H., Xu, M., Yin, H., & Deng, Y. (2017). Soil bacterial quantification approaches coupling with relative abundances reflecting the changes of taxa. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-017-05260-w>
- Zhao, Z., Gonsior, M., Schmitt-Kopplin, P., Zhan, Y., Zhang, R., Jiao, N., & Chen, F. (2019). Microbial transformation of virus-induced dissolved organic matter from picocyanobacteria: coupling of bacterial diversity and DOM chemodiversity. *ISME Journal*, 13(10), 2551–2565. <https://doi.org/10.1038/s41396-019-0449-1>
- Zimmerman, A. E., Howard-Varona, C., Needham, D. M., John, S. G., Worden, A. Z., Sullivan, M. B., Waldbauer, J. R., & Coleman, M. L. (2020). Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nature Reviews Microbiology*, 18(1), 21–34. <https://doi.org/10.1038/s41579-019-0270-x>

Appendix / Glossary of Terms

CPR, Candidate Phyla Radiation

DPANN, the group of archaea originally defined as an acronym based on the lineages the first included phyla: Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota and Nanohaloarchaeota

HMM, Hidden Markov Model

qSIP, quantitative stable isotope probing

Virome, viral metagenome

APE, atom percent excess

MAG, Metagenome-Assembled Genome

vOTU, viral Operational Taxonomic Unit

zOTU, zero-radius Operational Taxonomic Unit

B12, Vitamin B12 or cobalamin

pB12, Pseudocobalamin

Cre, Cresol-cobamide

Cbi, Cobinamide