UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Corrinoids as Model Nutrients for the Study of Soil Bacterial Interactions

Permalink

https://escholarship.org/uc/item/21j7d23j

Author

Alvarez Aponte, Zoila I

Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Corrinoids as Model Nutrients for the Study of Soil Bacterial Interactions

By

Zoila I. Alvarez-Aponte

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor in Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michiko E. Taga, Chair
Professor Britt Koskella
Professor William B. Ludington
Professor Matthew F. Traxler

Fall 2024

Abstract

Corrinoids as Model Nutrients for the Study of Soil Bacterial Interactions

by

Zoila I. Alvarez-Aponte

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Michiko E. Taga, Chair

Microbial communities inhabit nearly every ecosystem on Earth and have large effects on their environment, including impacts on crucial processes like host health and global nutrient cycling. Understanding the biotic and abiotic factors that influence the structure and function of these communities is crucial for predicting their continued impacts on a changing planet. A challenge to studying microbial communities in detail is their complexity. These communities often contain hundreds to thousands of species that are growing and metabolizing different compounds, making it challenging to deduce how individual species impact one another or how particular nutrients may affect the functioning of a community.

The model nutrient approach poses a solution to the challenges of studying microbial community interactions. Corrinoids, the vitamin B_{12} family of enzyme cofactors, are synthesized by bacteria and archaea and required by organisms in all domains of life for different metabolic processes. Genomic analysis and microbial co-culture studies have shown that these nutrients are shared among microbes. Further, bacteria have preferences for specific corrinoids that affect their growth and may lead to corrinoids impacting community structure. For these reasons, the corrinoid model is ideal for studying interactions among bacteria across multiple scales, spanning molecules to communities.

In Chapter 1, I present a review of corrinoid biology across scales of increasing complexity, summarizing knowledge recovered over the last century that informs the model nutrient approach. At the molecular scale, corrinoid biology spans enzymes, transporters, regulatory riboswitches, and an elaborate biosynthesis pathway. These components give rise to corrinoid diversity and preferences for different corrinoids observed at the organismal scale. At the organismal scale, defining the categories of “producer,” “dependent,” “independent,” and “provider” is the genomic and experimental basis for characterizing and predicting interactions. These definitions contribute to the community scale, where nutrient sharing can be studied in co-cultures of producers and dependents, and different corrinoids have been shown to impact community structure. These scales come together to enable the study of microbial interactions and continued research will enable expansion into new scales.

In Chapter 2, I applied the model nutrient approach at the organismal scale through the isolation of bacteria from a soil environment and characterization of producers, dependents, and independents. The results revealed a bias for B_{12} among isolates; tested producers all synthesized B_{12} and dependents preferred it over other corrinoids. Further, a set of producers were classified as providers after corrinoid was detected in culture supernatants. These observations, combined, led to the hypothesis that these isolates may interact via corrinoid sharing in their native environment. As part of this research, I evaluated the phylogenetic dispersion of producer, dependent, and independent traits among sequenced bacteria and found that corrinoid metabolism can be predicted from taxonomy for around half of the genera studied.

In Chapter 3, I applied the model nutrient approach at the community scale with a focus on co-cultures and three-species consortia. Comparison of the monoculture growth of corrinoid-dependent isolates to their growth in co-cultures showed that the outcome of competition is predictable based on corrinoid preferences when preferences are sufficiently distinct. As part of this work, I tested interactions between producers and dependents to evaluate whether these isolates interacted by sharing corrinoids, a hypothesis generated in Chapter 2. While producers supported dependent growth by sharing B_{12}, interactions in co-culture were not easily predictable based on monoculture growth of dependents provided with producer supernatant. Finally, tests in tri-cultures showed that dependents compete for the corrinoids made available by producers. Despite emergent properties arising, results were consistent with corrinoid-based interactions being the predominant interaction in these consortia. Together, Chapters 2 and 3 illustrate the applicability of the model nutrient approach in individual bacteria and small communities and establish a system in which interactions can be further tested.

In Chapter 4, I pivot to a diversity, equity, inclusion, and belonging project focused on the qualifying exam experience of graduate students. The Inclusive Excellence in Quals Prep (IEQP) pilot program focused on increasing support for students as they go through their qualifying exam by creating a support network and providing structure during the preparation process. The program's effectiveness was evaluated through a series of surveys presented here. IEQP was demonstrated to be effective in its first year, with peer mentorship identified as a key component.

A mi mamá,
Por su mentoría, apoyo y amor incondicional.

-

To my mom,
For her mentorship, support, and unconditional love.

Table of contents

Acknowledgements

The last 5.5 years in Berkeley have been a gift. In addition to completing a PhD in an incredibly supportive lab, I found a village, met my partner, and experienced the PMB department, which is full of people committed to DEIB and community building. As I reflect on this journey, the word *community* keeps surfacing in the scientific and personal sense.

First, I must acknowledge Prof. Michi Taga, who has been an exceptional advisor. Michi tailors her mentoring to different stages of the PhD and the needs of different mentees; she has held my hand when I needed guidance and challenged me when necessary. She has been an inspiring mentor. Michi meets weekly with every lab member despite her numerous commitments, she makes time for celebrations -big and small-, and has been committed to community nurturing, during in-person and COVID-19-pandemic times. Michi has encouraged me to develop my mentoring skills by training people at different stages, and she has allowed me to collaborate with her on grant and review writing. She has always taken my ideas seriously and allowed me to grow in many aspects. I initially chose the Taga Lab because I enjoyed the science but mostly because I enjoyed Michi's mentorship, and that was the best decision I could have made.

My thesis committee members, professors Will Ludington, Britt Koskella, Matt Traxler, and Mary Firestone, have offered invaluable advice over the years. I am thankful for the expertise they have eagerly shared. Throughout my time here, I learned so much about ecology, soil microbiology, and mathematical modeling from them. Additional mentors have included Professors Karine Gibbs, Ceci Martínez-Gomez, Rachel Brem, Diana Bautista, Denzil Streete, and Trent Northen. I am also thankful for my mentors before grad school: Mrs. Luz Burgos, and Professors Gustavo Martínez, Arturo Massol, Larry Forney, Aaron Wright, and Natalie Sadler.

I had exceptional collaborators and co-authors: Drs. Hans Carlson and Valentine Trotter made it possible for me to acquire 16S rRNA gene and genome sequences. Over the years, Hans has shared important advice, perspectives on data, and much-needed words of encouragement. Dr. Eoin Brodie shared bacteria with us that enabled initial experiments in this project. Prof. Mary Firestone allowed us to sample soil from her long-term experiment site and taught me so much about soil microbiology in her role as my quals and thesis committee member. Prof. Devin Coleman-Derr and Dr. Citlali Fonseca shared valuable knowledge about soil bacteria isolations and gave me access to precious resources in their lab. Alekhya Govindaraju initiated the computational analysis of corrinoid roles in bacteria during her rotation in the lab, and I am glad we got to collaborate on this aspect of the project.

The Taga Lab has been my home over the past 5.5 years, and I could not be more thankful for the beautiful community that Michi has brought together. I had many senior graduate students to look up to when I joined the lab. Olga Sokolovskaya and Amanda Shelton left a legacy of corrinoid knowledge and computational work, without which navigating my own project would have been extremely difficult. Kris Kennedy, Gordon Pherribo, and Alexa Nicolas contributed greatly to the lab culture; Kris taught me to approach problems very methodically, and Gordon taught me to be a better facilitator of difficult conversations. As a mentor and the elder soil microbiologist in the lab, Alexa shared heaps of knowledge, literature, and advice for my project and grad school that I continue to reference. Zachary Hallberg was a great rotation mentor, soil project colleague, and sometimes pipetting robot. Janani Hariharan was a postdoc in the lab for a

short time, during which she taught me a lot about phylogenetic trees and pedagogy. In more recent years, it has been a delight to see Dennis Suazo and Eleanor Wang grow as scientists and bring so much joy to the day-to-day. I learned so much from my mentees, Myka Green, Hazel Anagu, Dennis Suazo, and Cami Scantlin, all of whom reminded me of the excitement of doing science. Kenny Mok is an incredible scientist and has become a great friend. I have felt validated as a scientist whenever he asks me about experiments, and I will miss picking his brain and asking him for help when things stop working. Becky Procknow and I joined the lab at the same time, and I could not have asked for a better cohort-mate, lab-mate, co-GSI, co-first author, and friend. One of the most gratifying grad school experiences was writing a review article in co-first authorship with Becky and with the assistance of her baby, Charlotte. I'm overjoyed because we get to finish this experience together and I'm so thankful for her constant presence and support.

The PMB 2019 cohort graciously navigated the experience of completing classes and joining labs in May 2020 during the height of the COVID-19 pandemic. I am grateful to have had their company during those rough times. Christine Qabar has become a lifelong friend and part of my little California family. I am grateful for her coaching on flow cytometry and perspectives on data analysis, her endless curious questions, and the joy she brings to life. Through being her friend, I have learned about all the parallels between the people of Puerto Rico and Palestine and grasped that we are connected in many ways; the fight for freedom is just one of them.

The Boricuas in Berkeley community has been an important constant during my time in Berkeley, and I am grateful for the students who founded it in 2017 and those who keep it running. It has been a delight to have a little piece of home here.

Being more than 3,000 miles from my home in Puerto Rico has been made easier by constant contact with friends and family and consistent shipments of the coffee beans that have powered this PhD (thanks, Mom!!!). I have had a lot of family support. My mom always encouraged my academic side, and my dad encouraged my scientific curiosity. My dad helped me with my first science fair project and was the first person I ever saw get genuinely excited about science. My mom has given my brothers and me infinite love and support. She was there through grad school applications, answered many calls over the last 5.5 years, and made every effort to understand my project. My Titi Tati has also been present every step of the way, almost as a second mom, with perfect attendance at all important events. My grandfather has been a constant supportive presence. My brothers Tony and Juan have been there for me in many ways, including but not limited to personal training and ginger beer supplying. My older sisters have been invested in my happiness since I was born, and they have been infinitely caring. My friends and almost sisters Frances, Cecilia, Mariela, and Berliza have played a giant role in maintaining my sanity.

I met my partner, Darryl, a few months after moving to California in 2019, and we could not have imagined the journey that awaited. We survived a pandemic together, and he has been my rock every single day. His patience and "it's okay, we've got this" attitude ground me daily. Having his support has been the best part of this journey. He also contributed to my work directly by helping me with coding and accompanying me on a soil sampling trip. I love that we make a fantastic team, and I cannot wait to see what we do next. ☺

Finally, I thank my dad Carlos, my sister Karla, and my grandma Alice for inspiring me and watching over me, wherever they may be. AMO.

1. The corrinoid model for dissecting microbial community interactions across scales

Zoila I. Alvarez-Aponte[1]*, Rebecca R. Procknow[1]*, Michiko E. Taga[1#]

[1] Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA, 94720, United States of America

* These authors contributed equally to this work.

Corresponding author: Michiko E. Taga, Department of Plant and Microbial Biology, University of California, Berkeley, 111 Koshland Hall, Berkeley, CA 94720, United States of America. taga@berkeley.edu

Under review at *Annual Review of Microbiology*

Abstract

Microbial communities in different environments have significant impacts on global nutrient cycling and the health of host organisms. However, the complexity of microbial communities makes it difficult to investigate how the interactions between numerous microbial species, each with distinct features and metabolic capabilities, impact global processes. In this review, we describe the corrinoid model for investigating microbial community interactions across scales, from individual microbes to complex natural communities. Corrinoids, the cobalamin family of organometallic cofactors, are required for numerous metabolic processes across all domains of life but are produced only by a fraction of bacterial and archaeal species. This structurally diverse set of shared nutrients influences community structure in different ways. Knowledge about corrinoid biology at each scale informs and reinforces a robust model that can be expanded to increase our understanding of microbial communities.

1.1 Introduction

Microbial communities inhabit nearly all of Earth's environmental and host-associated niches. They are key drivers of Earth's biogeochemical cycles, the health of agricultural crops, digestion in animals, and the human immune system. The complexity of these communities leaves much to be discovered about how they function: how do numerous individual microbes, each with distinct genetically encoded capabilities, interact with one another to form communities capable of impacting global processes?

One key aspect of microbial communities that governs their function is nutrient-sharing interactions. Many microbes rely on metabolic byproducts of other community members for use as carbon, nitrogen, and energy sources and use amino acids or cofactors produced by others to fulfill their metabolic needs (1–3). Interactions between microbes shape their communities, and in turn, influence the surrounding environment (4). Understanding how the metabolic capabilities

of individual microbes influence these interactions will be a key to harnessing the potential of microbial communities to address global challenges such as climate change.

In this review, we focus on corrinoids – the cobalamin family of cofactors – which have emerged as a model for studying nutrient-sharing interactions within microbial communities (5–7). This model is particularly beneficial for investigating the mechanisms that drive community structure because it can be applied across scales of complexity. At the molecular scale, production and use of corrinoids and the genes, enzymes, and regulatory systems that control them can be dissected in specific organisms. At the organismal scale, corrinoid-specific metabolic capabilities influence how microbes grow in different conditions. At the community scale, corrinoid-based interactions collectively shape community structure and function. We present the corrinoid model as a framework for dissecting microbial interactions and community structure across scales.

1.2 Corrinoids as a model nutrient

1.2.1 Structurally diverse corrinoids are produced by different microbes and have been detected in communities

Cobalamin (vitamin B_{12}) was first discovered a century ago as an essential component of the human diet (8, 9). A cobalt-containing modified tetrapyrrole, cobalamin functions as a cofactor that harnesses the reactivity of the central cobalt ion to perform radical-based and methyltransfer reactions (10, 11) (Fig. 1.1A). Unlike other vitamins, cobalamin is produced only by certain bacteria and archaea (11–13). Also unusual is that while some microbes produce cobalamin, others produce variants with different lower ligands – collectively known as cobamides – that can carry out the same chemical reactions (14) (Fig. 1.1B). Cobamides are a subset of corrinoids that are distinguished by the presence of an upper and a lower axial ligand and are active as enzyme cofactors. Here, we use the term corrinoid to include cobamides and late precursors such as cobinamide (Cbi), while the term cobamide specifically refers to the active cofactor form (Fig. 1.1A).

Reactions catalyzed by cobamide-dependent enzymes include those involved in methionine synthesis, carbon and nucleotide metabolism, reductive dehalogenation, natural product biosynthesis, and methanogenesis (15–18). Diverse corrinoids have been detected in various host-associated and environmental samples, indicating they are widely used metabolites (19) (Fig. 1.1C). Nonetheless, cobalamin and Cbi are the only commercially available corrinoids; other corrinoids must be purified from bacterial cultures for use in experimental studies (20).

1.2.2 Corrinoids are shared nutrients

Similar to other cofactors such as thiamin, biotin, and folate, corrinoids are produced only by a fraction of the microbes that require them, indicating that they are shared among microbes (21–23). Genomic analyses of different sets of bacteria have all concluded that those predicted to be capable of *de novo* corrinoid biosynthesis (“producers”) are a minority of the species present in a particular environment and across bacteria as a whole (Fig. 1.1D). The largest bacterial genome dataset evaluated for corrinoid production and use to date estimated that 86% of sequenced bacterial species require corrinoids and of these, less than half are predicted corrinoid producers (23). Experimental studies of co-cultured microbes have demonstrated sharing of corrinoids

between producers and dependent microbes (24–33). The mechanisms of corrinoid sharing at the community scale are unknown but are an active area of research.

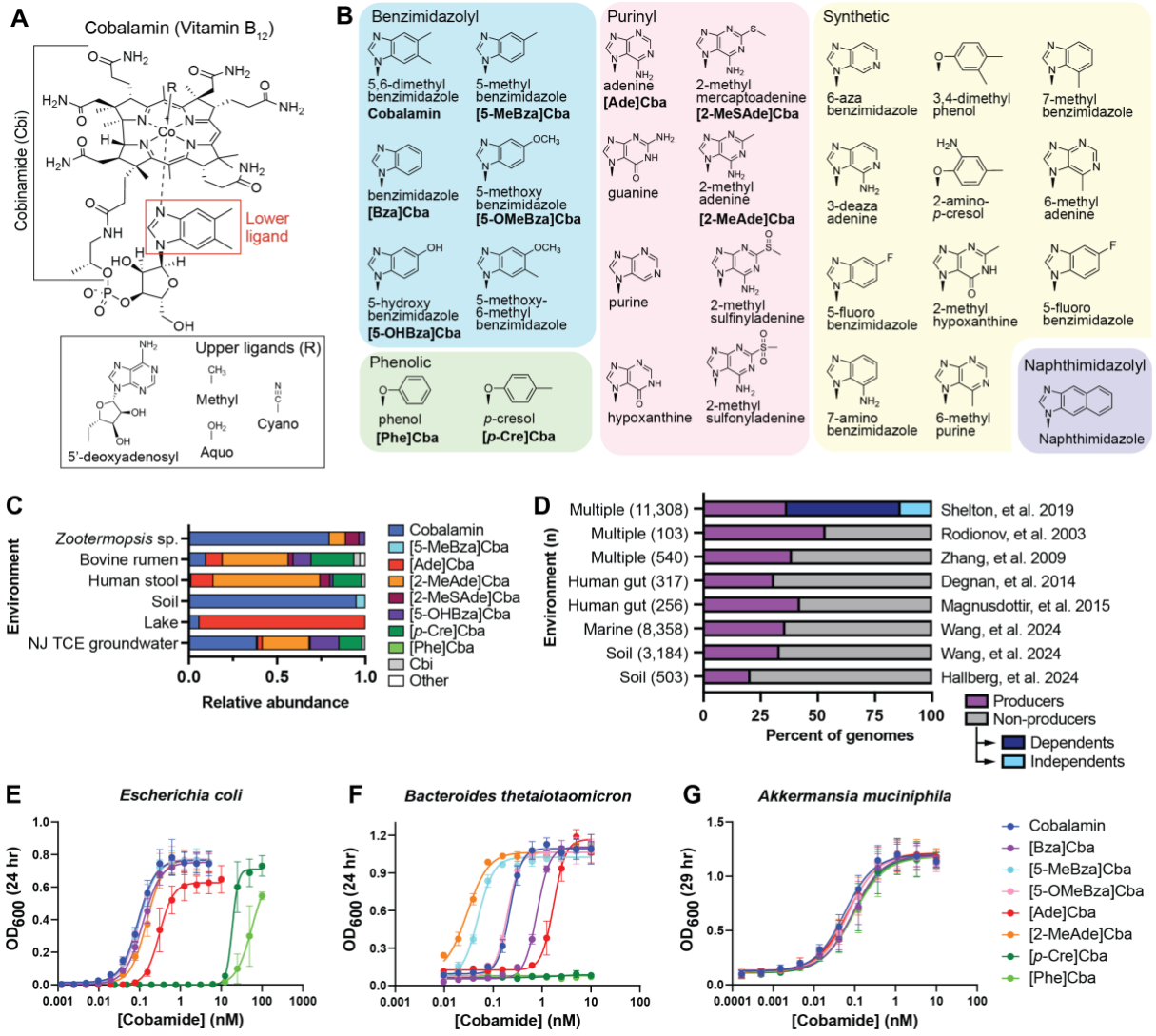


Figure 1.1. Corrinoids as a model nutrient. (A) The chemical structure of cobalamin. Cobinamide (Cbi) is an incomplete corrinoid that lacks part of the nucleotide loop including the lower ligand. Upper ligands (Adenosyl, Methyl, Cyano, and Aquo) are shown in the square. (B) Cobamide lower ligands and the the structural categories to which they belong (14, 34–38) The name of the lower ligand is given below each structure and the abbreviated cobamide name is shown in bold lettering for those that are included in other parts of this figure. (C) Corrinoid diversity as measured in different environments (19). (D) Fraction of genomes that were classified as corrinoid producers in different studies. Only studies that used genome-resolved methods are included. The environment is shown on the left and the number of genomes assessed is shown in parentheses. Shelton, 2019 differentiated the non-producer category into dependents and independents (23). (E) Corrionid-dependent growth of *Escherichia coli,* (F) *Bacteroides thetaiotaomicron,* and (G) *Akkermansia muciniphila* illustrate different corrinoid preferences (39).

1.2.3 Corrinoid preferences are reflected by differential growth and metabolism

Just as humans are metabolically primed to use cobalamin more readily than other corrinoids, most microbes have corrinoid preferences that are reflected in their growth (39). For example, *Escherichia coli* and *Bacteroides thetaiotaomicron* can use all or some tested corrinoids but vary in the relative concentrations required to support growth (Fig. 1.1E, F). An exception is *Akkermansia muciniphila*, which does not exhibit corrinoid preferences because it remodels all cobamides to a single form (40) (see section 3.5) (Fig. 1.1G). The corrinoid preferences observed in laboratory cultures suggest that the combinations of corrinoids in different environments can have distinct impacts on bacterial growth and metabolism (19, 39) (Fig. 1.1C).

1.3 The molecular scale: corrinoid biosynthesis, dependence, and specificity

Environmental corrinoid diversity can influence microbes at the molecular level in multiple ways. It is thought that microbial corrinoid preferences are due to the specificity with which cobamide-dependent enzymes, transporters, adenosyltransferases, and regulatory systems interact with particular corrinoids (36). These molecular factors are discussed separately in this section.

1.3.1 Cobamide-dependent enzymes are diverse in their functions

Cobamide-dependent enzymes are found in all domains of life and are involved in diverse metabolic processes (15–18). While the lower ligand can impact binding and reactivity, the upper ligand (Fig. 1.1A) is the part of the molecule involved in catalysis. Cobamide-dependent isomerases, typically involved in fermentation, use cobamides with a 5'-deoxyadenosyl upper ligand to generate a radical to initiate carbon skeleton rearrangements and other reactions (15, 18). Methyltransferases use cobamides to transfer methyl groups from a methyl donor to a substrate, transiently generating a cobamide with a methyl group as the upper ligand (methylcobamide) during each catalytic cycle (15, 18). These enzymes include methionine synthase, the most widespread cobamide-dependent enzyme, and multiple enzymes involved in methanogenesis and other one-carbon metabolisms (23). A third cobamide-dependent enzyme class uses cobamides with no upper ligand, instead directly using the cobalt ion to facilitate catalysis (16). These enzymes include reductive dehalogenases, involved in the detoxification of aromatic and aliphatic chlorinated organic compounds, and epoxyqueuosine reductase, which performs the final reaction in the synthesis of queuosine, a modified nucleoside in tRNA (16, 41).

The B_{12}-dependent radical S-adenosylmethionine (B_{12}-rSAM) enzymes represent a fourth class of cobamide-dependent enzymes. B_{12}-rSAM enzymes enlist two cofactors, SAM and a methylcobamide, to catalyze a variety of reactions (42). More than 200,000 predicted B_{12}-rSAM enzymes have been found, predominantly in bacterial genomes (42). For most, the substrate, product, and catalytic mechanism are unknown (17). Known reactions are functionally diverse and include the biosynthesis of vitamins, cofactors, and antibiotics, as well as protein post-translational modifications (17, 42). In the 20 years since the B_{12}-rSAM enzymes were first characterized (43), much progress has been made towards understanding this novel enzyme

class. However, the structure, chemistry, cobamide preferences, and metabolic functions of most B_{12}-rSAM enzymes remain to be elucidated.

Cobamide preferences have been studied in several enzymes. *In vitro* studies have shown that methylmalonyl-CoA mutase (MCM) orthologs from different organisms have different cobamide preferences (36, 44, 45). These studies have shown that Even small differences in lower ligand structure can greatly impact corrinoid-enzyme binding (36). Further, MCM orthologs from different bacteria and from humans have distinct binding affinities for different cobamides, suggesting that certain sequence differences between orthologs exist that confer distinct preferences (36, 37). Other examples of corrinoid preference in enzymes have been observed for MetH, glutamate mutase, ethanolamine ammonia lyase, and 2-methyleneglutarate mutase (45–49). These observations highlight the influence of cobamide structure on enzyme function and implicate cobamide preferences in enzymes as major contributors to preferences at the organismal scale.

1.3.2 Biosynthesis of corrinoids results in diverse structures

De novo corrinoid biosynthesis in bacteria requires approximately 30 genes (11, 50). The first set of steps involves assembling uroporphyrinogen III, a precursor shared with other tetrapyrrole biosynthesis pathways (11, 50). These steps are followed by corrin ring synthesis via either an anaerobic or aerobic route (11, 50). The cobalt ion is installed prior to ring modifications in the anaerobic pathway and after completion of these modifications in the aerobic pathway (11, 50). Adenosylation at the upper (β) ligand position occurs next to form adenosylcobyric acid (50). Finally, nucleotide loop assembly is followed by the attachment of the lower (α) ligand to form the cobamide (11, 50). Some archaea are known to make corrinoids using biosynthesis genes homologous to those from the bacterial pathway, but all the enzymes involved have yet to be characterized (51–55).

The diversity of cobamides produced by different organisms stems from the ability to synthesize and attach different lower ligand bases (Fig. 1.1B). Unlike the purine and phenolic bases, which have other roles in metabolism, the benzimidazole bases are thought to be specific to cobamides and have enzymes dedicated to their production. 5,6-dimethylbenzimidazole (DMB), the lower ligand of cobalamin, is synthesized aerobically from a flavin cofactor by the BluB enzyme (56–58). One quarter of sequenced bacterial genomes contain *bluB* homologs (23), suggesting widespread production of cobalamin, though it is not known how many are non-functional pseudogenes, as recently characterized in a *Roseovarius* species (32). The anaerobic biosynthesis of DMB requires the *bzaABCDE* or *bzaFCDE* genes, responsible for conversion of the purine precursor 5-aminoimidazole ribotide to DMB (59, 60). The three intermediates in this pathway, 5-OHBza, 5-OMeBza, and 5-OMe-6-MeBza, are also found as lower ligands, and organisms that produce these cobamides have the corresponding sets of *bza* genes in their genomes (59) (Fig. 1.1B and C).

Approximately 13% of bacterial species are predicted “salvagers” that import corrinoid precursors and use them as building blocks to make cobamides (23, 61). This strategy is thought to reduce the cost of biosynthesis. The genomes of salvagers contain an incomplete corrinoid biosynthesis pathway, lacking one or more initial steps but possessing genes for later steps in the

pathway (23). Experimentally, salvagers are defined by the ability to grow and produce a cobamide when certain precursors such as 5-aminolevulinic acid or Cbi are available (7, 23, 26, 61–64). α-ribazole, the activated form of DMB, can also be salvaged to form cobalamin when late steps in the pathway are absent (65). Cbi and α-ribazole have been detected in environments such as soil, animal gastrointestinal tracts, and ocean waters, suggesting they are available for uptake by salvagers (19, 32, 66, 67).

1.3.3 Corrinoid uptake is an active process

Efficient corrinoid uptake from environments with limiting corrinoid concentrations requires active transport. The Btu(B)FCD system is the most widespread corrinoid uptake system in bacteria (68), consisting of the ABC-type transport complex BtuFCD and, in Gram-negative bacteria, the TonB-dependent outer membrane transporter BtuB (11, 69–71). *In vitro* studies of *E. coli* BtuB show that it interacts with the corrin ring, and thus is not specific for corrinoids with particular upper or lower ligands (72, 73). Likewise, a *Bacillus subtilis* strain that overexpresses *btuFCD* imports cobamides and Cbi non-specifically (47). An alternative ECF-type transporter coupled with the corrinoid-specific substrate-binding protein CbrT was found to transport corrinoids in *Lactobacillus delbrueckii* (74). Additionally, the nonspecific transporter BacA can transport cobalamin bidirectionally in *Mycobacterium tuberculosis* (75, 76). It remains unknown whether these transporters are specific for particular corrinoids.

Research on corrinoid uptake in *Bacteroides* species has introduced additional components of the Btu system. BtuG is an outer membrane protein that binds cobalamin and cobinamide with remarkably high (femtomolar) affinity and is thought to transfer extracellular corrinoids to BtuB for transport (77, 78). BtuH is a cobalamin-binding protein found in *Bacteroides* that has no known function but is often encoded in operons containing BtuB (79). BtuM, an inner membrane protein in *Thiobacillus denitrificans,* has the ability to transport cobalamin and is thought to decyanate it (80).

The influence of transport on corrinoid preferences has been studied mostly in *Bacteroides* species. Most human gut *Bacteroides* genomes contain multiple *btuB* paralogs with highly divergent sequences, suggesting they may be specialized for different corrinoids (68). These paralogs can confer corrinoid-dependent competitive advantages for cells grown with different corrinoids (see section 5.2).

1.3.4 Adenosylation occurs following uptake

Adenosylation is a step in biosynthesis that readies corrinoids for catalysis by enzymes that use cobamides with a deoxyadenosyl upper ligand (adenosylcobamides, Fig. 1.1A) (16). The reactivity of adenosylcobamides, which uniquely enables them to facilitate radical chemistry, also makes the upper ligand labile, dissociating readily upon exposure to light (46, 81). Thus, the most common form of cobalamin in the environment is hydroxocobalamin (OHCbl) (82). Cyanocobalamin, the vitamin form of cobalamin, has a more stable cyanide ion as the upper ligand (Fig. 1.1A).

Most bacteria adenosylate imported corrinoids via one of three types of adenosyltransferases, which vary in sequence and bind substrates in different conformations. These enzymes, present in 76% of sequenced bacterial species, have been mostly studied in *Salmonella enterica*, which encodes all three classes: CobA, EutT, and PduO (23, 83). CobA is involved in the *de novo* synthesis of AdoCbl, while EutT and PduO are co-expressed with adenosylcobamide-dependent enzymes required for the metabolism of specific substrates (83). An additional function of adenosyltransferases is that of chaperones that deliver the adenosylcobamide to cobamide-dependent enzymes (83, 84), which may explain their genetic association with specific cobamide-dependent metabolic processes.

The ability of bacteria to use diverse corrinoids suggests adenosyltransferases can act on many different corrinoids, but this has yet to be verified. In support of this inference, corrinoids with different lower ligand structures can be recovered in their adenosylated form from a *B. subtilis btuFCDR* overexpression strain cultured with cyanated corrinoids (47). Further, mutations that increase expression of the adenosyltransferase BtuR enable *E. coli* to grow more robustly with a less-preferred corrinoid (85). These observations suggest that bacteria can adenosylate diverse corrinoids, and that adenosylation may play a part in defining bacterial corrinoid preferences.

1.3.5 Remodeling enables conversion of different cobamides to a preferred form

Corrinoid-dependent organisms are reliant on the corrinoids available in their environment, which could pose a vulnerability if their preferred corrinoids are unavailable. Thus, some bacteria, archaea, and microalgae have developed cobamide remodeling as a strategy to make use of any available cobamide, regardless of its structure, while avoiding the metabolic burden of *de novo* biosynthesis. Remodeling involves removing a portion of the nucleotide loop, including the lower ligand, and replacing it with a different lower ligand (61, 86). The three remodeling enzymes characterized to date are each unique in their sequence, structure, reaction catalyzed, and substrate specificity (61). CbiZ hydrolyzes the amide bond in the nucleotide loop and is the most widespread, found in diverse bacteria and archaea (86, 87). CbiR, which hydrolyzes the ribose-phosphate bond in the nucleotide loop, was discovered in *A. muciniphila*, and homologs are present in a number of bacterial species (40). Following the CbiZ or CbiR hydrolysis reaction, additional enzymes involved in the final steps of corrinoid biosynthesis are required to rebuild the cobamide with a more favorable lower ligand. A third remodeling function has been discovered in *Vibrio cholerae*. A variant form of CobS, an enzyme required for the final step in corrinoid biosynthesis, is involved in "direct remodeling" by both removing and replacing the lower ligand in a cobamide (88, 89). The differences among the three remodeling enzymes suggest that each evolved independently from the others (40), highlighting the importance of lower ligand structure on bacterial fitness. A fourth remodeling enzyme likely exists in algae but has not been identified (90), further reinforcing microbes' need to obtain corrinoids that function robustly in their metabolism. The ability of microbes to alter corrinoid structure could impact environmental corrinoid profiles.

1.3.6 Regulation of corrinoid-dependent processes is mediated by riboswitches

Bacteria rely on a fine-tuned and rapid regulatory mechanism to respond to changing corrinoid availability. Riboswitches, segments of mRNA that regulate gene expression by directly binding

to an intracellular metabolite, are the predominant method of corrinoid-responsive regulation in bacteria (91, 92). Most corrinoid riboswitches are involved in maintaining homeostasis by reducing the expression of genes that become unnecessary when corrinoids are present in excess, such as those involved in corrinoid transport, biosynthesis, and corrinoid-independent redundant pathways (92–95).

While most riboswitch classes are defined by their ability to recognize a single molecule with high specificity, corrinoid riboswitches have the unique ability to respond to corrinoids with different lower ligand structures (47). Two corrinoid riboswitch classes are defined by their affinity for corrinoids with either a large (5'-deoxyadenosyl) or small (methyl, cyano, and aquo) upper ligand (93, 96, 97). Likewise, riboswitches can be classified into two subclasses based on their affinity for different lower ligands. The "promiscuous" subclass responds equally to Cbi and to cobamides with various lower ligands, while the "semi-selective" subclass responds only to certain cobamides and displays preferences for specific lower ligands (47). For one riboswitch studied in depth from *Priestia megaterium*, the corrinoid responsiveness is tuned to the preferences of the organism's cobamide-dependent enzyme (47). This suggests that bacteria have evolved to maintain a delicate balance between regulation and function in order to respond to different corrinoids in their environment.

Though the majority of corrinoid riboswitches repress gene expression to maintain homeostasis, a riboswitch that activates gene expression in response to corrinoids was recently discovered. This riboswitch responds preferentially to Cbi and regulates the expression of *cobT*, a gene required for activating the lower ligand before attachment to a precursor derived from Cbi (98). Thus, the riboswitch is tuned for maximal activation in response to a CobT substrate. Additional corrinoid riboswitches have been found upstream of unknown genes and genes involved in processes not known to use corrinoids (95), suggesting there is more to learn about corrinoid-based regulation and function.

1.4 The organismal scale: Defining ecological roles of individual organisms

The corrinoid-specific metabolism an organism encodes at the molecular scale influences growth and metabolic capabilities at the organismal scale. In this section, we use the categories of "dependent," "independent," and "producer" as a framework for nutrient interactions between species. Each of these roles can be defined genomically based on the presence or absence of specific genes, and experimentally based on the results of laboratory growth assays (99). A new subcategory of producers, "providers," has been differentiated by the ability to release corrinoids to the extracellular environment (31, 99, 100). While corrinoid metabolism is common to organisms in all three domains, it has been most extensively characterized in bacteria.

1.4.1 "Dependents" use but cannot synthesize their own corrinoids

Dependents are defined as organisms that carry out corrinoid-dependent processes but cannot synthesize corrinoids. While the term "auxotroph" refers to organisms with an absolute requirement for a nutrient they are incapable of synthesizing, the dependent category additionally includes those for which the nutrient expands the organism's metabolic capabilities. Dependents can be characterized genomically based on the presence of one or more genes encoding

cobamide-dependent enzymes and absence of a complete set of genes necessary for biosynthesis (23). Experimentally, they can be defined as organisms that grow or carry out a particular metabolic process only when a corrinoid is provided (99). Corrinoid dependence is widespread, including most eukaryotes and approximately half of bacterial species (11, 23). Corrinoid "salvagers" are a subset of dependents that import and use corrinoid precursors to synthesize cobamides (see section 3.2).

Dependence can vary based on which cobamide-dependent reactions an organism carries out. First, some organisms encode corrinoid-dependent functions that are not essential, but enable them to access certain resources when corrinoids are available (23, 101). This is the case for the catabolism of ethanolamine, propanediol, and certain amino acids, which are not available in all environments, but having the ability to use them as a nutrient source represents an opportunity for niche expansion (102). In addition, corrinoid-independent alternatives exist for some corrinoid-dependent pathways such as methionine synthesis and ribonucleotide reductase (23, 103–106). Compared to some cobamide-dependent enzymes, the cobamide-independent alternatives can be less resilient because of an inability to function under certain stress conditions (107–111). This suggests an evolutionary tradeoff between reliance on external corrinoids or corrinoid-independent alternatives that function in limited conditions. Finally, the cobamide-dependent epoxyqueuosine reductase QueG, which catalyzes a post-transcriptional tRNA modification, can be inactivated without an observable growth phenotype in *E. coli* (41), despite its presence in over half of sequenced bacterial species (23). Given that bacteria exist that lack cobamide-dependent enzymes except QueG (23), QueG may have an unknown but important role under environmental conditions that are not mimicked by laboratory culture.

Organisms with no cobamide-dependent enzymes and no biosynthesis pathway are considered to be corrinoid-independent (23). They can be defined genomically by the absence of cobamide-dependent enzymes and biosynthetic genes, and experimentally by growth without added corrinoid and the absence of corrinoid production (99).

1.4.2 "Producers" synthesize corrinoids *de novo*

Corrinoid producers are defined genomically as encoding a complete biosynthesis pathway, and experimentally, by detecting corrinoids in microbial cultures without corrinoid supplementation (23, 99). Until recently, all genomically predicted producers were found to encode one or more cobamide-dependent enzymes in their genomes (23); the one exception is a soil archaeal metagenome-assembled genome (112). Many producers have corrinoid uptake genes which presumably allow them to circumvent the energetic burden of corrinoid biosynthesis by importing corrinoids when available and downregulating biosynthesis (68).

Corrinoid producers typically synthesize a single corrinoid, though some produce two or more, including organisms that produce different corrinoids under different conditions. An example of the latter is *Propionibacterium freudenreichii*, which synthesizes the corrin ring via the anaerobic pathway and the lower ligand DMB via the oxygen-dependent synthase BluB (113, 114). This leads to production of [Ade]Cba under anoxic conditions and cobalamin under microoxic conditions (113–115). Because corrinoid structure can depend on the growth condition, it

remains unknown which corrinoids microbes produce in their natural communities where conditions differ from laboratory culture.

In the laboratory, most producers can attach non-native lower ligands, including synthetic lower ligands, when they are added to growth media. This process is known as guided biosynthesis and was initially developed for production of commercially unavailable corrinoids (20, 66, 116). Guided biosynthesis can enable a producer to synthesize a corrinoid that functions more efficiently for its metabolism than its native corrinoid (117). Conversely, guided biosynthesis can result in production of a corrinoid that does not efficiently support metabolism, resulting in growth arrest (35, 118–120). Because DMB has been detected in soil (82), and DMB and other benzimidazoles have been found in multiple environments, it is possible that guided biosynthesis occurs in natural environments (121). Thus, production of free lower ligand bases, coupled with guided biosynthesis by producers, may influence the structures of corrinoids produced in microbial communities.

1.4.3 "Providers" are the subset of producers that share corrinoids with other microbes

Producers are the only source of corrinoids for dependents, yet the corrinoids they synthesize are not always available to other community members. Experiments with producer bacteria isolated from marine and soil environments found that only a subset of producers release corrinoids into culture supernatants or support dependent growth in coculture (99, 100). These producers form the novel subcategory of "providers" (31, 99, 100). Corrinoid providers likely play a key ecological role because they sustain dependents in communities.

The mechanisms of corrinoid release from provider cells remain unknown. Bacteriophage lysis has been proposed as a general mechanism of releasing intracellular nutrients, based on the ability of phage lysates to robustly support the growth of certain amino acid auxotrophs (122) and the temporal correlation between prophage induction of a marine *Roseovarius* and corrinoid providing to *Colwellia* (32). Additionally, the general transporter BacA was recently found to be capable of transporting cobalamin bidirectionally, suggesting it could mediate corrinoid release (75). Finally, it is possible that corrinoids leak out of cells or are actively exported, possibly to protect cells from harmful buildup of corrinoids or to support the growth of a mutualistic partner (123). When corrinoid providing is part of a co-evolved mutualism, specific partners may trigger corrinoid release through chemical or physical signals (25, 28).

1.5 The community scale: Corrinoid-based interactions among microbes in laboratory cocultures and natural communities

While 86% of all sequenced bacterial species are predicted to use corrinoids (23), only 21% to 53% are predicted producers across different environments (Fig. 1.1D) (22, 23, 68, 95, 112, 124, 125). This disparity suggests that few producers support the majority of the community through corrinoid production and release. Without access to corrinoids, many corrinoid-dependent bacteria would presumably cease to fulfill their ecological roles. Thus, it stands to reason that corrinoid-sharing interactions are essential for microbial community function. Corrinoid sharing has been studied in co-cultures and in multi-member communities.

Research on corrinoids at the molecular and organismal scales has enabled genomic predictions and hypotheses about corrinoid-based interactions at the community scale. Metagenomic analyses of soil, ocean, and skin microbial communities have revealed that producer species are relatively scarce, while dependent species abound (Fig. 1.1D) (22, 23, 68, 82, 95, 112, 124–126). While the focus of these studies is on ratios of species, more work is needed to understand the relationships between producers and dependents in different environments. The low ratio of producer to dependent species, which means an even lower ratio of providers to dependents, suggests that on average, each provider supports the corrinoid-dependent metabolism of several dependents. However, these studies analyzed corrinoid metabolism at the species level but did not account for the population size of each species. While a minority of bacterial species are corrinoid producers, it is possible that these species are present in high abundances in certain environments. Future research should elucidate the ratio of producer to dependent cells in communities.

1.5.1 Corrinoids are shared between producers and dependents in coculture

Laboratory co-cultures have demonstrated that corrinoid sharing occurs between pairs of microbes (24, 27, 29, 31). Several algae-bacteria co-cultures provide examples of corrinoid-mediated symbioses; bacteria provide a corrinoid to algae in exchange for nutrients such as fixed carbon or other B vitamins (25, 28). Bacteria have also been shown to provide corrinoids to other bacteria, which can have implications for environmentally relevant processes such as methanotrophy and bioremediation by allowing for growth of corrinoid-dependent organisms responsible for carrying out these processes (27, 30, 33). In addition to interactions between producers and dependents, two recent examples provide evidence of cooperation between salvager and dependent microbes. One study used engineered *E. coli* strains to show that a salvager could provide cobalamin to a dependent and support its growth (26). A second study on marine bacteria showed that two bacterial strains, one that provided α-ribazole and another that provided Cbi, cooperated to support each other's growth and that of a corrinoid-dependent diatom (32). These examples of cooperation by sharing corrinoids and corrinoid precursors illustrate that corrinoid sharing can occur between two microbes and suggest it could be an important feature of interaction networks in complex communities.

1.5.2 Microbial communities are shaped by corrinoid preferences

Given that the growth of individual bacteria is influenced by corrinoid structure (39) (Fig. 1.1E, F, G), an outstanding question is whether corrinoids can influence bacterial growth in the context of a community. It stands to reason that an increase in abundance of a certain corrinoid could alter the composition of a community by conferring a growth advantage to the microbes that prefer it. One study presented functional links between corrinoid transport and competition between bacteria (68). When mutants of *B. thetaiotaomicron* containing a single *btuB* paralog competed in media containing different corrinoids, clear corrinoid-specific advantages were conferred by different *btuB* paralogs (68). In Bacteroidetes, *btuB* homologs are highly divergent and are often found on mobile genetic elements (68, 127), suggesting that genetically sampling diverse *btuB* homologs among a population is advantageous. These observations suggest that competition for corrinoids is important for shaping gut microbial communities.

Recent studies have shown the effects of adding corrinoids to different microbial communities, though most have focused only on the impact of adding cobalamin. In ocean mesocosms, cobalamin addition altered community composition at the phylum level (128). In mice, cobalamin supplementation resulted in a dramatic decrease in Bacteroidetes and had minimal effects on other taxa in the gut (129). Additional studies that compared cobalamins with different upper ligands found that the upper ligand can also affect community composition, particularly Bacteroidetes levels (130, 131). Dramatic changes in the abundance of Bacteroidetes are consistent with the observation that organisms in this phylum are broadly corrinoid-dependent (23). Further, decreases in Bacteroidetes upon the addition of cobalamin may reflect a low preference for this corrinoid. The impact of adding cobalamin may be due to gut environments containing relatively low levels of cobalamin and other benzimidazolyl cobamides (19, 66, 67), though the corrinoid composition of the gut environments in these studies were not measured.

To date, one study has investigated the impact of lower ligand diversity on microbial communities (112). In microcosms derived from a grassland soil, where cobalamin is the most abundant corrinoid, the addition of corrinoids other than cobalamin resulted in a significant but transient shift in the bacterial 16S community composition. Similarly, in enrichment cultures derived from the same soil, 20% of microbial orders were influenced by corrinoid treatment (112). These results indicate that soil microbial communities respond to corrinoids not already present in their environment but recover from these perturbations over time. The marked effect of adding different corrinoids highlights the influence of corrinoid structure on bacterial growth, even in a complex community where numerous metabolic processes occur. Corrinoid amendment studies showcase the importance of corrinoids for bacterial growth and validate the utility of corrinoids as model nutrients. Building on this knowledge by dissecting how corrinoid addition influences individual microbes that shape the community could lead to future applications of corrinoids as tools to modulate community structure and function.

1.6 Corrinoids as a model for interaction studies across scales

In this review, we have demonstrated how the molecular, organismal, and community scales build on each other to provide a framework for understanding nutrient sharing interactions. Figure 1.2 illustrates how the model nutrient approach contributes to the understanding of microbial community interactions across scales of complexity. Using the corrinoid model, characterization of corrinoid production and use at the molecular scale enabled the development of the producer/dependent framework, which in turn serves to define interactions at the community scale. Thus, the corrinoid model has begun to generate a deeper understanding of microbial community interactions. Outstanding questions remain at each scale that represent areas for future research.

At the molecular scale, the ability to make sequence-based predictions about an organism's corrinoid preferences remains elusive. Although clear patterns have been observed in the structure-function relationships between corrinoid structure and binding to enzymes, identifying specific sequence motifs that confer these corrinoid preferences has not yet been achieved. Developing tools to predict corrinoid preferences for cobamide-dependent enzymes, transporters, and riboswitches will enable better predictions of metabolic interactions and ecological roles in communities.

At the organismal scale, many microbes encode both a cobamide-dependent enzyme and its cobamide-independent counterpart – for example, both methionine synthase genes *metE* and *metH* are found in 43% of bacterial genomes (23) – but the advantage of having both is unknown. Biochemical studies have shown that the cobamide-dependent methionine synthase is more efficient than its independent counterpart, and loss of *METE* confers a fitness advantage in the alga *Chlamydomonas reinhardtii*, suggesting a tradeoff between reliance on an externally supplied corrinoid and use of an inefficient enzyme (107, 132). Alternatively, it may be beneficial to have more than one enzyme capable of performing the same function. Additionally, while bacterial corrinoid biology is well characterized, knowledge about archaeal corrinoid biology lags behind, despite genomic evidence that archaea may represent a large proportion of corrinoid producers (133, 134).

Additional questions remain regarding the evolution of nutritional dependence that may be answered by continued investigation of corrinoid production and dependence. How the roles of producer and dependent emerge and are maintained are complex questions that require further study. The evolution and maintenance of nutritional dependence is supported by the Black Queen hypothesis, which suggests that the loss of the corrinoid biosynthesis pathway may be favored when corrinoids are made available by producers (135), though it is still not known how corrinoids are released from cells. Likewise, corrinoid structural diversity can be explained by the nutrient encryption hypothesis (136), which proposes structural diversification as a way to limit other organisms from accessing a nutrient. The impact of diverse corrinoids on community structure also requires further study. Ongoing work suggests that corrinoids could be used to modulate microbial communities by leveraging bacterial corrinoid preferences to promote or inhibit growth of specific organisms. Probing interaction mechanisms in increasingly complex synthetic consortia will be an important step toward this goal.

While application of the model nutrient approach at the three scales we have presented has offered extensive knowledge about microbial interactions, we have yet to apply the model to understand the effects of interactions beyond the community level. Microbial community interactions are known to impact the ecosystem scale by influencing host health and global nutrient cycling (Fig. 1.2). Future work across scales will likely reveal how corrinoid-based microbial community interactions impact host and ecosystem processes, expanding the corrinoid model to new scales.

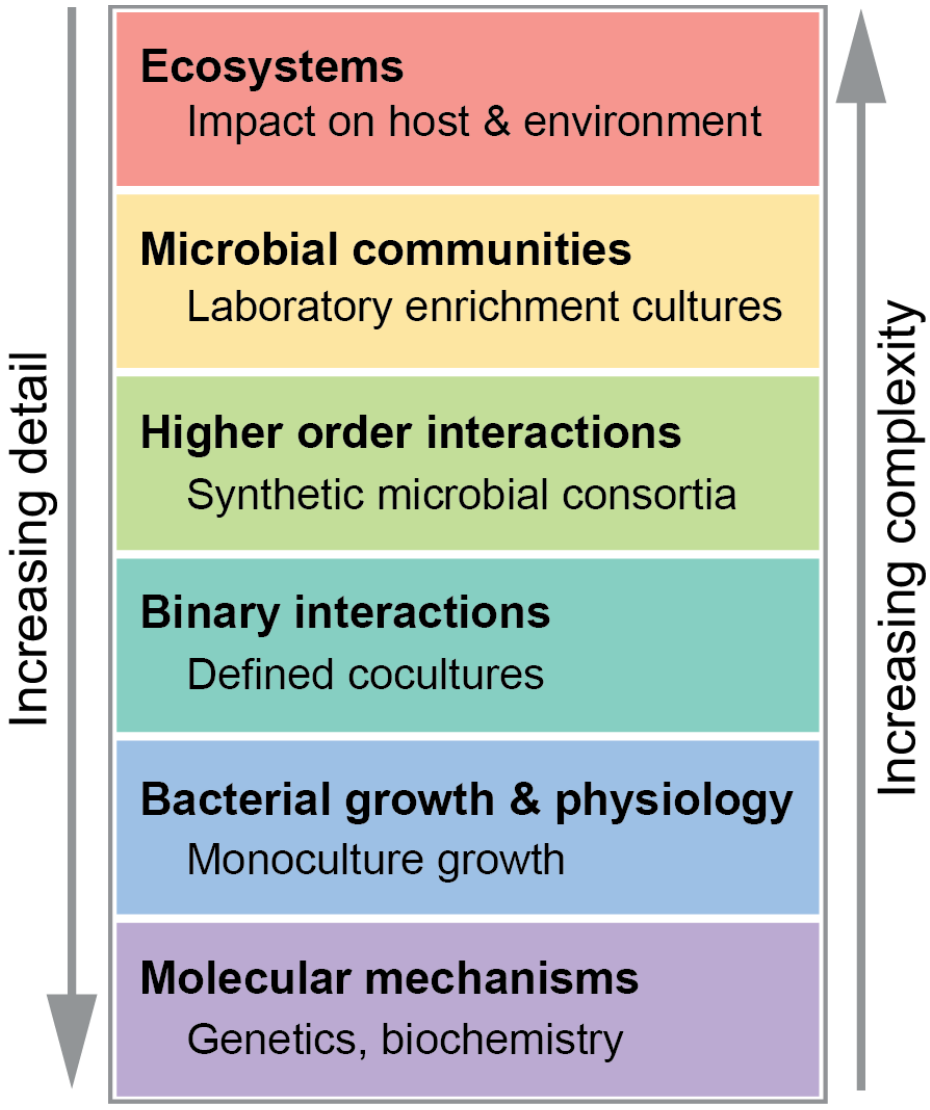


Figure 1.2. The corrinoid model increases understanding of nutritional interactions across scales.

Acknowledgements

We thank Bernhard Kräutler, Eleanor Wang, and Kenny Mok for pointing us to foundational literature. We thank members of the Taga Lab for helpful discussions and Kenny Mok, Eleanor Wang, Dennis Suazo, Farrah Sherdil, and Zachary Hallberg for critical reading of the manuscript. This work was supported by NIH grant R35GM139633 to MET.

2. Phylogenetic distribution and experimental characterization of corrinoid production and dependence in soil bacterial isolates

Zoila I. Alvarez-Aponte,[1] Alekhya M. Govindaraju,[1] Zachary F. Hallberg,[1] Alexa M. Nicolas,[1] Myka A. Green,[1] Kenny C. Mok,[1] Citlali Fonseca-Garcia,[1,3] Devin Coleman-Derr,[1,3] Eoin L. Brodie,[2,4] Hans K. Carlson,[2] and Michiko E. Taga[1#]

[1] Department of Plant & Microbial Biology, University of California, Berkeley, Berkeley, CA, USA
[2] Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[3] Plant Gene Expression Center, USDA-ARS, Albany, CA, USA
[4] Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, CA, USA

Corresponding author: Michiko E. Taga, Department of Plant and Microbial Biology, University of California, Berkeley, 111 Koshland Hall, Berkeley, CA 94720, USA. taga@berkeley.edu

Published in *The ISME Journal*, 2024
Supplementary tables can be found with the published manuscript online at https://doi.org/10.1093/ismejo/wrae068

Abstract

Soil microbial communities impact carbon sequestration and release, biogeochemical cycling, and agricultural yields. These global effects rely on metabolic interactions that modulate community composition and function. However, the physicochemical and taxonomic complexity of soil and the scarcity of available isolates for phenotypic testing are significant barriers to studying soil microbial interactions. Corrinoids—the vitamin B_{12} family of cofactors—are critical for microbial metabolism, yet they are synthesized by only a subset of microbiome members. Here, we evaluated corrinoid production and dependence in soil bacteria as a model to investigate the ecological roles of microorganisms involved in metabolic interactions. We isolated and characterized a taxonomically diverse collection of 161 soil bacteria from a single study site. Most corrinoid-dependent bacteria in the collection prefer B_{12} over other corrinoids, while all tested producers synthesize B_{12}, indicating metabolic compatibility between producers and dependents in the collection. Furthermore, a subset of producers release B_{12} at levels sufficient to support dependent isolates in laboratory culture at estimated ratios of up to 1,000 dependents per producer. Within our isolate collection, we did not find strong phylogenetic patterns in corrinoid production or dependence. Upon investigating trends in the phylogenetic dispersion of corrinoid metabolism categories across sequenced bacteria from various environments, we found that these traits are conserved in 47 out of 85 genera. Together, these phenotypic and genomic results provide evidence for corrinoid-based metabolic interactions among bacteria and provide a framework for the study of nutrient-sharing ecological interactions in microbial communities.

2.1 Introduction

Microorganisms engage in metabolic interactions that collectively define ecological networks in communities (4). Microbial interactions can be key mediators of community function, and disruptions to interactions can restructure whole communities (137, 138). Thus, it is crucial to disentangle microbial interactions and generate a predictive understanding of nutritional influences on communities and, in turn, on the environment.

Experimental and computational studies have shown that microorganisms commonly lack the ability to synthesize all of the metabolites they require (1, 139). For example, many microorganisms are unable to synthesize certain cofactors and amino acids, and therefore must acquire these nutrients from other organisms in the environment (5, 140, 141). As a consequence, microbial communities are composed of "producer" and "dependent" organisms that synthesize and require a given nutrient, respectively. The complexity of microbial communities is due, in part, to this network of interactions arising from interdependence among microorganisms that produce and require a range of different nutrients. Such interdependence may develop because loss of biosynthesis genes can be evolutionarily favored in contexts where required nutrients are abundant in the environment or can be acquired from other microorganisms (107, 135).

Investigating the molecular mechanisms and community impacts of nutritional interactions is challenging for many reasons. First, many microbial communities are functionally diverse and contain numerous metabolites that are produced, used, and chemically transformed by community members. Second, while some metabolic capabilities can be inferred from genomic data, these analyses currently lack spatial and temporal resolution, making it difficult to predict how interactions among community members may be impacted by the metabolic activities of a single microbe. Third, because most microorganisms have not been isolated in pure culture (142), metabolic predictions of the uncultured majority have yet to be confirmed. The challenges common to microbiome studies are amplified in soil due to its taxonomic diversity, physical and chemical heterogeneity, and environmental fluctuations, as well as disturbances due to animal or human activities (143–145). Nonetheless, generating mechanistic knowledge of nutrient cycling in soil communities is essential because of their broad impacts on the health of our planet (146, 147). We address these challenges by studying the ecological roles of microorganisms in relation to one class of model shared nutrients in a collection of newly isolated soil bacteria.

We took a reductionist approach to investigate nutrient production and dependence by focusing on corrinoids as a representative class of shared metabolites. Corrinoids are produced by a subset of bacteria and archaea and include the vitamin B_{12} (cobalamin) family of cobalt-containing cobamide cofactors and their biosynthetic precursors (Fig. 2.1A). Corrinoids are required cofactors for methionine synthesis and propionate metabolism in most eukaryotes, and additionally are used by prokaryotes for other diverse processes such as mercury methylation, natural product biosynthesis, nucleotide synthesis, and numerous carbon and nitrogen transformations (7). Corrinoid-sharing interactions between producer and dependent microorganisms have been observed in laboratory co-cultures of bacteria (27, 30, 33), bacteria-microeukaryote pairs (28, 100, 148), and in higher-order consortia (149), and are thought to be prevalent in the gut microbiome (141).

Computational predictions further support the hypothesis that corrinoids are broadly shared, as 37% of sequenced bacterial species are predicted corrinoid producers and 49% are dependents (23). The remaining 14% are predicted not to produce or use corrinoids, fulfilling their metabolic needs via corrinoid-independent pathways, and are therefore considered "independents" (Fig. 2.1B) (23). Given that dependents and producers coexist in the same environments (22, 112, 126), we hypothesize that many sharing interactions have yet to be described.

An aspect of corrinoids that influences their function as shared nutrients is their structural diversity (Fig. 2.1A), which has been shown to impact function: microbial preferences for particular corrinoid structures are apparent in their differential growth responses to corrinoids (7, 90). Distinct groups of corrinoids have been detected in a variety of host-associated and environmental microbial communities, suggesting that microorganisms encounter diverse corrinoids in nature (19). Because most corrinoids are not commercially available, nearly all research on corrinoids has been performed only with B_{12} (7). Given that some microorganisms require corrinoids other than B_{12} (35), it is likely that novel bacteria that could not have been isolated on B_{12} will be culturable on other corrinoids.

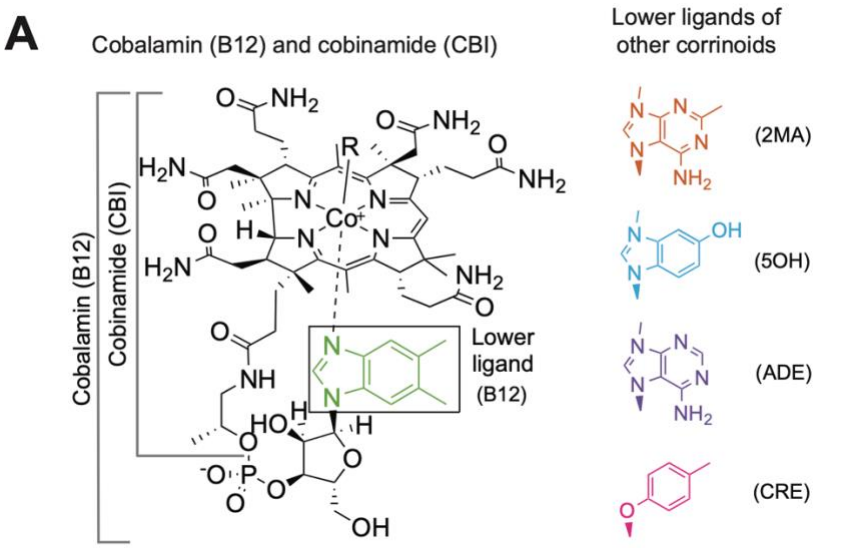
In this study, we investigated the impact of corrinoid structure on bacterial growth and isolation from a California annual grassland soil. We generated a diverse collection of soil bacterial isolates and experimentally confirmed that all three predicted corrinoid metabolism categories – producer, dependent, and independent – are represented in the soil microbial community. Further, the structure and amount of corrinoid released by some producers is compatible with the requirements of corrinoid-dependent isolates, suggesting that corrinoid sharing can occur between these bacteria in soil. A broad analysis of bacterial genomes revealed that it is possible to classify some bacterial taxa as corrinoid producers, dependents, or independents based on phylogeny alone. Our results provide an ecological framework for understanding nutritional interactions in soil through the lens of corrinoids.

2.2 Results

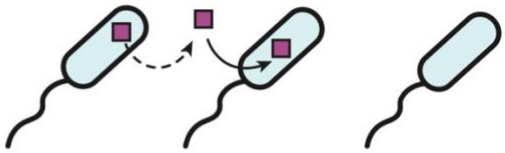
2.2.1 Generating a collection of 161 bacterial isolates from soil

To generate a collection of soil bacterial isolates with a diversity of corrinoid requirements, we performed the limiting dilution method (150) in 384-well plates containing media with one of six different corrinoids (Fig. 2.1A) or no corrinoid. The isolation was performed in two rounds (Fig. 2.2, Nov 2019 and Apr 2021). In the first isolation, 8,064 inoculated wells yielded only 2.4% of wells with detectable growth. 16S rRNA gene sequencing of the resulting cultures revealed that 78% were clonal. In the second isolation round, 8,064 wells were inoculated and 37% contained detectable growth, yet only 5.8% of cultures were clonal. By statistical metrics alone, we expected 80% of cultures to be clonal when 30% of wells contained detectable growth. These results suggest that smaller soil inocula lead to more clonal cultures, possibly because cell aggregation is more prevalent in soil than in the gut environment, where the same method led to the statistically predicted result (150). 20 phyla were found in the total collection (Fig. 2.3), but after selecting clonal cultures, reviving them from frozen stocks, purifying, and archiving, our

collection contained 161 bacterial isolates, representing four phyla, which were used for subsequent analyses (Fig. 2.2).



B



	Producer	Dependent	Independent
Experimental (E)			
Growth with no corrinoid	+	-	+
Corrinoid detected	+	-	-
Genome (G)			
Biosynthesis pathway	+	-	-
Corrinoid dependent enzymes	+*	+	-

Figure 2.1. Diversity of corrinoid lower ligands and metabolic roles. (A) Chemical structure of cobalamin (B12), lower ligands of other corrinoids used in this research, and 3-letter abbreviations. From the top, 2-methyladeninylcobamide (2MA), 5-hydroxybenzimidazolylcobamide (5OH), adeninylcobamide (ADE), and cresolylcobamide (CRE). Cobinamide (CBI), shown on the left, is an incomplete corrinoid that does not contain a lower ligand. (B) Corrinoid metabolism categories include producers, dependents, and independents. Producers may release corrinoids (dashed line). These categories can be assigned based on experimental results, denoted with superscript E, and genomic analysis, denoted with superscript G, as summarized in the table. *In principle, producers[G] are defined based solely on the presence and completeness of the biosynthesis pathway, but all corrinoid producer[G] genomes examined thus far encode one or more corrinoid-dependent enzymes (23).

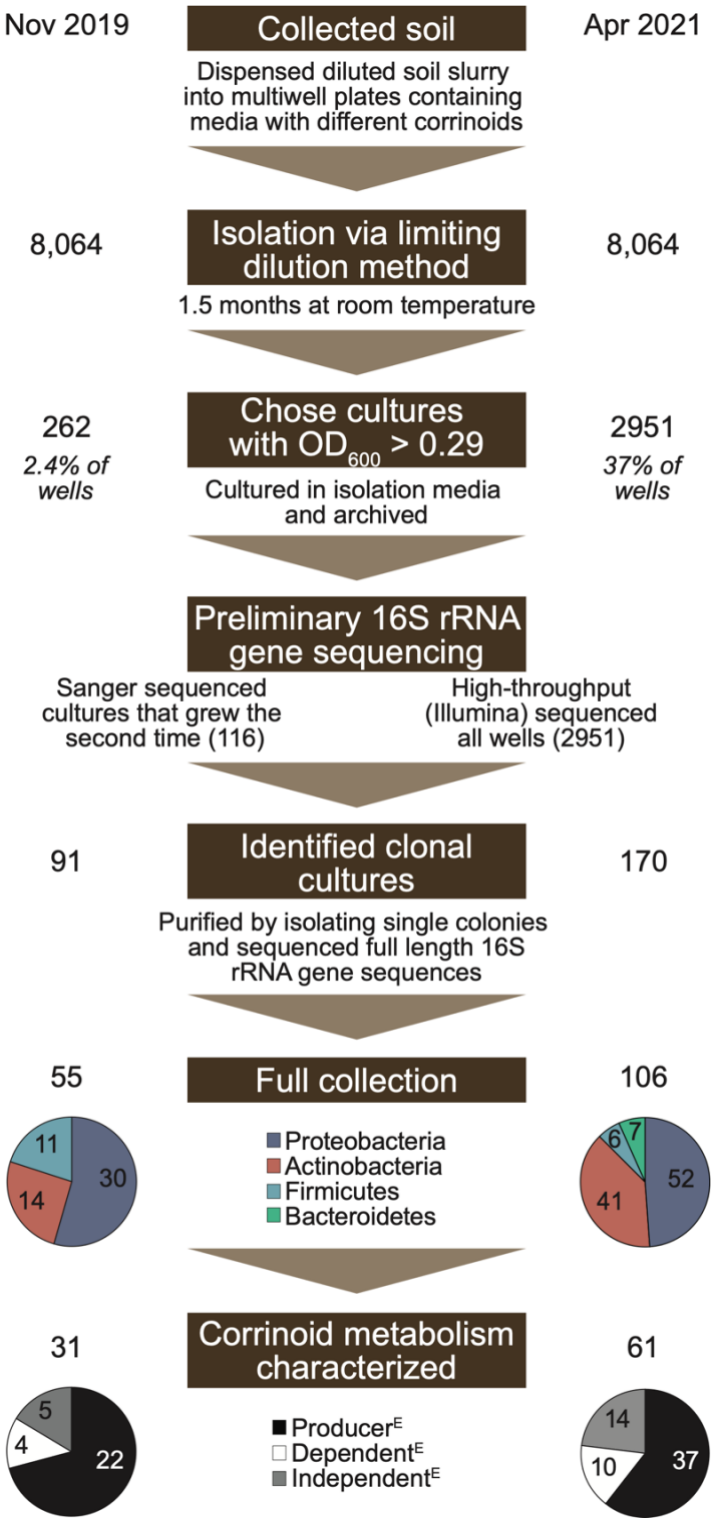


Figure 2.2. Overview of the experimental procedure. Soil bacteria were isolated by the limiting dilution method on media containing different corrinoids. Growth was detected based on OD_{600} and clonal isolates were distinguished from mixed cultures by 16S rRNA gene sequencing. After several purification steps, the final isolate collection contains 161 clonal isolates, of which 92 were successfully characterized as corrinoid producer[E], dependent[E], or independent[E].

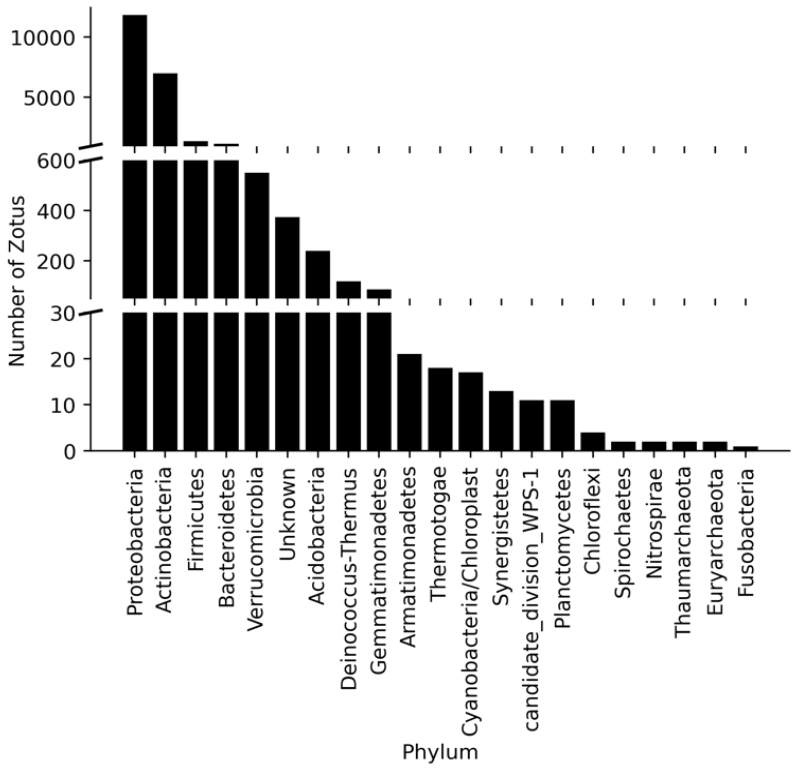


Figure 2.3. Phylum classification of microbes cultured by the limiting dilution method. Data are shown from Illumina sequencing of amplicons of the 16S V4-V5 region in the second isolation batch. 5.8% of these zOTUs were found to be in clonal wells, while the majority of wells contained two or more zOTUs.

The isolate collection is dominated by the phyla Proteobacteria and Actinobacteria, with fewer representatives from the Firmicutes and Bacteroidetes phyla (Fig. 2.4). This is similar to the relative abundances observed in bulk soil, where Proteobacteria and Actinobacteria are the dominant phyla (151–153). Of the 161 isolates, 23% (37 isolates) were considered to be novel species (154). The collection comprises 31 genera and 121 unique 16S rRNA gene sequences, with 11 genera each represented by a single isolate and three genera represented by 18 or more isolates. Despite this diversity, we have not sampled the bacterial diversity in this soil exhaustively (Fig. 2.5).

2.2.2 Taxonomic and phenotypic characterization of the isolate collection

To investigate whether there were phylogenetic trends for the observed phenotypes, we constructed a phylogenetic tree of the isolate collection annotated with the characteristics of each isolate (Fig. 2.4, Supplemental Table 4). We did not observe strong phylogenetic trends in the time required for each isolate to reach saturating growth, except that some clades of Proteobacteria contained only fast-growing isolates. Similarly, we did not observe a correlation between phylogeny and the corrinoid used for isolation. An exception was a clade of producers within Proteobacteria, all *Sphingomonas*, that were isolated on B12. The number of isolates recovered in B12, 5OH, and 2MA was higher than the number of isolates in the no corrinoid

(NOC) condition, while ADE, CBI, and CRE led to the recovery of fewer isolates than NOC (Fig. 2.4).

2.2.3 Classifying corrinoid metabolism phenotypes in the isolate collection

We experimentally classified each isolate as a corrinoid producer, dependent, or independent. We first assessed growth in the presence and absence of corrinoid. Isolates that stopped growing following serial transfer into media without corrinoid were classified as dependents[E]. Isolates that could grow in the absence of corrinoid were tested for corrinoid production using an *E. coli*-based corrinoid detection bioassay (20) to distinguish producers[E] from independents[E] (Fig. 2.1B; see Materials and Methods). Based on these results, all three categories are represented in the isolate collection, with the majority (64%) classified as producers[E], and 14 (15%) classified as dependents[E] (Fig. 2.4). The abundance of producers[E] in our collection contrasts with genome-based predictions that dependents outnumber producers across bacteria and specifically in soil (23, 112, 155).

To investigate potential phylogenetic trends in corrinoid metabolism categories, we overlaid the experimentally determined corrinoid phenotypes onto the phylogenetic tree (Fig. 2.4). At the phylum level, we observed mixed phenotypes. For example, all three categories are represented among characterized Actinobacteria and are interspersed across several clades. In contrast, in the Proteobacteria the corrinoid phenotypes are largely consistent with phylogeny. Most Proteobacteria clades are composed of only producer isolates, aside from one clade containing genera *Phenylobacterium* and *Caulobacter* that is composed of only dependents, while in a few other clades the phenotypes are interspersed. Thus, while trends are seen in some closely related isolates, large-scale phylogenetic trends in corrinoid phenotype are not apparent.

2.2.4 Corrinoid metabolism is conserved in a subset of genera, enabling taxonomy-based metabolic predictions

Although we did not observe strong phylogenetic trends for corrinoid metabolic categories in our isolate collection, we hypothesized that trends might exist across the bacterial domain at large. The existence of strong phylogenetic patterns for corrinoid metabolic categories would enable the prediction of corrinoid metabolism based solely on phylogeny, and having an isolate collection would allow us to confirm these predictions. We analyzed our previously published corrinoid metabolism classifications for over 11,000 bacterial species (23) to distinguish between the competing hypotheses that 1) corrinoid production, dependence, and independence show strong phylogenetic trends, enabling predictions of corrinoid metabolism based on taxonomy, or 2) corrinoid metabolism categories are phylogenetically interspersed, making it impossible to infer corrinoid-related ecological roles based solely on taxonomy. In our previous study, trends were not apparent at the phylum level except in the Bacteroidetes, which were nearly all dependents[G] (23). Therefore, we aimed to evaluate trends at lower taxonomic levels, starting with the genus level.

We searched for phylogenetic trends among the 85 genera in our dataset and classified each genus as producer[G], dependent[G] or independent[G] when possible. A corrinoid metabolism category could be assigned with high confidence for 47 out of 85 genera (Fig. 2.6, Supplemental

Table 5). In the remaining 38 genera, a single corrinoid metabolism category does not predominate, so corrinoid metabolism classifications could not be made.

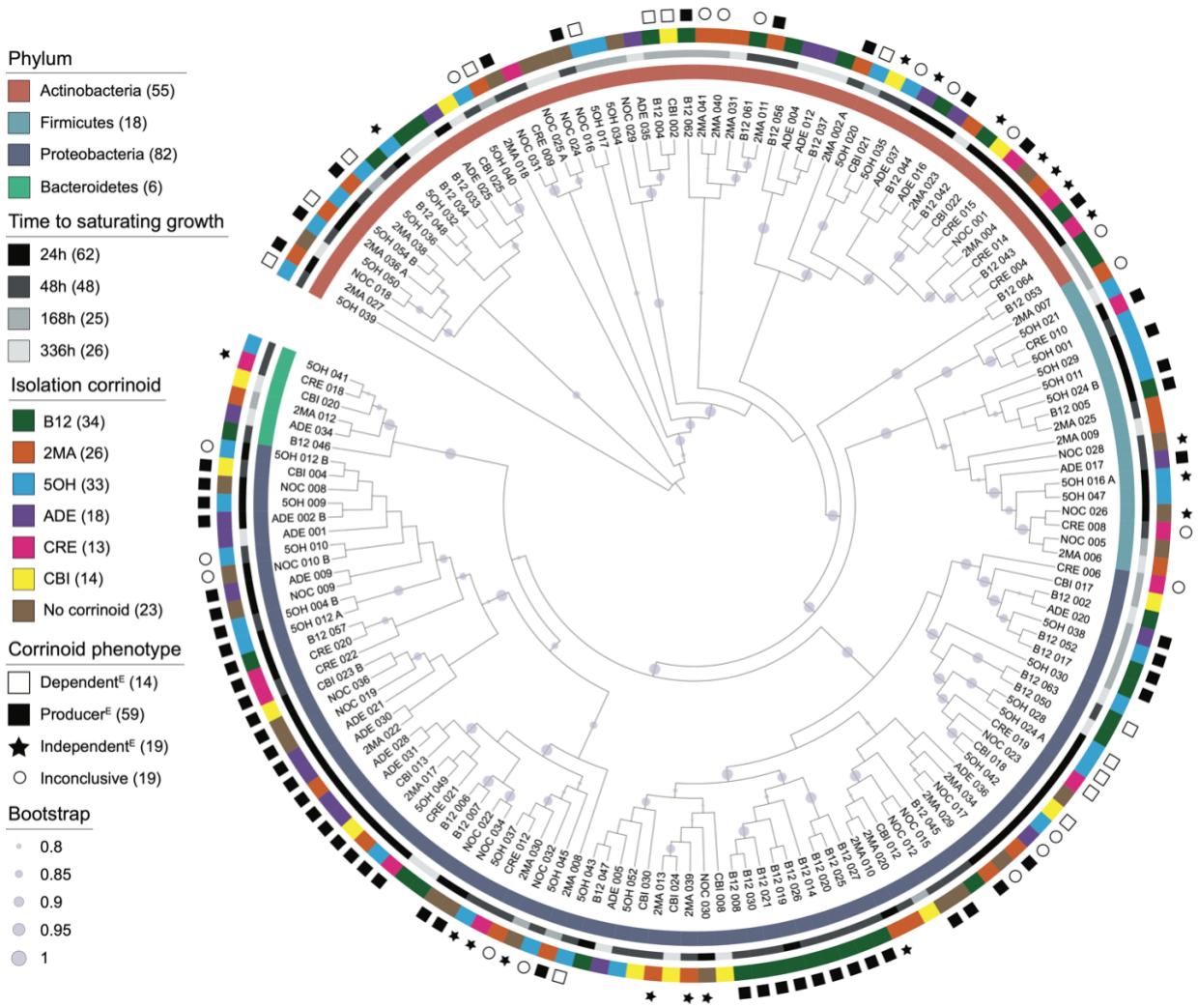


Figure 2.4. Overview of the isolate collection. A phylogenetic tree was built from the full-length 16S rRNA gene sequences of the isolates. The six- to seven-character ID is shown for each isolate. Light purple circles show bootstrap values above 0.8. Rings from the inside out show, for each isolate, (1) phylum as identified by SILVA taxonomy, (2) time to saturating growth determined from growth curves, (3) the corrinoid used in the isolation medium, and (4) the experimentally determined corrinoid metabolism category. Isolates with no symbol on the fourth ring were not tested due to slow or inconsistent growth. Numbers in parentheses correspond to the number of isolates belonging to each category.

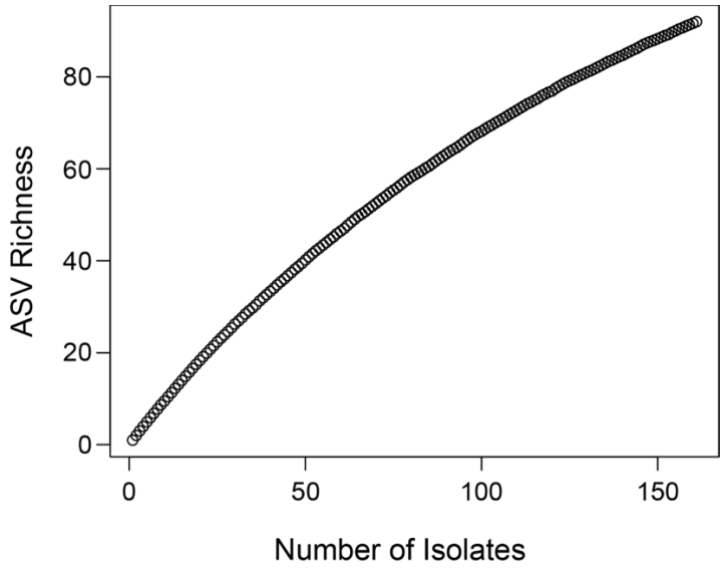


Figure 2.5. Collector's curve for the final isolate collection. The curve shows that our isolation effort was far from saturating the bacterial diversity of this soil.

To evaluate trends across higher taxonomic levels, we mapped the genomic predictions onto a phylogenetic tree constructed from the full 16S rRNA gene sequences of the type species for each genus (Fig. 2.6, Fig. 2.7). As expected, five of the six Bacteroidetes genera were predicted to be dependent[G] and the sixth Bacteroidetes genus (*Porphyromonas*) has a small percentage of independent[G] species. We observed phylogenetic trends at levels higher than the genus level in some cases. The Actinobacteria form two distinct clades, one dominated by producers[G] (clade I in Fig. 2.6) and the other by independents[G] (clade II). No Actinobacteria genera were classified as dependents[G], although some genera have low percentages of dependent species. All genera in one Proteobacteria clade are classified as producers[G] (clade III), with the notable exception of *Bartonella*, which has undergone genome reduction (156) and is classified as independent[G]. However, other Proteobacteria clades were mixed, aside from a few sister taxa that share corrinoid genotypes in some instances, such as *Bradyrhizobium and Methylobacterium* which are both producers[G] and *Xanthomonas* and *Lysobacter*, which are both dependents[G]. For phyla that had fewer genera, the few classified genera were independent[G]. This may be due to a bias in the dataset, which is composed of over 90% cultured bacteria. Because cultured bacteria tend to have larger genomes and fewer auxotrophies than uncultured bacteria (139), they are less likely to be corrinoid-dependent. Future analysis of metagenomes may reveal more dependence among phyla with fewer sequenced representatives.

Upon comparing the genomic classifications to our experimental results for the isolate collection, we found that 19 isolates belong to genera for which genomic classifications were possible. All isolates except one matched the genomic classification (Fig. 2.6). The exception was a *Mesorhizobium* isolate that was predicted to be a producer[G] but found to be independent[E], suggesting either that it is not capable of synthesizing a corrinoid or did not produce a detectable amount under our growth conditions. The confirmation of our genomic classifications with experimental data from our isolate collection lends support to the phylogenetic predictions made for certain genera.

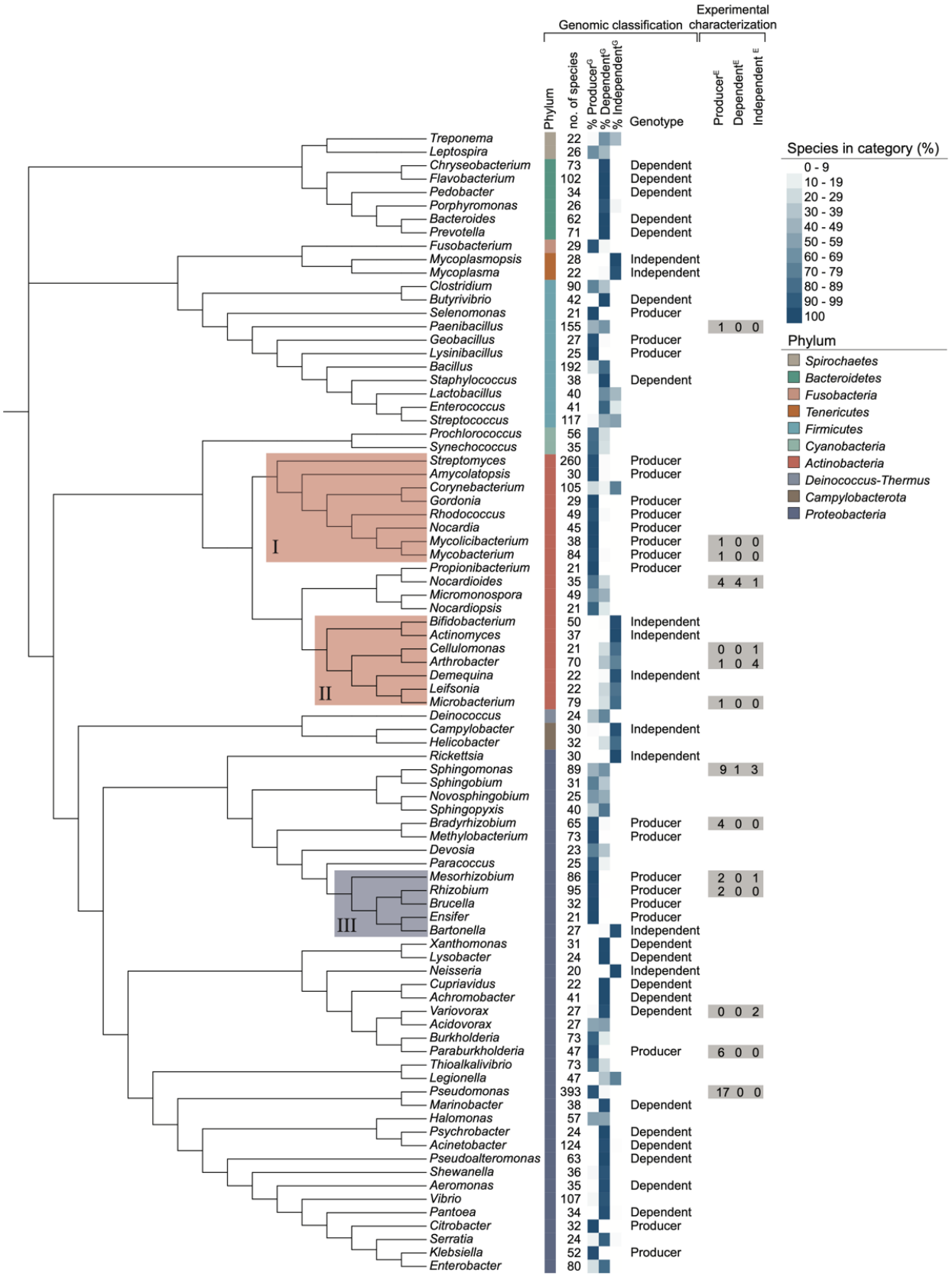

Experimental characterization
Genomic classification
Phylum
no. of species
% Producer^G
% Dependent^G
% Independent^G
Genotype
Producer^E
Dependent^E
Independent^E
Treponema 22
Leptospira 26
Chryseobacterium 73 Dependent
Flavobacterium 102 Dependent
Pedobacter 34 Dependent
Porphyromonas 26
Bacteroides 62 Dependent
Prevotella 71 Dependent
Fusobacterium 29
Mycoplasmopsis 28 Independent
Mycoplasma 22 Independent
Clostridium 90
Butyrivibrio 42 Dependent
Selenomonas 21 Producer
Paenibacillus 155 1 0 0
Geobacillus 27 Producer
Lysinibacillus 25 Producer
Bacillus 192
Staphylococcus 38 Dependent
Lactobacillus 40
Enterococcus 41
Streptococcus 117
Prochlorococcus 56
Synechococcus 35
Streptomyces 260 Producer
Amycolatopsis 30 Producer
Corynebacterium 105
Gordonia 29 Producer
Rhodococcus 49 Producer
Nocardia 45 Producer
Mycolicibacterium 38 Producer 1 0 0
Mycobacterium 84 Producer 1 0 0
Propionibacterium 21 Producer
Nocardioides 35 4 4 1
Micromonospora 49
Nocardiopsis 21
Bifidobacterium 50 Independent
Actinomyces 37 Independent
Cellulomonas 21 0 0 1
Arthrobacter 70 1 0 4
Demequina 22 Independent
Leifsonia 22
Microbacterium 79 1 0 0
Deinococcus 24
Campylobacter 30 Independent
Helicobacter 32
Rickettsia 30 Independent
Sphingomonas 89 9 1 3
Sphingobium 31
Novosphingobium 25
Sphingopyxis 40
Bradyrhizobium 65 Producer 4 0 0
Methylobacterium 73 Producer
Devosia 23
Paracoccus 25
Mesorhizobium 86 Producer 2 0 1
Rhizobium 95 Producer 2 0 0
Brucella 32 Producer
Ensifer 21 Producer
Bartonella 27 Independent
Xanthomonas 31 Dependent
Lysobacter 24 Dependent
Neisseria 20 Independent
Cupriavidus 22 Dependent
Achromobacter 41 Dependent
Variovorax 27 Dependent 0 0 2
Acidovorax 27
Burkholderia 73
Paraburkholderia 47 Producer 6 0 0
Thioalkalivibrio 73
Legionella 47
Pseudomonas 393 17 0 0
Marinobacter 38 Dependent
Halomonas 57
Psychrobacter 24 Dependent
Acinetobacter 124 Dependent
Pseudoalteromonas 63 Dependent
Shewanella 36
Aeromonas 35 Dependent
Vibrio 107
Pantoea 34 Dependent
Citrobacter 32 Producer
Serratia 24
Klebsiella 52 Producer
Enterobacter 80
I
II
III
Species in category (%)
0 - 9
10 - 19
20 - 29
30 - 39
40 - 49
50 - 59
60 - 69
70 - 79
80 - 89
90 - 99
100
Phylum
Spirochaetes
Bacteroidetes
Fusobacteria
Tenericutes
Firmicutes
Cyanobacteria
Actinobacteria
Deinococcus-Thermus
Campylobacterota
Proteobacteria

Figure 2.6. Genome-based predictions of corrinoid metabolism at the genus level. The phylogenetic tree was built from the full-length 16S rRNA gene sequences of the type species of 85 genera from the dataset in Reference (23) that met our cutoff by having 20 species or more. The first two columns show the phylum and the number of species analyzed for each genus, respectively, which total 10 phyla and 4,720 species. The next three columns show the percent of species in each genus predicted to belong to each corrinoid metabolism category. A corrinoid-specific genotype is indicated if 95% or more species in a genus belong to the same category. The columns labeled Experimental characterization show the number of isolates in the collection found to belong to each category based on experimental results. The unpruned tree that was used to generate the figure is shown in Fig. 2.7 and sequences obtained from the NCBI Reference Sequence database can be found in Supplemental Table 3.

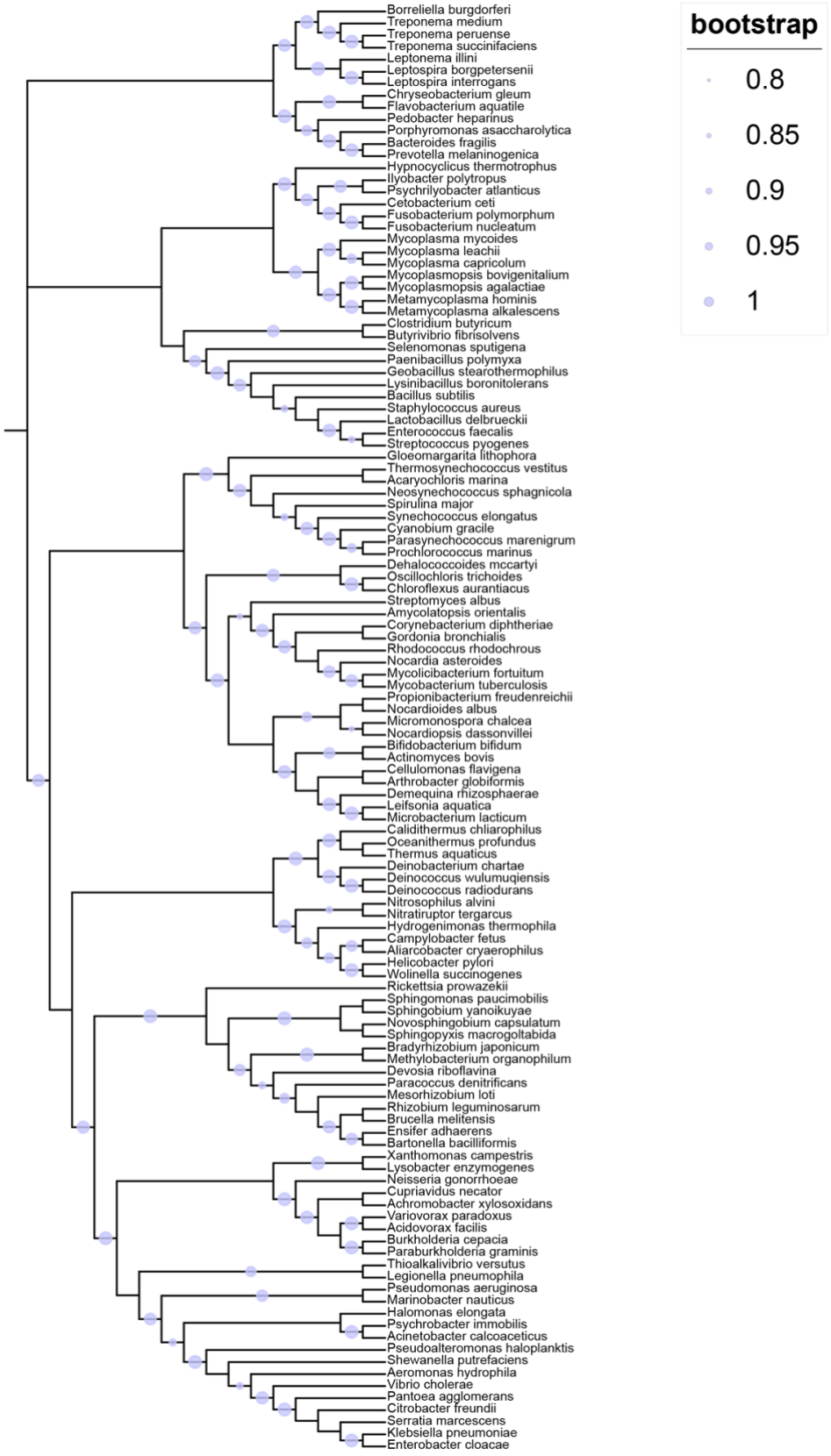


Figure 2.7. Phylogenetic tree of all type species used to generate the pruned tree in figure 2.6.

2.2.5 Corrinoid preferences of dependent isolates reveal diverse corrinoid use capabilities

Our observation that the corrinoid used for isolation did not correlate with phylogeny (Fig. 2.4) led us to investigate the corrinoid preferences of the 14 dependent isolates in our collection. We measured growth in media containing a range of concentrations of different corrinoids and calculated the corrinoid concentration that resulted in half-maximal growth (EC_{50}); the corrinoid with the lowest EC_{50} is considered the most preferred (Fig. 2.8A and Fig. 2.9). The corrinoid used for isolation was not the most preferred corrinoid in many cases, likely because the corrinoid concentration in the isolation medium was in 100- to 1,000-fold excess of isolate requirements (Fig. 2.8B). Further, despite our previous finding that most corrinoids other than B12 have not been detected in this soil (19), we found that all of the isolates can use at least one corrinoid in addition to B12, with B12 preferred by almost all isolates. ADE, CRE and CBI could not be used at any of the tested concentrations by some isolates. These were the same corrinoids in which we recovered the lowest numbers of isolates, suggesting that only some isolates can use the complete corrinoids ADE and CRE as cofactors or salvage CBI to make a complete corrinoid under our experimental conditions (62–64).

2.2.6 B12 is the main corrinoid produced by isolates in the collection and it is only provided by a subset of producers

Given that dependents need to obtain corrinoids from producers in their community, we sought to determine whether there is compatibility between the corrinoids produced and required by isolates in our collection. To that end, we extracted corrinoids from cultures of 12 fast-growing producers and analyzed them by HPLC. We detected a corrinoid by HPLC in 11 of these producers. The other isolate showed no corrinoid signal when re-tested with the *E. coli* bioassay and was reclassified as inconclusive. Comparison with authentic standards revealed that B12 was the dominant corrinoid synthesized by the 11 tested producers (Fig. 2.10A).

Although it is unknown to what extent dependents acquire corrinoids directly from producer cells, a recent report showed that some producers cultured in the laboratory can release corrinoids into the growth medium (100). We used the *E. coli*-based bioassay to quantify the corrinoids in culture supernatants and cell pellets of the 11 producer isolates. Seven of the producers were found to be “providers” (100) – producers for which corrinoids were present in the culture supernatant – while corrinoids were detected exclusively in the cell pellet fraction in the remaining four producers (Fig. 2.10B). The amount of provided corrinoid ranged from 1.3 to 21% of the total corrinoid, and the provided amount did not correlate with the total amount of corrinoid produced. The concentrations of provided corrinoid are 1 to 1,000 times higher than the EC_{50} values calculated for the dependents (Fig. 2.8B, Supplemental Table 6), suggesting that these isolates have the capacity to provide sufficient or excess corrinoid to all of the dependents in our collection. Thus, our measurements of corrinoid production and providing, in the artificial conditions of laboratory culture, coupled with prior genomic studies (23, 112), support our hypothesis that corrinoid sharing can occur within the communities of this soil.

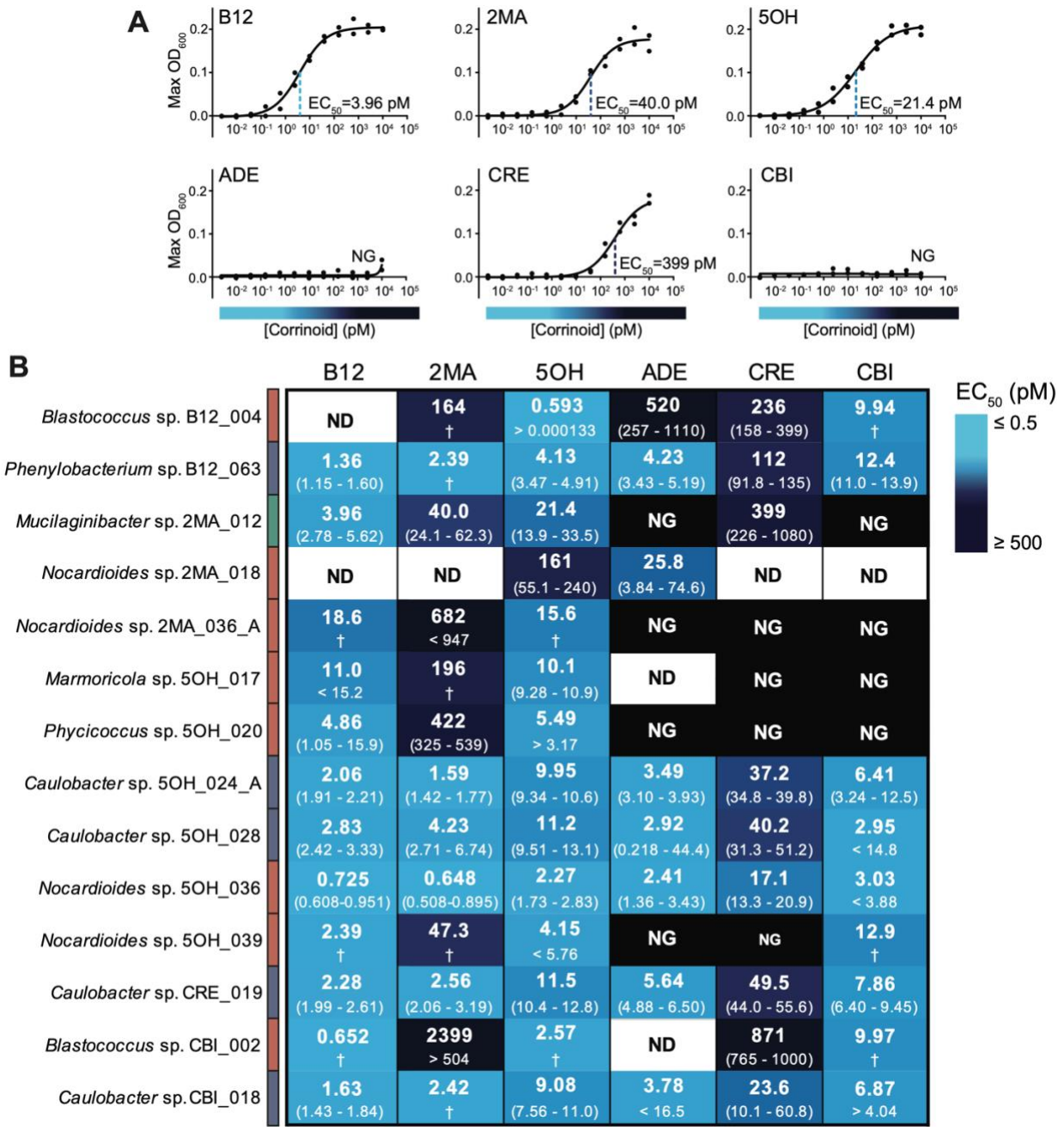


Figure 2.8. Corrinoid dependence in the isolate collection. (A) Representative dose-response curves belonging to isolate 2MA_012. Lines show non-linear fit for each corrinoid (n=2). Dotted line marks the EC_{50} for each corrinoid and the EC_{50} is shown when one was obtained. (B) The corrinoid concentrations resulting in half-maximal growth (EC_{50}) are shown for all 14 corrinoid-dependent isolates on the six corrinoids used in this study. For each isolate and corrinoid combination, the top number is the EC_{50} and the numbers in parentheses represent the 95% confidence interval as calculated by a four-parameter non-linear fit on GraphPad Prism (v9.5.1). Greater than and less than symbols were used when the upper or lower bound of the confidence interval could not be determined, respectively. The phylum each isolate belongs to is shown in the bar next to the isolate names, colors correspond to the legend in Figures 2.2, 2.4, and 2.6. NG: No Growth, ND: No Data when a regression was not possible due to poor data quality, † 95% confidence interval not determined. Curve fit data are summarized in supplemental table 7.

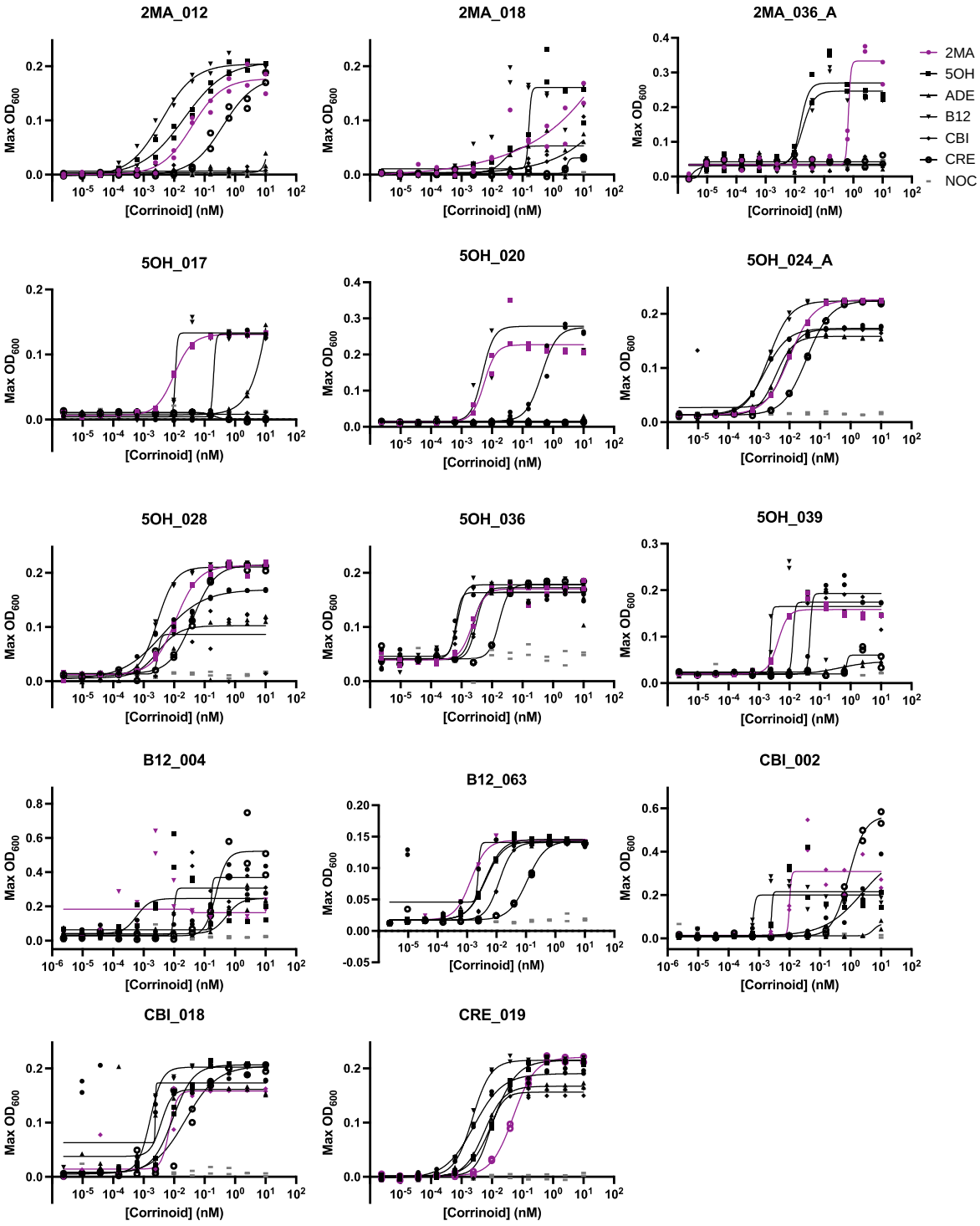


Figure 2.9. Dose-response curves for all tested dependents reveal widespread preferences for B12. The corrinoid used for isolation is shown in purple, and the no corrinoid condition is shown in gray. EC_{50} values calculated from these curves are shown in Figure 2.8B.

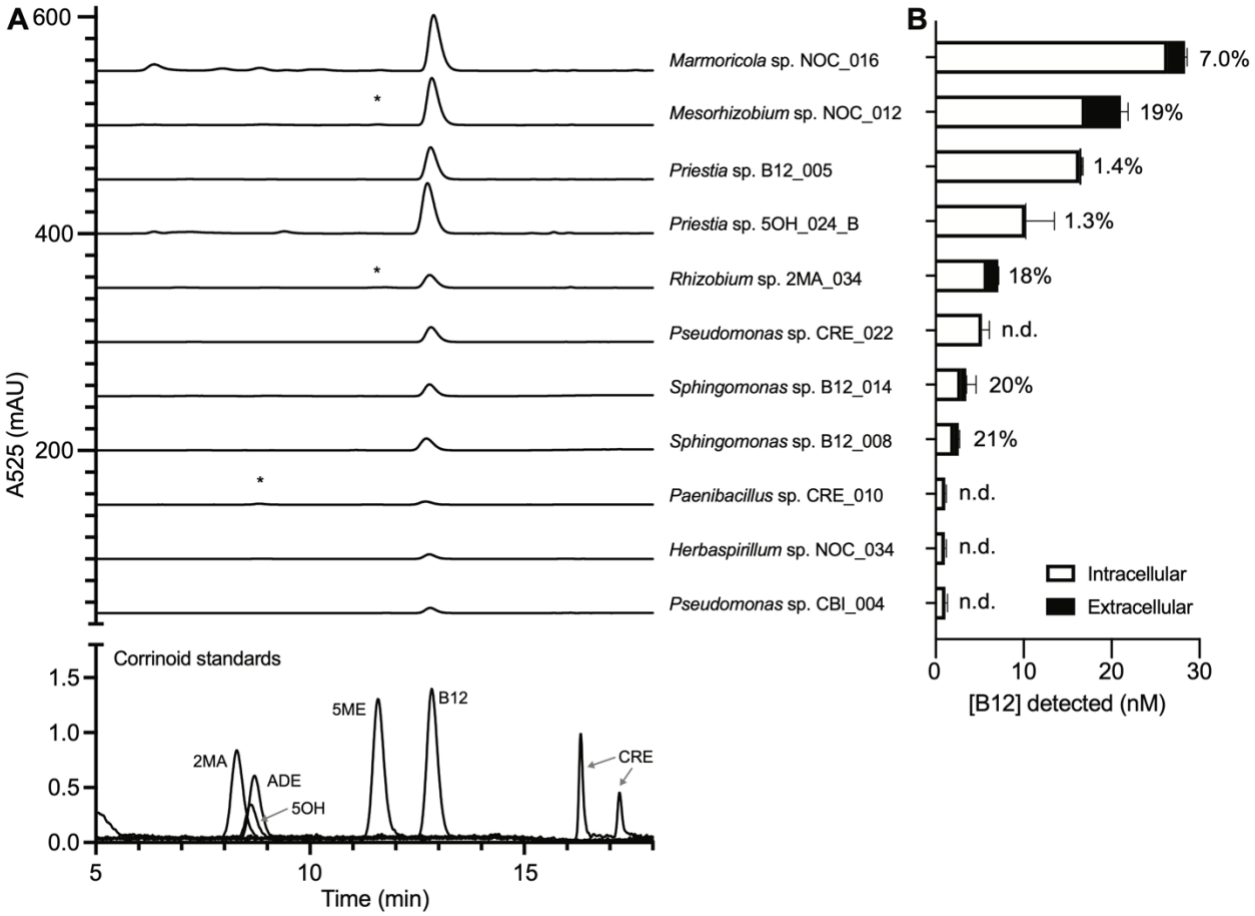


Figure 2.10. Corrinoid production and providing in the isolate collection. (A) HPLC analysis of corrinoid extracts of 11 selected producers shows B12 is the major corrinoid produced. Authentic corrinoid standards are shown at the bottom. Asterisks denote small peaks that indicate the presence of a second complete corrinoid. (B) Quantification of corrinoids in the cell pellet (intracellular) and supernatant (extracellular) fractions of each isolate as detected by an *E. coli*-based corrinoid bioassay. The percent of corrinoid provided (extracellular corrinoid as a fraction of the total corrinoid) is given to the right of each bar. Bars and error bars show the average and standard deviation of three technical replicates, respectively. n.d., extracellular corrinoid was not detected.

2.3 Discussion

Microbial nutritional interactions play pivotal roles in establishing community structure and function. Characterizing and predicting the ecological roles of microorganisms as nutrient producers and dependents can contribute to the understanding of microbial interaction networks and their influence on the whole community. Here, we investigated the ecological roles of microorganisms by overlaying experimental and computational approaches. We were able to characterize specific functional roles of bacteria by focusing on a single class of model nutrients, corrinoids, the sharing of which is thought to be widespread in microbial communities (7, 22, 23, 126). The importance of corrinoids for soil bacteria has long been recognized (155, 157–159). Here, we report the first systematic isolation and characterization of soil bacteria on corrinoids

other than B12, allowing us to consider the ecological roles of corrinoid producers and dependents in the context of soil microbial ecology.

Microorganisms typically have preferences for different corrinoids that are reflected in their EC_{50} values (7, 36, 90). These preferences result from corrinoid transport efficiency, the affinity of corrinoids for the enzymes that use them, and what corrinoid dependent processes are in use (36, 68, 160). Corrinoid preferences, combined with the availability of corrinoids in a given environment and competition for corrinoids in the community, can impact microbial fitness (7). After experimentally determining the preferences of the corrinoid-dependent bacteria in our collection, we found that our isolates have considerably lower EC_{50} values for their preferred corrinoids than bacteria for which EC_{50} measurements have previously been reported, indicating lower corrinoid concentrations are required for growth (20, 40, 68, 161). The EC_{50} values of our isolates are comparable to those of aquatic algae (90, 162), some of which live in environments with corrinoid concentrations in the picomolar range (134) (Fig. 2.8B and Table 2.1). The ability to use corrinoids at low concentrations could be a useful adaptation to the soil environment where corrinoids may be limiting due to the physical heterogeneity of soil microbial communities, long distances between cells, and fluctuations in water content throughout the year, which make nutrient availability highly variable (144, 163, 164).

All of the dependent isolates were able to use corrinoids that have not been detected in this soil (19). The concentration of corrinoid chosen for the isolation media was four-fold higher than the highest EC_{50} and over 16,000-fold higher than the lowest EC_{50} we measured, which explains why isolates were often recovered in their non-preferred corrinoid and why we detected no taxonomic trends in the corrinoid used for isolation. Despite the presence of excess corrinoid in our isolation media, we recovered fewer corrinoid-dependent isolates than expected (112, 155), which may be due to corrinoid-dependent bacteria having additional nutrient dependencies not satisfied by our isolation medium or requiring specific partners for their survival. Indeed, in an analysis of auxotrophies in gut microbiome genomes, most predicted B12 auxotrophs had at least one other vitamin auxotrophy, (140) and among a set of marine isolates 10/13 B12 auxotrophs were also auxotrophic for at least one other B vitamin (165).

When considering how our classifications of corrinoid metabolism fit into the context of soil microbial ecology, we must consider functional diversity (166). A contemporary question regarding microbiome function relates to whether groups of microorganisms with shared functions are composed of phylogenetically close organisms or unrelated organisms that share similar metabolic capabilities. Traits such as photosynthesis, methanogenesis, maximum growth rates, and response to soil wet-up tend to be strongly correlated with phylogeny, while others, such as use of specific carbon sources, have weak or no phylogenetic signals (167–169). Here, we found some phylogenetic trends in corrinoid traits, but overall, the distribution of these traits is patchy across the phylogenetic tree, suggesting that gene loss has occurred at various evolutionary points, possibly due to the frequent emergence of corrinoid dependence and independence (135), or that horizontal gene transfer (HGT) is important in sustaining corrinoid biosynthesis and use. Indeed, corrinoid uptake genes in human gut Bacteroidetes are commonly found on mobile genetic elements (127), and *Salmonella typhimurium* and *Lactobacillus reuteri* biosynthesis genes are thought to have been acquired by HGT (11, 170). The evolutionary history of corrinoids should be explored further to identify which processes have impacted the

biosynthesis and use of these cofactors. Using our isolate collection, we were able to carry out the crucial step of validating some genus-level predictions because seven of the genera for which we predicted a genotype were represented among the isolates (Fig. 2.6).

Table 2.1. Known EC_{50} values for bacteria and aquatic eukaryotic algae.

Domain	Organism	B12 EC_{50}	Reference
Bacterial	*Akkermansia muciniphila*	56.1 pM	Mok 2020
Bacterial	*Escherichia coli* Δ*metE*	~0.1 nM	Mok 2022
Bacterial	*Bacteroides thetaiotamicron*	< 0.4 nM	Degnan 2014 Cell Host Microbe
Bacterial	*Clostridium difficile*	~1 nM	Shelton 2020 J Bacteriol
Eukaryotic	B_{12} dependent mutant of *Chlamydomonas reinhardtii*	28 pM	Helliwell 2016 Curr Bio
Eukaryotic	*Karenia mikimotoi*	13.1 pM	Tang 2010 PNAS
Eukaryotic	*Aureococcus anophagefferens*	3.49 pM	Tang 2010 PNAS
Eukaryotic	*Rhodomonas salina*	0.36 pM	Tang 2010 PNAS
Eukaryotic	*Fibrocapsa japonica*	0.28 pM	Tang 2010 PNAS
Eukaryotic	*Chattonella marina*	0.19 pM	Tang 2010 PNAS
Eukaryotic	*Prorocentrum minimum*	0.02 pM	Tang 2010 PNAS
Eukaryotic	*Pavlova lutheri*	~ 18 pM	Helliwell 2016 Curr Bio
Eukaryotic	*Ostreococcus tauri*	< 70 pM	Helliwell 2016 Curr Bio
Eukaryotic	*Amphidinium carterae*	< 70 pM	Helliwell 2016 Curr Bio
Eukaryotic	*Thalassiosira pseudonana*	< 70 pM	Helliwell 2016 Curr Bio
Eukaryotic	*Aureococcus anophagefferens*	< 70 pM	Helliwell 2016 Curr Bio
Eukaryotic	*Lobomonas rostrata*	< 70 pM	Helliwell 2016 Curr Bio
Eukaryotic	*Euglena gracilis*	< 70 pM	Helliwell 2016 Curr Bio

The characterization of this isolate collection provides key insights about corrinoid-based microbial interactions in soil. We found that dependent isolates were all able to use B12, with most preferring it, and the producers we characterized all synthesize B12 under laboratory conditions, indicating compatibility between corrinoid production and preferences of the dependents. Because oxygen availability is highly variable in soil, and corrinoids other than B12 and ADE are found primarily in anoxic environments, it is possible that these producers synthesize other corrinoids under anoxic conditions. However, the observation that B12 is the major corrinoid found in this soil (19) suggests that most corrinoid producers synthesize B12 in their native environment. Among the producers, however, only some release corrinoid into culture supernatants, suggesting that the corrinoid provider role is fulfilled by a distinct subset of producers (100). Based on our observation that these providers release corrinoids at levels sufficient to support many dependents in laboratory cultures, we speculate that a small fraction of

the community disproportionately provides corrinoids to the dependents. In a similar vein, we previously found that amino acid auxotrophs can be supported by producers at a ratio of over 40:1 (122). Because corrinoid release cannot be predicted from genomes, it is necessary to combine genotypic predictions of corrinoid production with phenotypic characterizations when studying interactions. This collection of isolates assembled from the same study site will enable further investigation of corrinoid-based interactions via culture-based studies. For example, the mechanisms of corrinoid release, partner specificity in interactions, and competition for corrinoids among dependent isolates can be explored. Focusing on corrinoids simplifies community interactions to only one nutrient and does not take into account other possible interactions that are prevalent in the soil environment, including those involving other shared nutrients (171, 172) or cross-domain interactions (153), which may be affected by environmental fluctuations in the native environment (163). Nonetheless, focusing on this important class of shared nutrients enabled us to study the diversity of metabolic capabilities that may be prototypical of interactions among soil bacteria and provides a framework to expand the study of other nutrient-sharing interactions.

2.4 Methods

2.4.1 Isolation of bacteria from soil by the limiting dilution method

We collected soil samples from top 10 cm of an annual grassland at the Hopland Research and Extension Center in Hopland, CA (39.004056 N 123.085861 W) in November 2019 and again in April 2021 to increase the total number of isolates. Site characteristics and soil physicochemical properties for our sampling site were previously documented (151, 173). Soil pH, determined by preparing a slurry with 1 part soil to 2 parts deionized water and measuring (n=3) with an Orion Star A111 pH Meter (Thermo Scientific, Waltham, MA, USA), was 5.87 ± 0.42 in November 2019 and 6.28 ± 0.07 in April 2021.

To separate microbial biomass from the soil, we resuspended 2.5 g of soil in 25 ml phosphate buffered saline with 2.24 mM sodium pyrophosphate (Alfa Aesar, Heysham, England), stirred for 30 minutes, and allowed the slurry to settle for 15 minutes before diluting the supernatant. All isolations and subsequent growth steps were performed using a modified VL60 medium (pH 6.0) (174) (Supplemental Table 1), amended with 0.1 g/L each of xylose, xylan, N-acetylglucosamine, and glucose, as well as 10 nM B_{12} (B12), 2-methyladeninylcobamide (2MA), 5-hydroxybenzimidazolylcobamide (5OH), adeninylcobamide (ADE), *p*-cresolylcobamide (CRE), or cobinamide (CBI) when indicated. Cultures were grown at room temperature unless otherwise noted. The corrinoids 2MA, 5OH, and ADE were produced by guided biosynthesis in *Propionibacterium acidi-propionici* and CRE in *Sporomusa ovata*; extracted from bacterial cultures; and purified as previously described (20, 118). We conducted a most probable number (MPN) count to determine the soil slurry dilution required to reach growth in approximately 30% of wells, a density that is expected to yield 80% clonal cultures, based on a Poisson distribution and previously reported isolations from human stool samples (150). 80 µl aliquots of the soil solution diluted in each medium were dispensed into the wells of 384-well plates using a Biomek liquid handler (Beckman Coulter, Indianapolis, IN, USA). Three plates per corrinoid condition were inoculated, and an uninoculated plate was prepared for each condition, for a total of 28 plates. Plates were covered with BreatheEasy (Diversified Biotech, Dedham, MA, USA)

membranes for this and all subsequent steps, and incubated statically at room temperature for 44 days. Despite using the MPN calculation to determine the dilution, a surprisingly low number of wells showed growth in the isolation from soil collected in November 2019 (2.4%). Cultures from wells in which the OD_{600} (measured on a Tecan Spark plate reader (Grödig, Austria)) exceeded 0.29 were transferred into fresh medium at the end of the initial incubation and grown for up to 40 days. The cultures were then split into two portions, one stored at -80ºC in 25% glycerol and another prepared for sequencing. The full 16S rRNA gene was amplified by PCR from each well with primers 27F and 1492R (175) (IDT, Coralville, Iowa, USA) and DreamTaq polymerase (Thermo Scientific). PCR purification and Sanger sequencing of all amplicons using the same primers was done at the UC Berkeley DNA Sequencing Facility. Sanger sequence trimming with a 0.01 error probability cutoff and *de novo* assembly of reads were performed on Geneious Prime (2022.1.1). Cultures with a single, high-quality 16S rRNA gene sequence were considered clonal.

Prior to the second isolation from soil collected in April 2021, the soil sample was stored at 4ºC for one month, brought to 20% moisture from an original 3.33 ± 0.58% with sterile deionized water, and incubated for one week. After diluting the soil and dispensing into 96-well plates, the plates were incubated at room temperature for 49 days. The percentage of wells showing growth was much higher than in the previous isolation round (37% of total wells). Therefore, high-throughput sequencing of the 16S rRNA gene V4/V5 region was performed as previously described (176) with 600 bp v3 reagents. This step allowed us to select wells containing a single amplicon sequence variant, which were assumed to be clonal cultures. To determine the total phyla represented across all wells, 16S rRNA gene sequences were searched against the RDP database (release 8.0) (177) (Fig. 2.3).

Liquid cultures prepared from glycerol stocks were purified by streaking on 2X modified VL60 solidified with 14 g/L Difco noble agar (BD, Sparks, MD, USA). Nystatin (63 ng/ml) was added to the medium in cases where fungal contamination was observed. Cultures were serially purified by streaking until all observed colonies were of uniform morphology. For each isolate, liquid cultures inoculated from a single colony were stored at -80ºC in 40% glycerol. After purifying, the identity of each isolate was confirmed by a second round of Sanger sequencing.

We identified 23 sequences with higher than 99% pairwise identity to a sequence in an uninoculated well. These were considered potential contaminants and removed from the dataset. After removal of isolates with chimeric sequences, the final collection is composed of 161 isolates.

Growth curves were generated to classify isolates into groups based on the time they required to reach saturating growth (24, 48, 168, or 336 hours). Isolates were inoculated from glycerol stocks into 96-well plates in triplicate and grown for 168 hours at 28ºC, shaking at 800 rpm in a plate shaker (Southwest Science, Roebling, NJ, USA), and separately at room temperature with no shaking. Growth curves were generated by measuring OD_{600} at 0, 6, 12, 24, 36, and 48 hours, and every 24 hours until 168 hours for the shaken cultures and 216 hours for standing cultures. Because isolates grew more consistently in the shaking condition, cultures were shaken at 28ºC for all subsequent steps.

2.4.2 Experimental characterization of isolates as corrinoid producers, dependents, or independents

To determine whether isolates were dependent on corrinoids for growth, isolates in the 24-, 48-, and 168-hour groups were inoculated into 96-well plates with 200 µl of media containing the corrinoid used for isolation. Following growth to saturation, each culture was diluted into two wells, one amended with the same corrinoid and the other with no corrinoid, using a multi-blot replicator that transferred approximately 3 µl per well (V&P Scientific, San Diego, CA, USA). Cultures were serially passaged three additional times into the same media to eliminate corrinoid carryover. OD_{600} was measured before and after each passage. Isolates that did not grow reproducibly in media with corrinoid were not pursued further (24 isolates). Isolates that continued to grow in media with corrinoid but stopped growing after being transferred into media with no corrinoid were classified as dependents^{E}, while those that continued to grow in both conditions were considered to be either producers^{E} or independents^{E} (Fig. 2.1B). Superscript E is used to distinguish experimental results from genomic predictions (superscript G, discussed below). To evaluate the effect of corrinoids on the growth of isolates, we calculated the corrinoid-specific growth enhancement as $\log_2 [1+((OD_{\text{with corrinoid}} - OD_{\text{no corrinoid}})/(OD_{\text{no corrinoid}}))]$ (178) and determined a threshold for corrinoid dependence based on the growth of bacteria isolated in the no corrinoid condition that also underwent serial transfer (maximum value obtained from the equation plus standard deviation). If two or three of the three replicates were classified as corrinoid dependent, corrinoid dose-response assays were performed to confirm dependence and determine the corrinoid preferences of each isolate. For dose-response curves, isolates were pre-cultured in media with 10nM corrinoid, followed by a second pre-culture step in media with no corrinoid, and inoculated into 12 different concentrations of each corrinoid (ranging from 10 nM to 2.4 fM). Final ODs were measured and curves were fitted using a four-parameter non-linear fit on GraphPad Prism (v9.5.1) to determine each EC_{50}.

To distinguish producers^{E} from independents^{E} (Fig. 2.1B), 100 µl of each culture were collected at the end of the fourth passage with no corrinoid addition and lysed by incubating at 98ºC for 20 minutes. An *E. coli*-based corrinoid detection bioassay was conducted as previously described (20) to determine the presence or absence of corrinoid in each sample. Data were processed to yield a "growth due to corrinoid" metric by subtracting growth due to methionine (as measured by the Δ*metE*Δ*metH* control strain) from growth of the Δ*metE* bioassay strain and normalizing to growth of the wildtype *E. coli* strain. An isolate was characterized as a producer^{E} if the normalized result was greater than or equal to 2 or if the OD_{600} of the Δ*metE* bioassay strain was greater than or equal to 0.1. Conversely, an isolate was characterized as a non-producer, and thus an independent^{E}, if the normalized result was less than 2 and the *E. coli* Δ*metE* OD_{600} was less than 0.1. Our method was validated using a set of previously isolated soil bacteria (179) that were genomically predicted to be corrinoid producers (Supplemental Table 2). Isolates that repressed growth of *E. coli* (2 isolates), grew to an OD_{600} less than 0.1 (7 isolates), or for which the three replicates or results for the dependence and production were inconsistent (11 isolates) were deemed inconclusive.

Data processing and analysis were performed using Python 3.7 on Jupyter Notebooks (version 6.2.0) and all code is on GitHub under DOI: 10.5281/zenodo.10815172.

For further characterization of producers, we grew 1 L cultures of each in VL60 medium with no corrinoid and 200X amino acids and extracted corrinoids as described previously (20). Corrinoid extracts were analyzed by high-pressure liquid chromatography (HPLC) on a 1200 series HPLC system equipped with a diode array detector (Agilent Technologies, CA, USA) and compared to authentic corrinoid standards using Method 2, as described previously (35).

2.4.3 Genus-based predictions of corrinoid metabolism

We used a previously developed dataset (23), which reports predictions of corrinoid biosynthesis and dependence for 11,436 bacterial species, and used these existing predictions to further genomically characterize bacterial species into the producer[G], dependent[G], and independent[G] categories. Species that were previously classified as very likely, likely, or possible producers were considered producers[G] (23). Corrinoid-dependent[G] species were defined as those previously classified as very likely or likely non-producers that also had at least one corrinoid-dependent function, regardless of whether their genomes encoded specific corrinoid-independent alternative enzymes (23). Corrinoid-independent[G] species were defined as those that were likely or very likely non-producers and had no corrinoid-dependent functions. After classifying each species, we grouped all species into their respective genera (JGI IMG taxonomic metadata was downloaded on July 18, 2023, to update any reclassified genomes). To establish a reliable cutoff for our predictions, we chose genera containing 20 species or more and made a genomic classification when 95% or more of the species in a genus corresponded to the same category.

2.4.4 Phylogenetic tree building

The phylogenetic tree of the isolates was constructed from full-length 16S rRNA gene sequences (Fig. 2.4). Isolate taxonomy assignment and tree building were done using the Silva Alignment, Classification, and Tree (ACT) service (180). To determine whether isolates were likely novel, we used BLAST to search assembled sequences against the NCBI Reference Database using Geneious (2022.1.1). If the pairwise identity between the isolate sequence and the top hit was lower than 98.6%, we considered the isolate to be novel (154).

The genus level phylogenetic tree was generated using the full-length 16S rRNA gene sequences for the type species of each genus (85 species) and 35 additional type species that were added for context and later pruned (Fig. 2.6, Fig. 2.7, Supplemental Table 3). A MUSCLE (181) alignment and FastTree (182) were used to generate the tree on Geneious (2022.1.1) using default settings. Tree pruning and annotation for both trees were performed on iTOL (183).

Acknowledgements
We thank Janani Hariharan, Mary Firestone, Britt Koskella, Will Ludington, Amanda Shelton, Victor Reyes-Umana, and all members of the Taga Lab for helpful discussions. We are grateful to Victoria Innocent for providing corrinoids, Heejung Cho and Shi Wang for providing strains, Mary Firestone for allowing us to access her study site at Hopland Research and Extension Center, Katerina Estera Molina for soil sampling support, and Darryl Balderas for computational support. We thank Rebecca Procknow, Dennis Suazo, Janani Hariharan, and Eleanor Wang for critical reading of the manuscript. Sequencing was performed by QB3 Genomics, UC Berkeley, Berkeley, CA, RRID:SCR_022170 and the UC Berkeley DNA Sequencing Facility.

This work was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), Genomic Sciences Program (GSP) Grant DE-SC0020155 (M.E.T.), National Institutes of Health (NIH) Grant 3R35GM139633-03S2 (M.E.T.), the GEM Foundation (Z.I.A.A.), the Kase-Tsujimoto Foundation (Z.I.A.A.), and the Sponsored Projects for Undergraduate Researchers program at UC Berkeley (M.A.G.). E.L.B. was supported in part by the DOE, BER, GSP LLNL 'Microbes Persist' Soil Microbiome Scientific Focus Area SCW1632.

We acknowledge that this work was conducted on the ancestral and unceded land of the Ohlone and Pomo people.

Data Availability Statement
The sequencing data generated and analyzed during the current study are available in the NCBI GenBank repository under accession numbers OR878823-OR878983. Code generated during the current study is available from DOI: 10.5281/zenodo.10815172.

3. Corrinoid-based interactions between soil bacteria can be predicted based on monoculture growth

Zoila I. Alvarez-Aponte[1], Hazel Anagu[1], Valentine Trotter[2], Michiko Taga[1]

[1] Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, California, USA, 94706
[2] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, 94706

Abstract

Bacterial interactions in microbial communities are crucial for the structure and functioning of the environments they inhabit. Studying how bacteria interact is key for generating a predictive understanding of these communities. Using a model nutrient approach enables the study of a subset of interactions involving a single nutrient class and can shed light on general interaction mechanisms. Here, we focus on corrinoids, the vitamin B_{12} family of enzyme cofactors, to study nutrient competition and nutrient-sharing interactions in co-cultures and tri-cultures. We used bacteria that were previously isolated from a grassland soil and were characterized as corrinoid-dependent (require corrinoids but cannot synthesize them) or corrinoid providers (synthesize and release corrinoids). We report that the outcomes of competition for different corrinoids are predictable based on monoculture growth characteristics. We also found that corrinoid producers can support the growth of corrinoid-dependent bacteria in coculture and can influence the outcome of competition between dependents in triculture. Producer-dependent interactions were predictable for only one of our two producers, likely due to emergent properties of these interactions. We analyzed the metabolic capacity encoded in the genomes of these bacteria and found that corrinoids are the main metabolite driving their interactions. These results highlight the utility of the model nutrient approach to characterize and predict interactions in bacterial consortia of increasing complexity.

3.1 Introduction

From soils to the human gut, microbial communities shape processes crucial for human health and for the health of the planet. A growing area of interest is in predicting how different biotic and abiotic factors impact microbial communities (184–188). Interactions between microbes are strong determinants of community structure and function, with single microbes or single metabolites having the ability to impact the whole community (178, 189–192). Thus, understanding how microbe-microbe interactions occur and are regulated is crucial for our understanding of whole communities. However, microbial communities encompass complex networks with many metabolites being produced and consumed at a given time, making it difficult to observe and disentangle specific interactions (193–196).

Applying a model nutrient approach, the focus on interactions involving a single shared nutrient, allows for the mechanistic study of specific microbial interactions, while setting aside other interactions that could have confounding effects, as described in Chapter 1 and previously (7).

Corrinoids, the vitamin B_{12} family of cofactors (Fig. 3.1), are a useful model nutrient for several reasons. First, they are predicted to be shared, based on genomic data showing an estimated 86% of sequenced bacteria require corrinoids yet only 36% have the complete biosynthesis pathway (23). They have also been shown to be shared in bacterial co-cultures (7, 27, 31–33). Second, they are required enzyme cofactors for many metabolic processes that are relevant in different environments, from carbon and nitrogen metabolism to reductive dehalogenation (23, 197). Third, bacteria have distinct preferences for different corrinoids that impact their growth and metabolism (39). These characteristics, combined with thorough computational and experimental characterization of corrinoid biosynthesis and use, give us the ability to study corrinoid metabolism and corrinoid based interactions in microorganisms. The investigation of bacterial interactions through the reductionist lens of a model nutrient can enhance our ability to predict how different bacteria may cooperate and compete in consortia of different sizes.

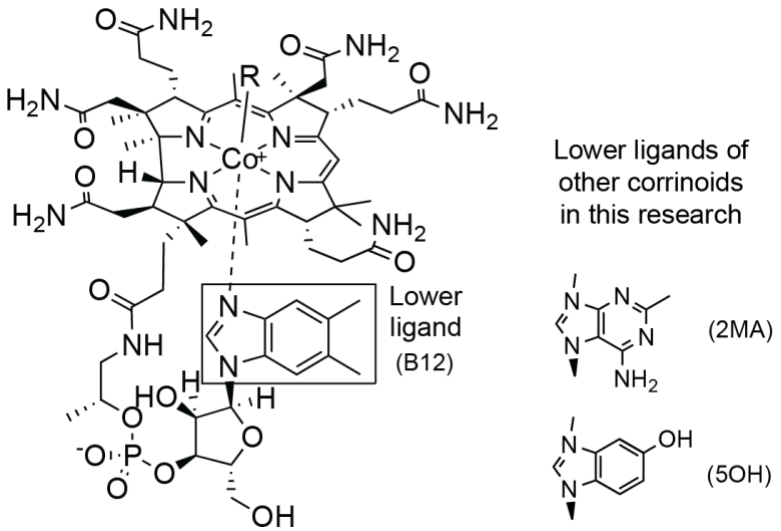


Figure 3.1. Cobalamin (B12) (left) and lower ligands of 2-methyladenine cobamide (2MA) and 5-hydroxybenzimidazole cobamide (5OH) (right). The three-character code used for each corrinoid is in parentheses next to its corresponding lower ligand.

As reported in Chapter 2, we previously isolated bacteria from a California grassland soil and characterized their corrinoid metabolism, gathering information necessary for application of the model nutrient approach to soil communities. After classifying isolates as "producers" and "dependents," we observed that corrinoid-dependent bacteria in our isolate collection preferred B12 but had the capacity to use different corrinoids to various extents. The low corrinoid concentrations required for growth of these soil isolates, compared to bacteria from other environments, suggest that soil contains a relatively low concentration of bioavailable corrinoid to which soil microbes have adapted (39, 112). If bioavailable corrinoids are limiting, dependent bacteria in soil might need to compete for these nutrients. Because different corrinoids can greatly influence the growth and metabolism of bacteria (39), it stands to reason that competition between dependents is shaped by the identity and amount of corrinoid present. Thus, we hypothesized that the outcome of competition can be predicted based on monoculture growth in consortia where competition for corrinoids is the predominant interaction.

All the corrinoid producers tested in Chapter 2 synthesize B12 in monoculture. Of this set, all contained intracellular B12 but only a subset of producers had detectable amounts of B12 in culture supernatants, indicating B12 release into the extracellular environment. We labeled this subset as "corrinoid providers" – producers capable of providing corrinoids to dependents (99, 100). Some producers provided B12 at levels exceeding the requirements of the dependents, leading us to postulate that these providers would support the growth of dependents in coculture. Further, we hypothesized that the extent to which each provider could support dependents would be predictable based on the amount of extracellular corrinoid detected in monoculture.

Table 3.1. Data collected previously on the four isolates used in this study.

Isolate	Previously published ID	Taxonomy	Corrinoid metabolism category	Extracellular corrinoid
*Ca*19	CRE_019	*Caulobacter* sp.	Dependent	N/A
*Ca*24A	5OH_024A	*Caulobacter* sp.	Dependent	N/A
*Pr*24B	5OH_024B	*Priestia* sp.	Producer	0.12 ± 0.002 nM[†] (1.3% of total corrinoid)
*Me*12	NOC_012	*Mesorhizobium opportunistum*	Producer	4.0 ± 0.82 nM[†] (19% of total corrinoid)

[†] *These corrinoid concentrations were measured in medium supplemented with additional amino acids, which results in increased producer growth. The concentrations in the medium used here are presumed to be lower.*

Here, we characterized corrinoid-based interactions among four bacterial isolates identified in Chapter 2 (Table 3.1). We found that the outcome of corrinoid competition between pairs of dependent bacteria could largely be predicted from differences in their growth in monoculture, with the dominant microbe determined by the corrinoid concentration. We also found that providers could support the growth of dependent bacteria in coculture. When grown in tri-cultures, producers could support the growth of both dependents and the dependents competed against one another. Further, we sequenced the whole genomes of these bacteria and found that corrinoids were the main nutrient predicted to be shared between producers and dependents. Knowledge about corrinoid sharing and competition among soil bacteria contributes to better comprehension of pairwise bacterial interactions and enhances our understanding of how corrinoid diversity may impact microbial community dynamics at a large scale.

3.2 Results

3.2.1 Dependent isolates compete for corrinoids

Because corrinoid-dependent bacteria are uniquely influenced by corrinoid structure and concentration (39), we hypothesized that the outcome of competition between two dependent bacteria could be predicted based on the corrinoid preferences observed in monoculture growth. We tested this hypothesis on a pair of corrinoid-dependent soil isolates, *Ca*19 and *Ca*24A, initially selected because they grew reproducibly with no aggregation in liquid culture. We

measured the maximum growth rates of the two dependent isolates in media containing a range of concentrations of 2MA, 5OH, and B12 to determine how these corrinoids influence growth in monoculture (Fig. 3.2A-D). Ca19 has a higher growth rate than Ca24A at low concentrations of 5OH and B12 but grows more slowly at high concentrations (Fig. 3.2B-C). The corrinoid concentration required to reach half-maximal growth (EC_{50}) was also lower for Ca19 on these two corrinoids (Fig. 3.2D) suggesting that it is adapted to lower corrinoid concentrations. A comparison of the data generated from the two isolates was used to predict the outcome of competition at specific concentrations. Specifically, the isolate with the higher growth rate at a given corrinoid concentration was predicted to outcompete the other. We differentiated between the two isolates based on their distinct colony morphologies (Fig. 3.2E). Through the genome sequencing effort described in a later section, we found that the isolates belong to two different species of *Caulobacter*, as reflected by a whole genome average nucleotide identity (ANI) of 89% (198). The species were not identified by GTDB-Tk (199, 200), which suggests that they could be novel species.

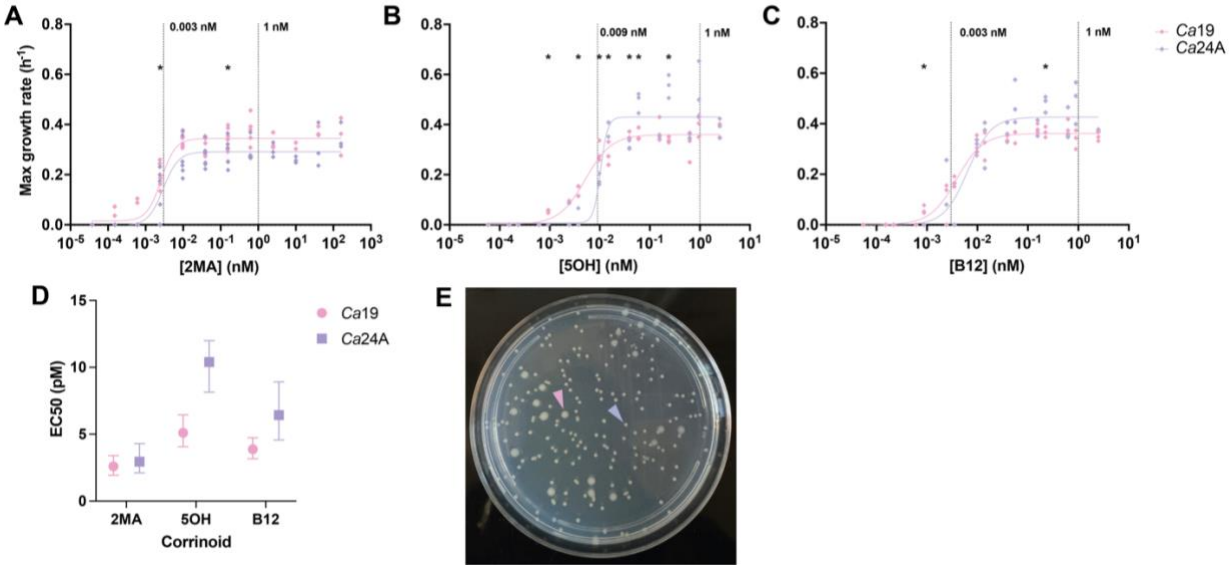


Figure 3.2. Dependent isolates *Ca*24A and *Ca*19 display different growth rates at varying concentrations of (A) 2MA, (B) 5OH, and (C) B12. Curve fits are for a four-parameter least squares fit. Asterisks denote Welch's t-test $p<0.05$ and dotted vertical lines denote the concentrations selected for further experiments. (D) The corrinoid concentration that leads to half maximal growth (EC_{50}) for each isolate on 2MA, 5OH, and B12 highlights the differences in their corrinoid preferences. Error bars represent 95% confidence interval. (E) Colonies of isolates *Ca*24A (small colonies, purple arrow) and *Ca*19 (large colonies, pink arrow) after 72 hours of growth on R2A medium containing B12.

Based on their growth rates in different concentrations of 5OH and B12, we predicted that *Ca*19 would outcompete *Ca*24A when co-cultured at concentrations less than 0.012 and 0.010 nM, respectively. Conversely, because *Ca*24A had higher growth rates at higher concentrations of both corrinoids, it was predicted to outcompete *Ca*19 at corrinoid concentrations higher than 0.012 and 0.010 nM. We considered predictions for 5OH to be most reliable because isolate growth rates were significantly different at seven concentrations (Fig. 3.2B). Predictions of co-

culture outcomes in 2MA were less clear because the average growth rate measurements for the two isolates were very similar at all concentrations (Fig. 3.2A). Nevertheless, we predicted *Ca*19 to be more abundant at all concentrations.

We tested our predictions by co-culturing *Ca*19 and *Ca*24A in media containing a high (1 nM) or low (0.003 or 0.009 nM) concentration of each corrinoid and passaging every 24 hours to maintain cultures at exponential phase. The viable cell number of each strain was determined for the inocula and co-cultures on days 1, 3, 5, and 7 (Fig. 3.3).

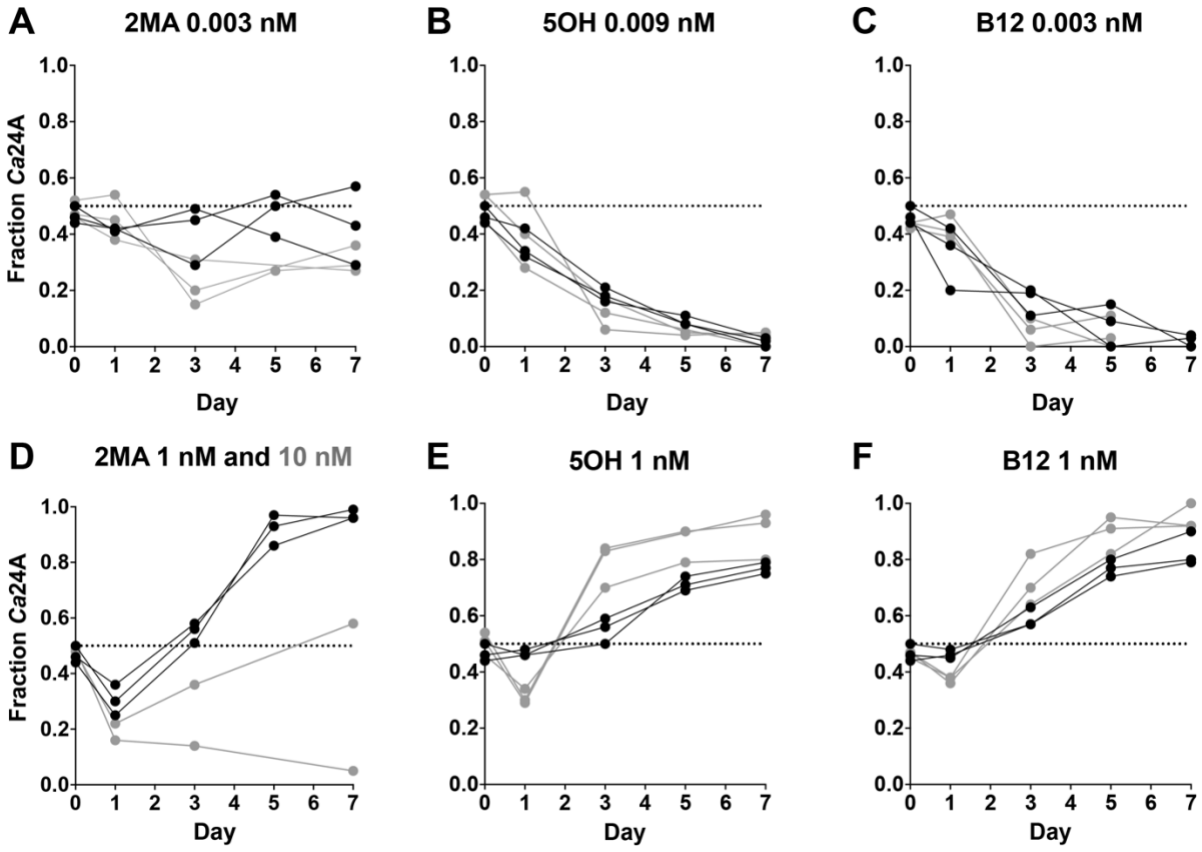


Figure 3.3. Growth competition between dependents Ca19 and Ca24A in different corrinoid conditions. Graphs show the relative abundance of *Ca*24A in co-culture with *Ca*19 passaged daily for seven days, based on the number of colonies on solid media from samples taken at the indicated time points. Black and gray lines show data for five to six biological replicates from experiments performed on two different days under the same conditions with the exception that 2MA was repeated using a different high corrinoid concentration (gray lines, 2MA). The dotted line indicates 50% abundance.

Predictions for the outcome of competition between the two dependents were accurate in the high and low concentrations of 5OH and B12, with *Ca*19 having a higher abundance at low corrinoid concentrations and *Ca*24A becoming predominant at high concentrations. These results were consistent with our hypothesis that the outcome of competition could be predicted from maximum growth rates in monoculture, which suggests that competition for corrinoids is the main interaction driving dynamics in these co-cultures. Overall, the competition for these two

corrinoids is determined by concentration, suggesting that *Ca*19 is adapted to lower concentrations of 5OH and B12 and *Ca*24A to higher concentrations.

In the case of 2MA, the outcome of competition was not predictable from monoculture growth rates at the high or low concentrations. The isolates remained at equal proportions in the low concentration condition, while *Ca*24A outcompeted *Ca*19 in the high concentration condition. When the experiment was repeated, 5OH and B12 outcomes remained the same, but 2MA cocultures trended differently (Fig. 3.3 (gray lines)). This is likely because the differences in growth rates or EC_{50} were not large enough to define a corrinoid dependent outcome of competition.

3.2.2 Producers can support dependents by providing a corrinoid

Next, we hypothesized that each dependent isolate could be sustained by a producer in co-culture, and that the amount of corrinoid the producers supply would determine the outcome of competition between the two dependents in triculture. We selected two isolates from the same collection that were previously found to produce and release different amounts of B12 (Table 3.1). *Pr*24B, an isolate with 99.9% 16S rRNA identity to *Priestia arhyabattai,* was originally isolated in the same culture as the dependent *Ca*24A, and the two were later cultured separately. Thus, these bacteria may interact in their natural soil environment and/or be physically associated. A second producer, *Me*12, a strain of *Mesorhizobium opportunistum,* was found in our previous study to release the highest amount of B12 into monoculture supernatants (Table 3.1). Additionally, the colony morphologies of these two producers were easily distinguishable from those of the dependents, making it possible to determine their abundances based on colony numbers.

We found that both dependents could grow in corrinoid-free medium containing cell-free supernatants from the producers (Fig. 3.4A). The two producers supported dependent growth to different extents; *Pr*24B supernatant supported less than one dependent cell per producer cell, while *Me*12 supernatant supported between 1 and 27 dependent cells per producer cell (Fig. 3.4).

When grown in co-culture with each dependent, *Me*12 supported approximately 1 dependent cell per producer cell, on the lower end of the number of dependents supported by Me12 supernatant (Fig. 3.4B). In contrast, *Pr*24B supported approximately 20 dependent cells per producer cell in coculture, 1000-fold higher than supernatants of the same isolate, indicating that emergent properties arise when these isolates are co-cultured. The two dependents were supported to similar extents in both supernatants and cocultures, suggesting that their corrinoid requirements were satisfied similarly. It is possible that *Me*12 provides less corrinoid in co-culture than in monoculture due to competition or inhibition by a partner. The difference between supernatant and co-culture growth with *Pr*24B may be the result of higher corrinoid providing by the producer when a partner is present.

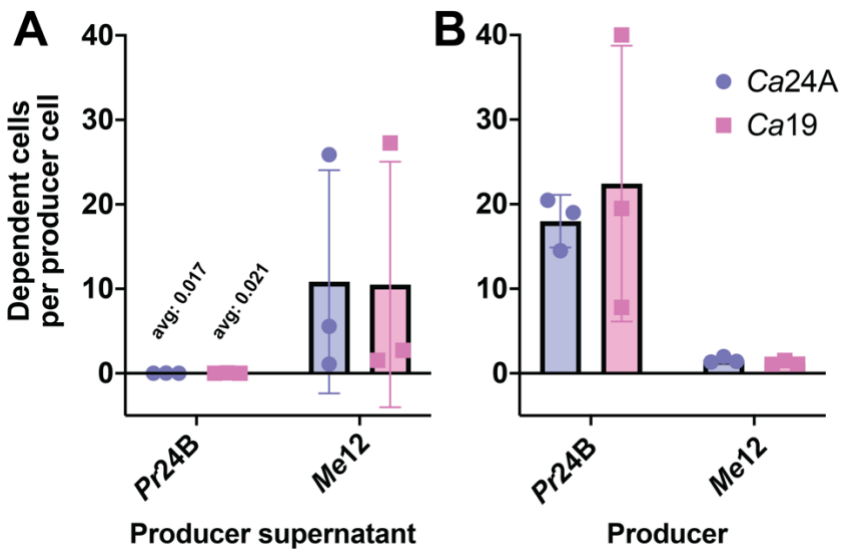


Figure 3.4. Growth of corrinoid dependent isolates (A) with supernatant from producer cultures, and (B) in co-culture with a producer. Bars represent the average of three replicates shown as individual points, and error bars represent standard deviation.

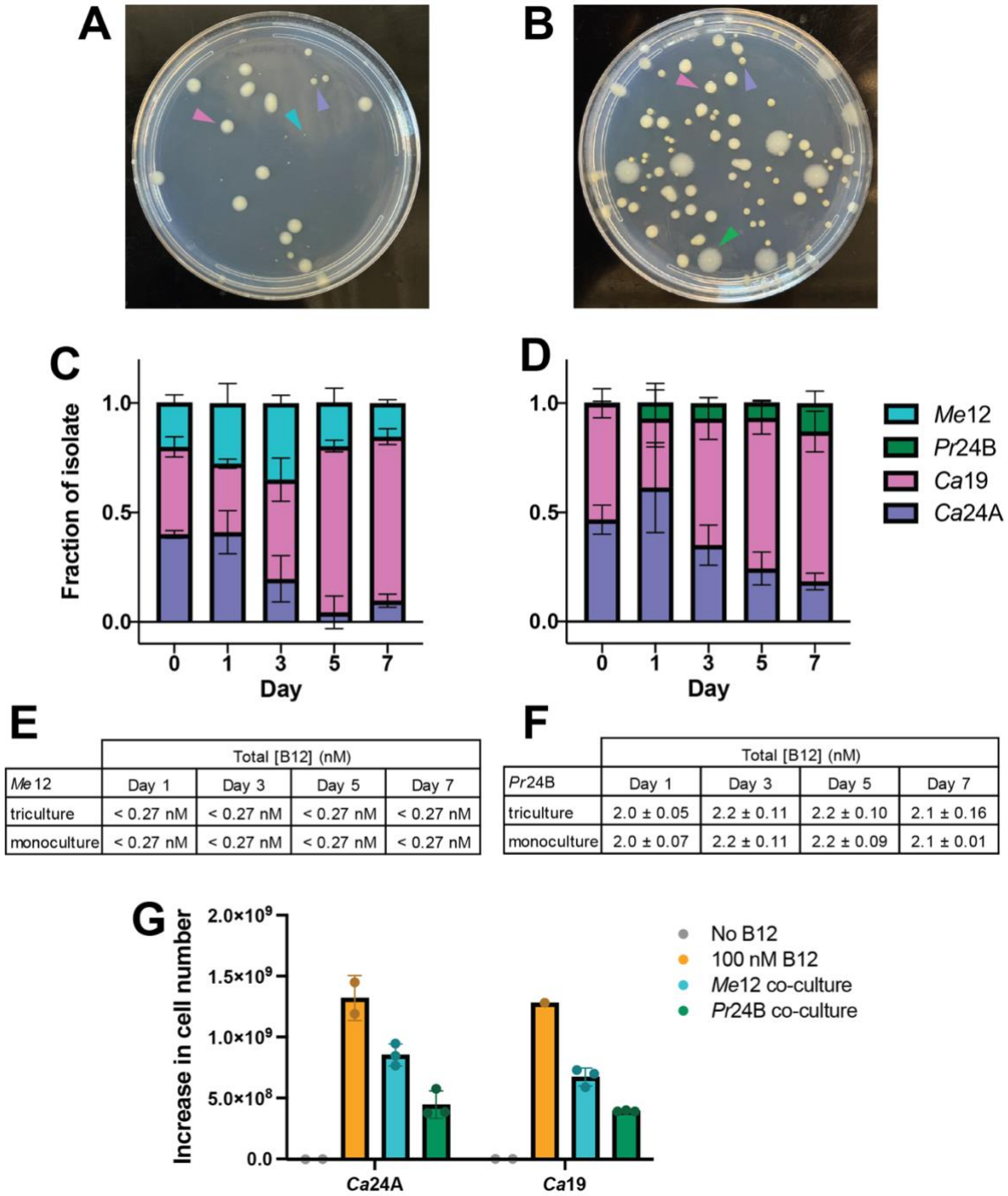
3.2.3 Corrinoid sharing and competition occur in three-species consortia

Upon confirming that each producer could support both dependents, we sought to determine how producers impact competition between the dependents. We did this by growing tri-cultures of the two dependents with each producer in a corrinoid-free medium with daily passaging into fresh medium. On days 1, 3, 5, and 7, the cultures were plated to determine the abundances of each strain based on their unique colony morphologies (Figs. 3.5A, B).

In both tri-cultures, *Ca*19 was the more abundant of the two dependents by day 3 and its relative abundance increased at later timepoints. We lysed the tricultures and producer monoculture controls and measured total corrinoid via an *E. coli*-based bioassay to evaluate if the outcome of competition correlated with the corrinoid concentrations detected (20). In the *Me*12 triculture, B12 was below the bioassay detection limit (0.27nM) at all timepoints (Fig. 3.5C). Low corrinoid concentrations explain why *Ca*19 outcompeted *Ca*24A in this tri-culture, consistent with coculture results where *Ca*19 was found at high abundance at low B12 concentrations. Thus, the *Me*12 triculture result was predictable from previous co-culture and monoculture results.

The *Pr*24B triculture had a stable corrinoid concentration of approximately 2 nM over multiple passages, which was also sustained in the producer monoculture, suggesting that the dependents do not trigger higher corrinoid production in *Pr*24B (Fig. 3.5D). From our dependent co-culture experiments we expected *Ca*24A to outcompete *Ca*19 in these high B12 concentrations, but the opposite occurred. The simplest interpretation of these results is that only a small amount of the measured corrinoid was available in the supernatant, which would explain why *Ca*19 was the dominant dependent strain in these conditions. However, it is also possible that other interactions occurred in the tri-culture that benefited *Ca*19 over *Ca*24A, such as sharing of other nutrients.

Growth of *Ca*24A did not seem to be inhibited in co-cultures with *Pr*24B or *Ca*19, which makes it unlikely that it was inhibited in the tri-culture, although we cannot discard the possibility.



	Total [B12] (nM)			
*Me*12	Day 1	Day 3	Day 5	Day 7
triculture	< 0.27 nM	< 0.27 nM	< 0.27 nM	< 0.27 nM
monoculture	< 0.27 nM	< 0.27 nM	< 0.27 nM	< 0.27 nM

	Total [B12] (nM)			
*Pr*24B	Day 1	Day 3	Day 5	Day 7
triculture	2.0 ± 0.05	2.2 ± 0.11	2.2 ± 0.10	2.1 ± 0.16
monoculture	2.0 ± 0.07	2.2 ± 0.11	2.2 ± 0.09	2.1 ± 0.01

Figure 3.5. Outcome of competition between *Ca*19 and *Ca*24A in tri-culture with each producer. (A, B) Colony morphologies of both dependent isolates and (A) *Me*12 and (B) *Pr*24B after 96 and 72 hours of growth on solid medium, respectively. Arrows correspond to the color of each isolate in the rest of the figure. (C, D) Relative abundances of producer and both dependents in triculture with (C) *Me*12 and (D) *Pr*24B are shown. Bars represent the average of three biological replicates and error bars represent standard deviation. The tables in E and F show average and standard deviation for total B12 measured via

bioassay in the tricultures and in monocultures of (E) Me12 and (F) Pr24B, respectively (n=3). <0.27 nM indicates the measurement was below the bioassay detection limit of 0.27 nM. (G) Increase in dependent CFU/ml after growth with 100 nM B12 or in co-culture with Me12 or Pr24B. Bars represent the average of one or two replicates and errors bars represent standard deviation.

3.2.4 Corrinoids appear to be the only shared nutrient in these consortia

To determine if other interactions were occurring between producers and dependents, we compared growth of the dependents in co-culture with each producer to growth of the dependents in excess B12 (100 nM). The cell counts of both dependents reflected a larger increase when B12 was provided in excess, suggesting that B12 is the main shared nutrient, and the amount made available by the producers is limiting (Fig. 3.5E).

To further investigate other metabolic capabilities that could impact growth of the isolates in tri-cultures, we turned to isolate genome sequencing. We successfully obtained sequences for three of the four isolates (*Ca*24A, *Ca*19, and *Me*12), to investigate whether they could produce nutrients that are commonly shared, such as amino acids and vitamins (Fig. 3.6). We also considered the possibility that certain isolates could degrade carbon sources that others could not, making byproducts available, so we investigated the catabolism pathways for the four carbon sources present in our medium.

After genome analysis, we found that the two dependents and the producer for which we had genome sequences were congruent in their metabolic capabilities. These include amino acid biosynthesis, where all three isolates are predicted to synthesize all the amino acids evaluated, except that *Ca*19 was missing one gene in the leucine biosynthesis pathway, possibly rendering it unable to produce leucine (Fig. 3.6). Regarding vitamin biosynthesis, all are predicted to synthesize pantothenate and tetrahydrofolate and lacked complete pathways for the synthesis of thiamine, riboflavin, biotin, and niacin (Fig. 3.6). The only difference was observed in cobalamin (B_{12}) biosynthesis, where the pathway presence and absence aligned with the phenotypic observations reported in Chapter 2, with the producer *Me*12 predicted to produce B12 and both dependents lacking the biosynthetic pathway. The three isolates also had the same pathways for catabolism of carbon sources in the growth medium; they shared the capability to degrade glucose, N-acetylglucosamine, and xylose, but lacked xylan catabolism genes (Fig. 3.6). From this genome analysis, we conclude that the main shared nutrient in our co- and tri-cultures is B12, and sharing of other vitamins or amino acids is not occurring in these consortia.

Since it was the main difference in metabolic capabilities and it agreed with our previous phenotypic characterization of the isolates, we further characterized the cobalamin biosynthesis and dependence genes in the three genomes. *Me*12 was genomically classified as a producer according to previously established thresholds (23). It has the 5,6-dimethylbenzimidazole synthesis gene *bluB*, in accordance with our previous observation that it synthesizes B12, reported in Chapter 2 (56, 58). The two dependent genomes lacked corrin ring and lower ligand biosynthesis genes, while they contained genes for tetrapyrrole precursor biosynthesis, which are shared with other pathways and not specific to corrinoid biosynthesis. However, they both had adenosylation, nucleotide loop assembly, and lower ligand attachment genes, which are required for salvaging late corrinoid precursors (23, 61, 201). The latter observation aligns with

experimental results in Chapter 2, which showed these two dependent isolates can grow when provided with cobinamide, a corrinoid precursor lacking the lower ligand. The two dependent genomes, however, lacked the known corrinoid lower ligand biosynthesis genes (23), suggesting that when they salvage corrinoid precursors they must also salvage DMB, or that they salvage cobinamide to produce a corrinoid other than B12.

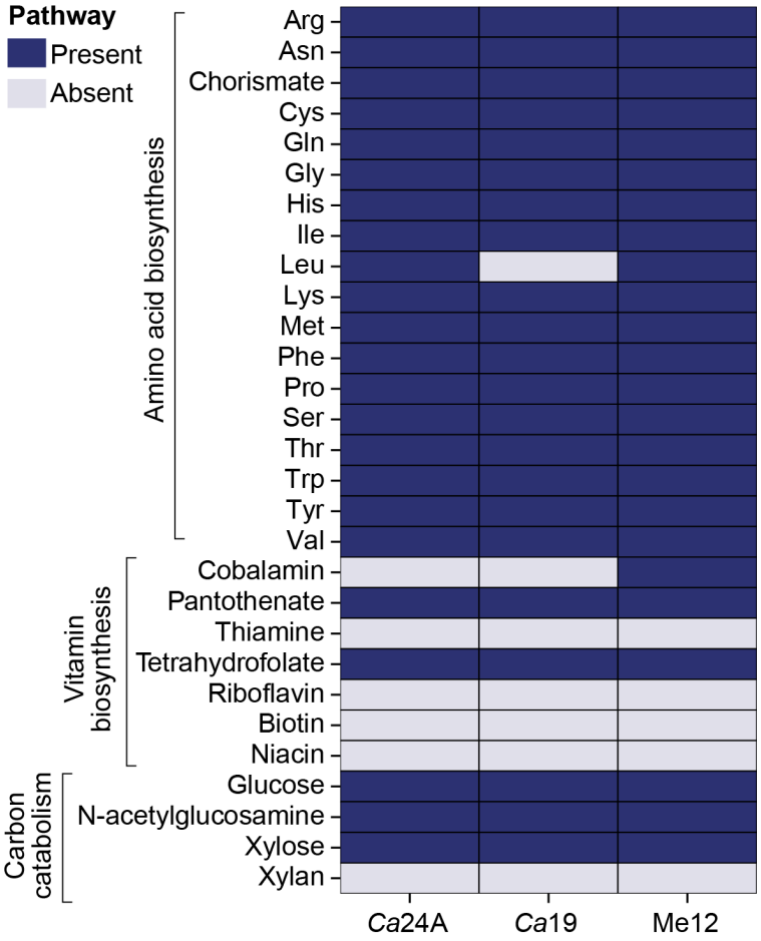


Figure 3.6. Amino acid, vitamin, and carbon metabolism pathways present or absent in the genome of each isolate.

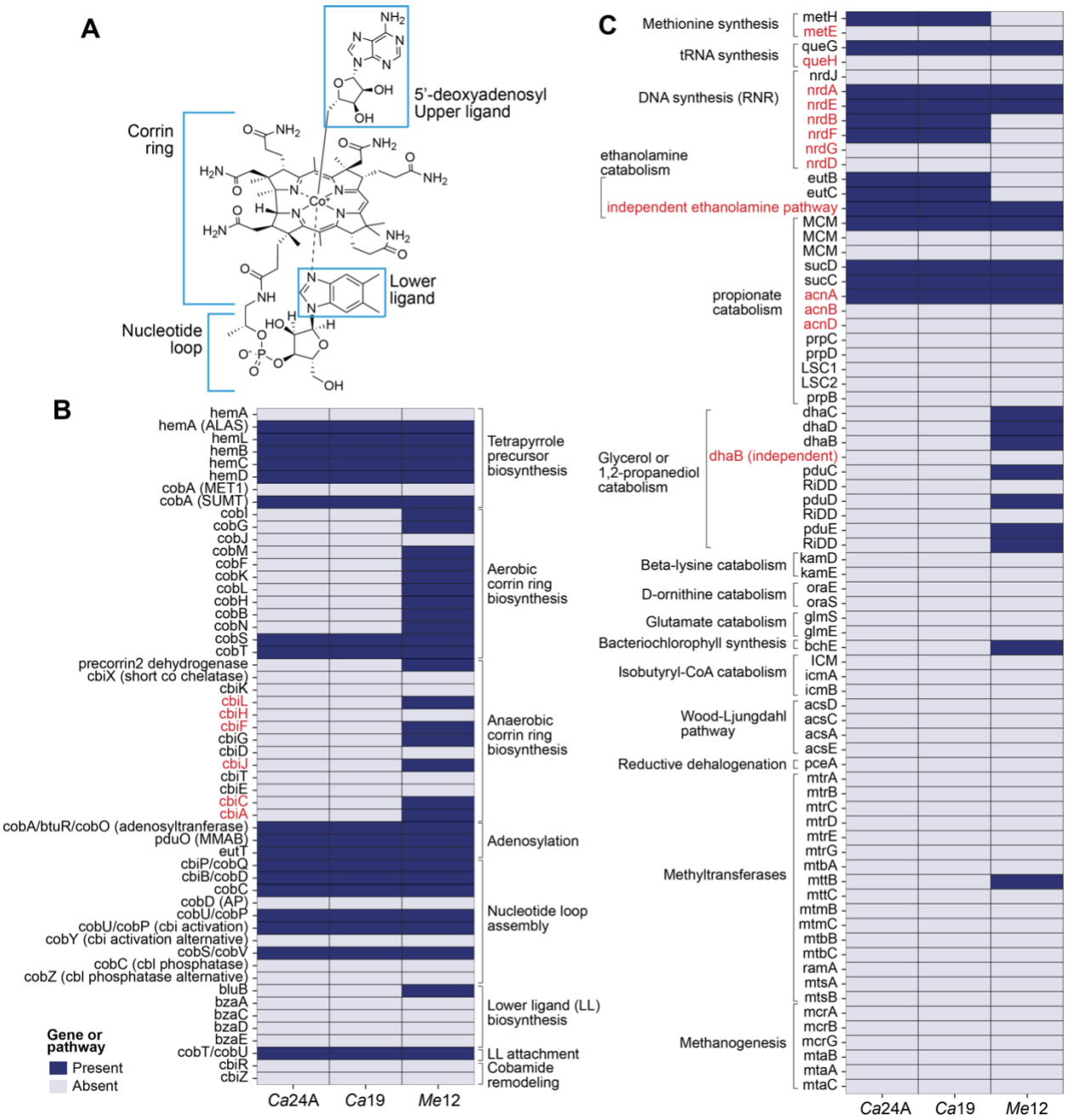


Figure 3.7. Presence or absence of corrinoid biosynthesis and corrinoid dependent genes and pathways. (A) Structure of adenosylcobalamin (B_{12}). (B) Corrinoid biosynthesis genes are on the left and the portion of the pathway they belong to is on the right of the heatmap. Gene names in red function in anaerobic corrin ring biosynthesis that share KOfams with genes in aerobic corrin ring biosynthesis. (C) Corrinoid dependent genes or pathways are shown in black text and corrinoid independent alternatives are shown in red text. Genes are grouped by the type of metabolism in which they are involved.

3.3 Discussion

Cooperative interactions such as cross-feeding and competitive interactions such as resource competition between pairs of bacteria can shape microbial community structure (202, 203). Understanding the contribution of particular interactions in detail, however, is complex because additional factors emerge as communities become more complex (196). Thus, studying small consortia with a focus on a single model nutrient can allow for the elucidation of specific interactions in a controlled environment.

Here, we set out to test predictions of corrinoid-based interactions in co- and tri-cultures of soil bacteria. By focusing our approach on a model nutrient, we successfully predicted the outcome of competition for corrinoids between dependent bacteria based on their differential concentration-dependent growth rates in monoculture. Our competition predictions were successful in the two corrinoids for which the dependent isolates had sufficiently distinct preferences. Further, we confirmed that corrinoid producers can support dependents by sharing corrinoids and can impact the dynamics of competition between dependents in a triculture. Interactions between *Pr*24B and both dependents were difficult to predict. Although we observed that the total corrinoid produced remained unchanged, our results suggest that changes in corrinoid release may be triggered when producers are grown with a partner. Finally, we concluded from whole genome analysis and growth experiments that corrinoid-dependent interactions are the main vitamin-based interaction between these bacteria. Together, these results reveal that corrinoid sharing and competition are central determinants of interaction dynamics among these isolates and suggest that these interactions may happen in the native environment.

Using pairwise interactions to predict growth in consortia of increasing complexity is difficult because of the emergence of higher order interactions (196). The corrinoid sharing interactions observed in co-cultures were conserved in both tri-cultures, suggesting that corrinoid sharing is maintained as community complexity increases. It appears that competition for corrinoids between dependents was also conserved in tri-cultures, although the outcome of competition could only be predicted based on tri-culture corrinoid concentrations in one case. The observed interactions may be the result of changes in gene expression that impact production or providing of this enzyme cofactor, thus affecting dependent partners. The conservation of corrinoid-based interactions in consortia of higher complexity would enable further predictions and shed light on the importance of corrinoid sharing and competition in communities.

Interactions between *Ca*24A and *Pr*24B were of particular interest because their joint isolation suggests they could have a pre-established metabolic partnership. They were indeed able to grow together in co-culture when *Pr*24B provided B12. Additionally, *Ca*24A was sustained in the tri-culture when *Pr*24B was the producer, suggesting it was somehow favored more by this producer than by *Me*12. However, the advantage was not sufficient for *Ca*24A to outcompete *Ca*19.

Our two dependents belonged to genus *Caulobacter* but are likely different species. Despite belonging to different species, they had the same corrinoid biosynthesis and dependence genes and their corrinoid metabolism differed solely in their adaptations to high or low concentrations of B12 and 5OH, which cannot be predicted from genome sequences. Both dependents encode corrinoid dependent genes involved in methionine synthesis, tRNA synthesis, ethanolamine catabolism, and propionate catabolism. Only the methionine and tRNA synthesis genes, *metH*

and *queG*, are relevant in our growth conditions. The epoxyqueuosine reductase encoded by *queG* has not been found to influence growth of *E. coli* (41), thus their dependence in these cultures is likely solely due to *metH*, because methionine is not present in the medium. It is curious that although the two dependents belong to the same genus and completely overlap in the corrinoid dependence and biosynthesis genes they encode, they have adaptations to different corrinoid concentrations. This could be due to experiencing distinct conditions in their native soil environment, such as close versus distant associations to corrinoid producers. However, it also shows that even for genera where corrinoid metabolism is conserved, different species or strains may still display distinct preferences.

It was previously shown that the bulk soil from which these bacteria were isolated contains large amounts of vitamin B_{12} (19, 112). However, because a large portion of the corrinoid is thought to be adhered to the soil matrix, we hypothesize that only a small fraction is available to the soil microbiota. Additionally, nutrient availability is known to vary greatly in different microenvironments within soil (145). The EC_{50}s observed for corrinoid dependent bacteria isolated from this soil were in the picomolar range. EC_{50}s in this range are common for microorganisms in ocean water, which contains corrinoid concentrations in the picomolar range, further suggesting that these bacteria have adapted to low corrinoid bioavailability in their native environment (39, 99, 134, 162). We hypothesize that despite the high concentration of B_{12} in this soil, corrinoids are not largely bioavailable and represent a growth-limiting nutrient for which bacteria must compete. This further highlights the importance of the corrinoid concentration adaptations in our two *Caulobacter* sp. isolates.

Our study was limited because the soil environment from which these bacteria were isolated contains 95% B_{12}, making it one of the least diverse environments in terms of corrinoid diversity (19). The dependent bacteria in this study have likely only experienced exposure to B_{12}, which might explain why the main difference we see between them is in terms of the concentration of corrinoid present, rather than the identity of the corrinoid. Hypotheses about competition for corrinoids and the niche diversity enabled by different corrinoids may be better suited in environments with a higher corrinoid diversity, such as the human gut (19, 66).

Together, we were able to predict the outcome of dependent-dependent and producer-dependent interactions based on growth in monoculture. This represents a significant advance in the applicability of corrinoids as model nutrients. Next steps include testing the predictability of corrinoid interactions between more bacteria pairs and evaluating the factors that make predictions possible. Further, elucidation of the mechanisms of corrinoid providing may enable genome-based predictions of producer-dependent interactions. Our prediction and evaluation of interactions is unique in that it was based on concentrations of a shared nutrient and performed using bacteria isolated from the same soil environment. This research adds to a body of literature that has shown that predictions about nutrient sharing are possible from a combined analysis of metagenomes, genomes, and phenotypic data (31, 165, 203, 204), and that nutrient availability impacts microbial competition and coexistence (205).

3.4 Methods

3.4.1 Isolate characterization and selection

Bacteria were grown at 28ºC with aeration in liquid VL60 medium as described in Chapter 2 or on solid R2 agar. Dependents were pre-cultured by growing to saturation in liquid medium containing 1 nM cyanocobalamin and transferred to medium containing no corrinoid at a 1:100 dilution after 24 hours. The latter cultures were used as the inoculum for experiments. Corrinoid producer isolates were pre-cultured by growing to saturation in medium containing no corrinoid.

Growth curves were performed by culturing the isolates in 200ul cultures in 96-well plates in a Biotek Synergy 2 plate reader with OD_{600} measurements taken every 15 minutes. Dose-response curves were acquired by inoculating into 12 different concentrations of each corrinoid as described in Chapter 2 with OD_{600} recorded every 15 minutes. The OD_{600} values were normalized to a 1 cm pathlength, blank subtracted in a replicate specific manner by subtracting the mean of the first two time points and the known initial abundance and smoothed with 0^{th} order smoothing with 3 neighbors on either side on GraphPad Prism (v10.4.0). Values below 0.009 were eliminated and maximum growth rates were obtained using AMiGA (206) and dose response curves were plotted and fitted on GraphPad Prism.

3.4.2 Co-cultures of corrinoid dependent isolates

Experiments were performed on three independent cultures of each strain. Each strain was inoculated at a starting OD of 0.0015 for monoculture and coculture experiments. 1 ml cultures were grown in 96-well plates with 2 ml wells covered with an Aeraseal (Sigma Aldrich) at 28ºC and 800rpm in a 96-well plate shaker (Southwest Science). 10µl of culture were transferred into 990ul of fresh medium containing the same corrinoid conditions every 24 hours for 7 days. Serial dilutions and plating on R2A medium were done on days 0, 1, 3, 5, and 7 and colonies were counted after 48 hours of growth at 28ºC.

3.4.3 Producer supernatant assay

Three independent cultures of each producer strain were inoculated from glycerol stocks, grown to saturation, and plated for viable cell counts. The cultures were centrifuged, and the supernatant was sterilized by filtration with a 0.22 µm syringe filter. Dependent pre-cultures were prepared as described above and inoculated into medium containing producer supernatant at various dilutions in a 96-well plate. Serial dilutions and plating on R2A medium were done for the inoculum and after 48 hours of growth of each culture to determine the viable cell count. Colonies were counted after 48 hours of growth and the number of dependents supported by each producer was determined by dividing dependent viable cell counts by producer viable cell counts adjusted for the supernatant dilution factor.

3.4.4 Producer-dependent co-cultures

Dependent cultures were inoculated in triplicate from glycerol stocks and pre-cultured as described above. Triplicate producer cultures were grown from glycerol stocks in medium lacking corrinoid for 24 or 48 hours for *Pr*24B and *Me*12, respectively. We diluted the dependent and producer liquid cultures 1:40 in fresh VL60 medium and mixed equal volumes of each as described previously (165). Cultures were grown in triplicate on a 96-well plate, serially diluted, and plated to determine the viable cell count in the inoculum and the culture after 48 hours. The

number of dependent cells supported by each producer was determining by dividing the endpoint dependent viable cell counts by the producer viable cell counts.

3.4.5 Tri-culture growth

Producers were cultured for 24 hours in medium lacking corrinoid, and dependents were pre-cultured as described above. Cultures were adjusted to and OD_{600} of 0.10 and mixed in equal volumes to prepare the inoculum. 10 µl of inoculum was added to 990 µl of medium with no corrinoid in a 96-well plate with 2 ml wells. Cultures were transferred and plated as described above. 150 µl of culture were collected on days 1, 3, 5, and 7 to determine the corrinoid concentration via an *E. coli* bioassay as described previously (20).

3.4.6 Genome sequencing and analysis

Genomes were sequenced by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China) and trimmed, assembled, and annotated using Kbase as described previously (207, 208), and filtered for >95% completeness and <5% contamination. Genomes were recovered for three out of the four isolates, as the genome for *Pr*24B had an insufficient number of reads. The DRAM metabolism summary was used to make calls about gene presence or absence for genes distilled by DRAM. This included most corrinoid biosynthesis and dependence genes. Raw annotations of other genes of interest acquired from DRAM were added to the analysis to determine gene presence and absence based on KOfams or Pfams. Genes were annotated as present when annotated by DRAM with a rank of at least C (209). The genes identified for corrinoid biosynthesis, dependence, and independent alternatives to corrinoid dependent pathways were the same we used previously (23, 112). Transport genes and B_{12}-binding domains were left out of the analysis due to poor annotations, and rSAM-B_{12} proteins were omitted because this category included the Radical SAM superfamily, which includes but is not restricted to B_{12}-dependent rSAM enzymes. Additionally, genes that are not essential for corrinoid biosynthesis or that are only found in archaea were not included.

Amino acid biosynthesis and small carbon source catabolism capability were assessed using GapMind (210). Pathway genes were considered to be present if they had at least a medium confidence hit as described previously (139). Synthesis capacity of B vitamins other than cobalamin was assessed by evaluating the biosynthesis pathways with previously established thresholds (165, 204). Genes of interest identified by previous work were cross-referenced against the DRAM annotations.

Acknowledgements

We are thankful to Eleanor Wang for assistance with genome analysis, Kenny Mok, Christine Qabar, and Will Ludington for help with experimental design, Zachary Hallberg for statistical recommendations, and Hans Carlson for assistance with sequencing efforts.

This work was funded by the National Institutes of Health (NIH) Grant 3R35GM139633-03S2 (M.E.T.). H.A. was supported by the Bay Area NSF RaMP (Grant # 2216550).

4. Design, implementation, and evaluation of a qualifying exam preparation program for STEM PhD programs

Zoila Alvarez-Aponte[1#*], Amanda J. González-Segarra[1#*], Denzil Streete[2], Diana Bautista[1], Rachel Brem[1].

[1]University of California, Berkeley, United States;
[2]Massachusetts Institute of Technology, Cambridge, United States;

*Contributed equally
[#]Corresponding authors

Abstract

The qualifying exam (QE), also known as the preliminary exam, is one of the two major evaluative milestones in graduate programs. It is usually administered in the second or third year of the program and is meant to evaluate if a student is prepared to undertake an individual thesis project. Preparation for the qualifying exam is often daunting, as many programs do not have structured guidelines, and it is the first evaluation of its kind that many students encounter. We developed and implemented a pilot for the Inclusive Excellence in Quals Prep (IEQP) program, designed to provide mentorship, community, and academic support for students from diverse backgrounds as they prepared for their QE. The program focused on three STEM PhD programs at the University of California, Berkeley, which have similar QE formats. The main components for IEQP included pairing students with graduate student mentors, academic and wellness workshops, and community-building events. Here, we evaluated the program's effect on the pilot cohort of 11 students. The most significant component of the program, as per student feedback, was peer mentorship. Overall, we saw an increase in students' perception of their preparedness, QE-related skills, the support received from their advisors, and the agency they felt over their proposed work. We conclude that the pilot program was successful as it increased the structure around QE prep and provided students with mentors in their field. We recommend broad implementation of similar programs to enhance equity in graduate education.

4.1 Introduction

In comparison to their share of the U.S. population, Women, Black, Hispanic, and American Indian/Alaska Native scientists are underrepresented in STEM. Racially minoritized groups comprise 16.9% of doctoral degree recipients in STEM (211), but 42.2% of the U.S. population (212). Multiple factors contribute to the lack of diversity in academia, with retention of students from minoritized groups being a crucial, often overlooked, component (213, 214). There are many stages at which retention can be affected, especially around major evaluations or program requirements. There are two major evaluation components in PhD programs: the qualifying exam (QE) or preliminary exam early in the program and the dissertation defense at the end of the program.

In 2015, the National Science Foundation Alliance for Graduate Education and the Professoriate (NSF AGEP) California Alliance (Berkeley, Caltech, Stanford, UCLA) conducted a wide-ranging survey of graduate students across STEM to pinpoint ways to improve the success of PhD students, and eventually increase diversity in STEM leadership positions. This study found that women and minority students were most likely to publish at rates comparable to their male majority peers if they felt that they were prepared for their graduate courses, accepted by their colleagues, and enrolled in a structured PhD program. Authors also found that stating what was expected of the students clearly, deeming that they were prepared for graduate level courses, and feeling accepted by their colleagues reduced student distress levels (215). Here, we identified a way to increase the retention of students from minoritized groups in STEM by creating a program that adds structure to the QE preparation and evaluation processes.

The goal of a QE is to demonstrate that the student has knowledge in their field of study and is prepared to undertake independent research. Once a student passes their QE, they advance to candidacy, after which they carry out thesis research until their committee deems they are ready to defend their dissertation and obtain the PhD (215, 216). Depending on the graduate program, a student usually has 1-3 attempts at the QE, after which they are dismissed from the doctoral program.

The QE can vary greatly depending on the graduate institution, and even within the same institution, different departments have different QE formats. For example, some QEs are a 2-3 hour oral examination, which may also include a written component. The committee can examine the student's knowledge on their proposed research, topics related to their field of study, and more general topics within their broader field (217). Meanwhile, some QEs are a written examination spanning several days, where the student answers questions provided by their committee. These questions can cover general topics and topics in the student's specialization (218). In the Helen Wills Neuroscience Institute (HWNI), Molecular and Cell Biology (MCB), and Plant and Microbial Biology (PMB) programs at UC Berkeley, the QE is administered by a committee of four professors. The exam consists of writing and orally defending a thesis proposal and presenting a separate breadth requirement: answering general questions about Neuroscience in HWNI, answering questions about major papers in the field in MCB, and writing an "outside proposal" on a topic not related to your research in PMB.

While the QE is a crucial point in the scientific development of a graduate student, it can be very difficult to navigate for most students and can be particularly isolating for students from minoritized backgrounds (219). Further, students with differing levels of undergraduate preparation, receive varying levels of support from their graduate programs and advisors, which was the case for students who participated in this program. Because for most students the QE is drastically different from exams they experienced during their undergraduate studies, and for many students it is the first time writing a proposal of the required caliber, students are often left not knowing how to best prepare for their QE. Due to a lack of courses in most graduate programs dedicated to developing the skills needed to succeed in the QE, students often find it difficult to define milestones to measure their preparedness for the QE. The process is part of a hidden curriculum (220), where students often need to know to ask faculty about their expectations, making the process inequitable and more challenging for students from minoritized backgrounds. In addition, the COVID-19 pandemic had a negative effect on access to community

building and peer mentorship in recent graduate cohorts, compounding the difficulties described above. We identified the need for a program dedicated to developing the skills assessed during the QE, demystifying the QE, and providing structure with concrete milestones. This program would greatly benefit all graduate students but could be particularly advantageous for students from minoritized communities in academia.

Here, we describe the creation and assessment of the Inclusive Excellence in Quals Preparation program (IEQP) at University of California, Berkeley. The aim of the IEQP program is to prepare graduate students for their qualifying exam by establishing structured goals, building a sense of community, providing a support system, and creating an environment where students are encouraged to view themselves as scientists, and feel valued and capable of achieving their goals. To our knowledge, a program dedicated to support graduate students' preparation for the QE has not been previously reported.

4.2 Methods

The Inclusive Excellence in Quals Prep (IEQP) program was developed as a pilot program to support graduate students in their qualifying exam preparation process. The first iteration of the program was focused on second-year graduate students in the Plant and Microbial Biology (PMB) and Molecular and Cell Biology (MCB) programs at the University of California (UC) Berkeley.

4.2.1 Curriculum design

The program was designed by two graduate students, one in PMB (ZAA) and one in MCB (AGS), two faculty members, similarly from PMB (RB) and MCB (DB), and the director of the Office for Graduate Diversity (DS). Curriculum design drew largely from our own experiences in these two programs and graduate students' collective experiences.

The main goals of the IEQP program were to prepare graduate students for their qualifying exam by building a sense of community, providing a support system, and creating an environment where students feel valued and capable of achieving their goals. As such, the core principles addressed by the IEQP curriculum (Fig. 4.1) were academic preparation, community building, and mentorship.

Academic preparation was addressed by panels and workshops led by invited speakers (from UC Berkeley and elsewhere), which covered topics such as strategies to read research articles efficiently, crafting and giving an elevator pitch, and what to talk about when meeting with committee members. These topics were specifically addressed because they are often overlooked and not incorporated into graduate curricula despite their importance (academia's hidden curriculum) (221). Additionally, a subset of workshops was focused on wellness, resilience, and stress management. These topics were crucial to the IEQP curriculum because students frequently find QE preparation very stressful, and we wanted to provide them with tools to help address anxiety during this stressful time and throughout their graduate education.

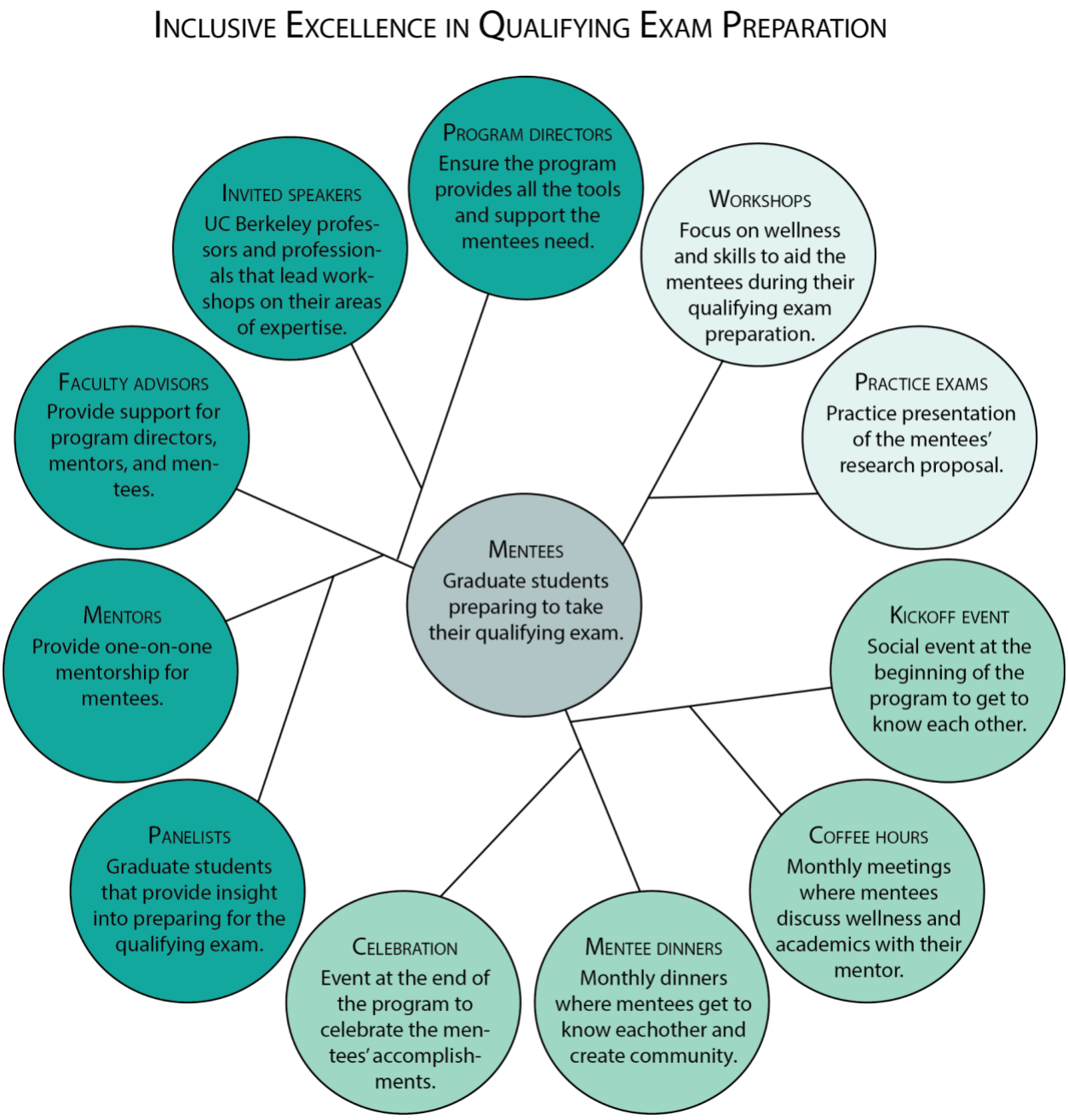


Figure 4.1. Curriculum map for the IEQP Program. The curriculum was designed around the core goals of academic preparation and community support. The academic preparation component consisted of workshops and practice exams, while the community support component consisted of coffee hours with mentors, monthly dinners, a kickoff event and a final celebration. Additionally, the IEQP community consisted of mentees, mentors, faculty advisors, and program directors.

We paired each student in the program with an older, post-QE graduate student mentor from the same program. Mentors formally engaged with their mentees through monthly one-on-one “coffee hours" where they discussed their progress and wellness. The program directors suggested guide topics and questions for each meeting that would match the mentees’ preparation stage. This peer mentoring allowed students to engage with older graduate students in their program who had shared interests and represented an additional source of support for each program participant. Additionally, mentors and mentees interacted informally during social events hosted by the IEQP program directors. These social events included monthly dinners, a kickoff event, and an end-of-program celebration. The events were meant to encourage mentees to build affinity as a cohort and provide opportunities for informal in-person interactions, which were missing during the COVID-19 lockdown.

Additionally, each IEQP mentee was required to have at least one practice exam in the month preceding their qualifying exam. This practice exam consisted of a mock examination led by trainees with expertise in each student’s research area. In most cases, practice exam participants were trainees in the laboratory of each student’s committee members. To decrease IEQP mentees’ workload, the program directors coordinated these practice exams.

A detailed curriculum containing lesson plans and full descriptions is included in the supplementary materials.

4.2.2 Timeline

Students in the target programs normally take their qualifying exam late in the Spring semester of their second year. For this reason, the program’s timeline was designed to support participants through this semester. The curriculum was designed during the Summer months, and invited speakers were contacted and confirmed between the Summer and Fall semesters. The program started officially in December 2021 with one introductory workshop, and weekly workshops were held from mid-January through the first week of April. The month of April was dedicated to practice exams.

4.2.3 Mentor and mentee recruitment and pairing

Graduate students were recruited via a general message sent by each program’s graduate advisor to the cohort of second-year students. Additionally, we encouraged students to apply to the program through one-on-one conversations. Students applied through a form that collected information regarding their identity and research interests. We received 13 applications (7 MCB, 4 PMB, 2 HWNI) and admitted 11 students to the pilot IEQP cohort. Criteria for selection were (1) availability during the semester and (2) self-proclaimed sense of preparedness for the exam (students that felt least prepared were prioritized). Care was taken to ensure that students belonging to underrepresented groups were represented in the mentee cohort.

We recruited potential mentors at the same time that we recruited mentees. The graduate student directors and faculty advisors personally reached out to graduate students in the PMB, MCB, and HWNI programs that were in their third year or above to invite them to participate as mentors in the program. We focused on inviting people with previous mentoring experience, and we ensured that students from underrepresented groups were represented in the mentor cohort. Prospective

mentors filled out an interest form where they detailed their research interests and indicated their availability for the Spring semester.

We received 18 mentor interest forms and carefully selected pairs of mentors and mentees with the following criteria in mind: (1) Mentor and mentee should be in the same program, (2) They should have similar or relevant research experience, (3) They should not be in the same lab to expand mentees' access to the graduate student community. Mentors were notified once they were matched to a mentee and they were required to attend an inclusive mentorship workshop before they officially interacted with their mentees. The inclusive mentorship workshop focused on helping mentors identify ways to connect with their mentees and make the QE experience one that enriches mentees' sense of belonging. Mentors were compensated for their participation in the program with a $1,000.00 stipend.

4.2.4 Program evaluation

We evaluated the program via a series of surveys, including a pre-program and end-of-program survey for mentees (Table 4.1), workshop surveys for mentees, and monthly check-in forms for mentors. We assessed the mentees self-perceived preparedness for the QE before and after completion of the IEQP program. The scale was 1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree.

4.3 Results

The program consisted of a cohort of 11 mentees from the PMB, MCB, and HWNI programs. We administered a survey at the beginning of the program (n=10) that evaluated the mentees' preparedness, wellbeing, wellness, and sense of community support. A second survey at the end of the program (n=11) assessed the progress of mentees and the program's impact.

Upon comparing results of the pre- and post-program surveys, there are some categories where the effects of general QE prep cannot be disentangled from the effects of the IEQP program. In general, students' QE preparedness and the support received from their advisors improved throughout the semester, as did their confidence in writing, critical reading, presentation, and communication skills. In some cases, the program may have had an effect on the observed improvements, but these are factors for which we would expect an increase due to time invested in QE preparation. We also observed an increase in metrics of sense of belonging in graduate school, and while some students attributed this to their participation in IEQP, the process of preparing for their QE and eventually passing it can contribute to feelings of belonging. We make this distinction throughout the results.

4.3.1 Demographics

One mentee identified as Black or African, four as Hispanic, Latinx, or Spanish Origin, two as Middle Eastern or North African, one as Native American or Alaska Native, two as South Asian, and three as White (Fig. 4.2A). One mentee identified as genderfluid or genderqueer, two identified as men, one as non-binary, and six as women (Fig. 4.2B). Four mentees reported being part of the LGBTQ+ community, four mentees reported having a physical, mental, or learning/cognitive disability, three mentees identified as international students, and four mentees

identified as first-generation students (Fig. 4.2C). Mentees' parents' education level was varied and no mentees reported having parents who had obtained a Ph.D. (Fig. 4.2D).

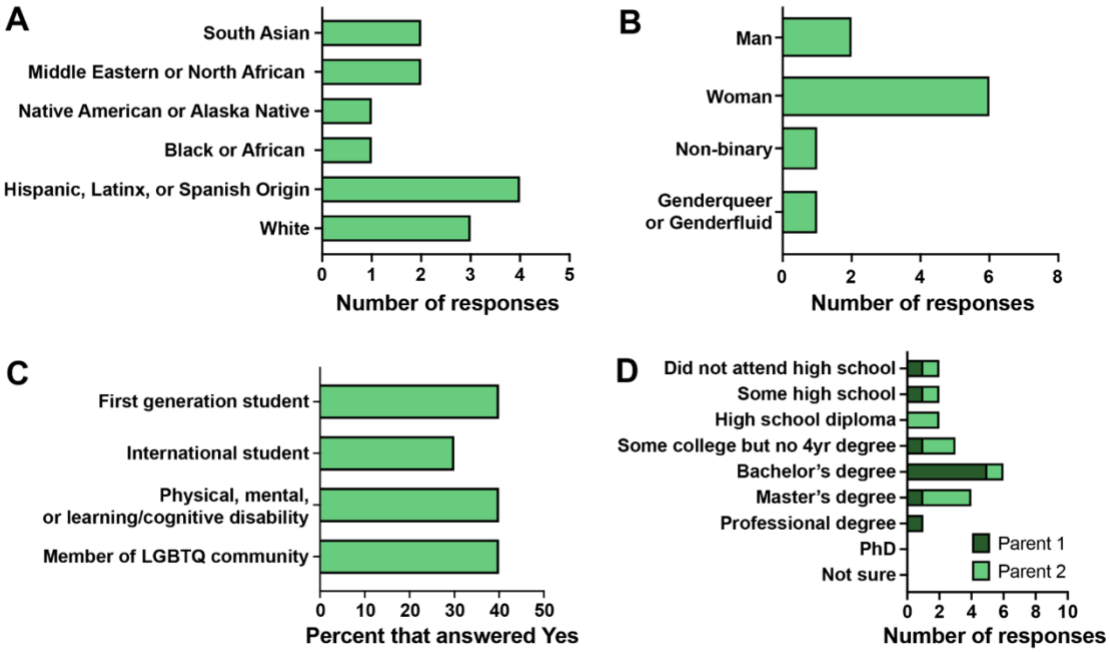


Figure 4.2. Program participant demographics. In a survey distributed before the start of the program, mentees reported (A) all races and ethnicities they identified with, (B) the gender they identified with, (C) groups they identified with, and (D) reported their parents' education level.

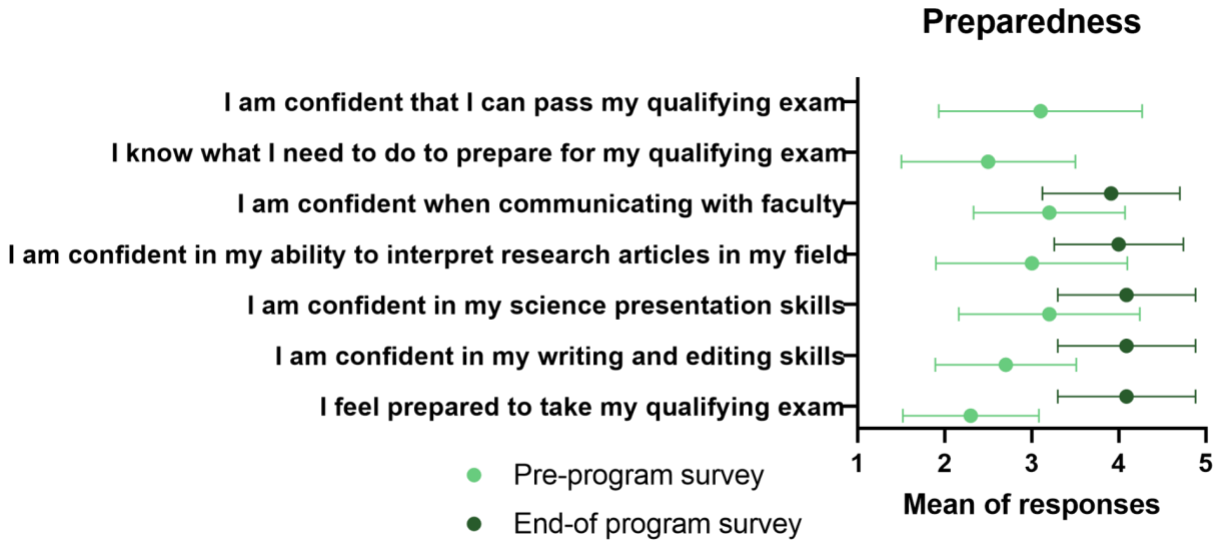


Figure 4.3. Academic preparedness. Mentees reported their self-perceived academic preparedness before the start of the program, light green, and after the program ended, dark green.

4.3.2 QE preparedness

Mentees reported their self-perceived preparation for the QE in a survey at the beginning and end of the program (Fig. 4.3). Nine mentees had written a research proposal before participating in the IEQP program. Overall, mentees disagreed that they knew what they needed to do to prepare for the QE (2.5 score) and felt neutral that they could pass the QE (3.1 score) before participating in the program.

We saw a drastic increase in agreement when comparing the following statements before and after the program: I feel prepared to take my qualifying exam, I am confident in my writing and editing skills, I am confident in my science presentation skills, I am confident in my ability to interpret research articles in my field, I am confident when communicating with faculty. Most mentees disagreed or felt neutral about these statements before the program and agreed after the program. Overall, the skills they honed throughout the workshops increased the mentees' self-perceived preparedness at the time of taking their QE. Still, additional preparation outside of the program likely also contributed to this effect.

4.3.3 Community support

We found that after the completion of the IEQP program, mentees had increased discussion of their proposal with their PI, felt their PI provided useful input on their proposal and felt increased agency over the content of their QE proposal (Fig. 4.4). We believe that weekly check-ins with mentees during the workshops likely increased the number of times the mentees met with their PI and thus increased their sense of agency over their proposal and their perceived usefulness of meetings with their PI. However, we cannot disentangle the effects of individual preparation for the QE on the increased agency over their proposal and discussions with their PI about the proposal. We found that mentees asked more members outside of UC Berkeley (friends, family) for help regarding the QE but asked fewer members of UC Berkeley for help regarding the QE (Fig 4.4). This is likely because the IEQP program already provided mentees with an extensive network of UC Berkeley members to help students with their QE. Additionally, mentees did not report an increase in asking lab members for help with the proposal, or discussions about topics other than the research proposal with lab members or PI (Fig. 4.4).

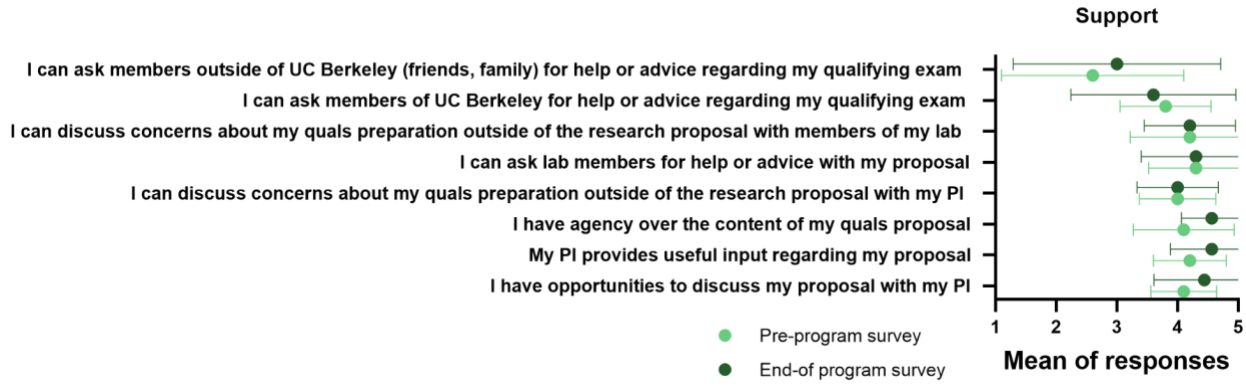


Figure 4.4. Support. Mentees reported the support received for qualifying exam preparation before the start of the program, light green, and after the program ended, dark green.

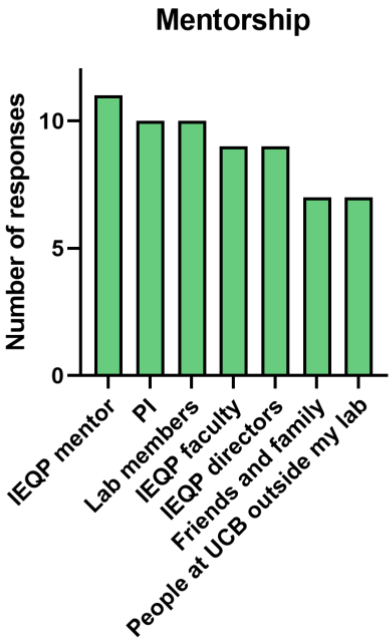


Figure 4.5. Mentorship figures. At the end of the program, mentees selected the mentors that they felt supported by during their qualifying exam preparation.

Two of the three participating departments provided a document with guidelines for the QE while one department did not provide students any guidelines for the QE. Eight mentees said their department organized a panel or seminar about the QE. However, mentees felt that they did not receive training similar to that covered in IEQP workshops in their program's curriculum. While the departments provided guidelines on the structure of the exam and a timeline of when documents need to be submitted, departments failed to provide guidelines for how to prepare for the exam.

When asked to rank sources of support at the end of the program, mentees reported feeling most supported by their IEQP mentor, followed by their PI and Lab members, followed by IEQP directors and faculty advisors, followed by friends and family, and people at UC Berkeley outside of their lab (Fig. 4.5).

4.3.4 Career Goals

When asked if the QE impacted their career goals, five mentees responded no, three responded yes, and three responded maybe. Some of the responses as to why the QE impacted their career goals:

"I felt that writing the proposal was difficult, so I am not sure if I want to continue in academia which requires writing grants a lot."

"It made me deeply consider how much I want to complete my PhD."

"[I] realized academia is a grind with too much work and too little pay; I don't want to do a postdoc. I'd rather go somewhere that values my time more highly."

"I don't think I want to feel this continuous scrutiny for the rest of my life…"

While the goal of the QE is to prepare students for graduate school, we found that it has a deep impact on students' self-perceived scientific identity. Unfortunately, students from minoritized backgrounds in academia may be more susceptible to the QE's impact on self-perceived scientific identity.

Four mentees said participating in IEQP impacted their career goals. One mentee wrote "IEQP has made me more confident about my place in STEM. Having people like me (latinX, minorities, etc) prepare and support me for the QE was very beneficial." Another mentee wrote "I realized that I don't actually love benchwork compared to reading the literature and coming up with ideas/testable hypotheses. I think I'd like a more social career that allows me to help mentor and manage other scientists, but without doing the primary bench work myself.

4.3.5 Wellness

We found that after the completion of the IEQP program mentees increased their feelings of belonging in graduate school, were better equipped to deal with stress, and felt they could overcome hardship (Fig. 4.6). However, mentees may have sought external help to deal with stress, further contributing to these results. We found that feelings that they can complete their PhD and that their degree is preparing them for the type of career that they want decreased slightly (Fig. 4.6). These decreases might be due to a change in career choice, as we saw that the QE affected some students' career choices.

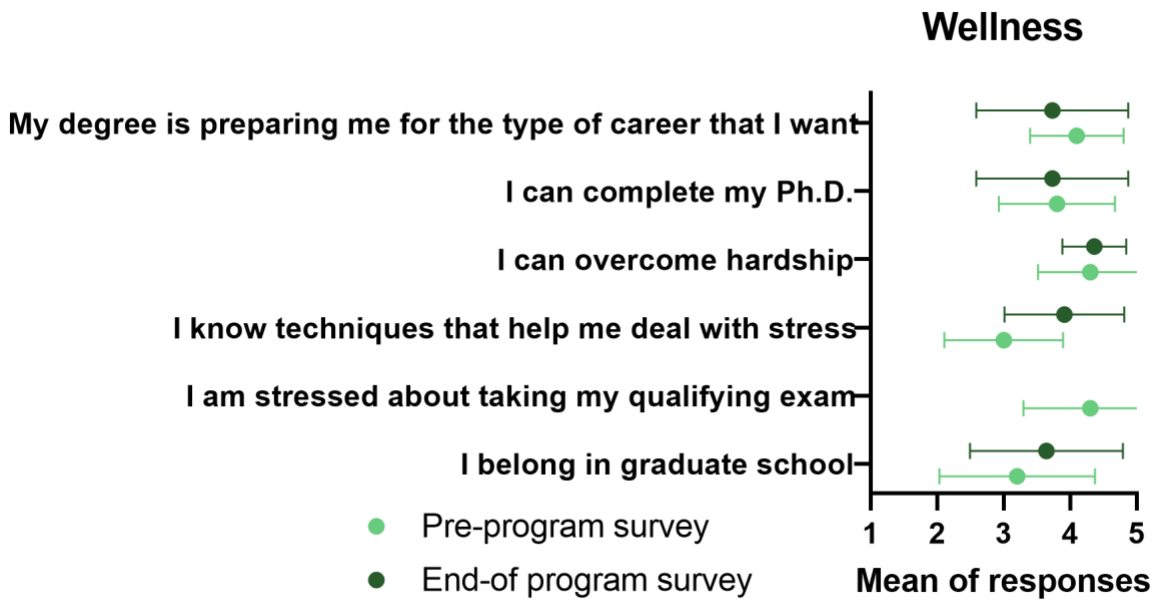


Figure 4.6. Wellness. Mentees reported their self-perceived wellness and belonging before the start of the program, light green, and after the program ended, dark green.

4.4 Discussion

We designed and implemented the Inclusive Excellence in Quals Prep program at the University of California, Berkeley. The pilot cohort consisted of eleven students, and we based our assessment of the program on their feedback through surveys and on our experience as program directors and faculty advisors. Ten out of eleven students succeeded at their QE on the first attempt and one student succeeded on their second attempt. Overall, we saw an increase in students' perception of preparedness, QE-related skills, the support received from their advisors, and the agency they felt over their proposed work. When asked about community support, 11/11 students stated feeling supported by their IEQP mentors. While mentees felt varied levels of support from their community, all mentees agreed that they felt supported by their IEQP mentor.

Our main takeaway is that peer mentorship was the most successful part of the Inclusive Excellence in Quals Prep program. The peer mentorship aspect of IEQP increased mentees' community in their programs, as many of them had not interacted much with older graduate students due to the COVID-19 pandemic. Having an older graduate student mentor provided mentees with support and guidance, an outlet for stress and anxiety, and an external source of feedback for their proposal, as well as someone to hold them accountable without judgement. Further, having a peer mentor contributed to tailored mentorship and an individualized approach where students developed plans for their QE prep that satisfied their needs, enabling them to thrive in the environment of semi-standardized tests. After this first iteration of IEQP, we recommend that peer mentors are retained as a crucial part of the program. This program component provided significant support for mentees and allowed more advanced graduate students to practice their mentorship skills and get compensated for their work. If we had to choose only one program component to retain, it would be peer mentorship. We further suggest that peer mentorship should be integrated broadly across graduate programs as an essential facet of retention in STEM.

While the participating graduate programs provide some training on soft skills, such as a class on giving scientific talks, the breadth of training provided by IEQP through academic-focused workshops is not provided by these programs' curricula. Students are generally expected to have these skills coming into their graduate program or acquire these skills indirectly. However, teaching them in a structured, hands-on manner proved to be beneficial for students. The skills taught through IEQP workshops are crucial at all stages of a graduate career and should be taught as early as possible in graduate programs to provide a strong foundation that students can continue to build upon. We suggest skills such as giving a chalk talk, giving an elevator pitch, skimming papers efficiently, and stress management be integrated into the first-year curriculum or taught as part of a structured program for first-year students.

Students expressed that they would have enjoyed having more structured 'group study time' to allow them to work on their proposals. By transitioning workshops to be part of the first-year curriculum, IEQP can be more focused on providing students with a structured, guided space to work towards their QE preparation goals with their peers, while building community and fostering accountability. By addressing core skills during graduate students' first year, we recommend only maintaining workshops that are QE specific in future iterations of IEQP, such as how to effectively communicate with committee members, planning a quals study strategy,

and proposal specific chalk talk and elevator pitch practices. Further, we recommend that these workshops be led by the IEQP community: program directors, faculty advisors, or mentors to not disrupt the process of community building by bringing in external people each week.

Finally, while some mentees relied heavily on IEQP as their main source of mentorship for QE prep, others had advisors that were very involved in the process. In the future, we recommend more communication between each mentee, their PI and the program directors in order to determine the best ways to support each student to make QE prep more equitable. This can be addressed by communicating with PIs at the beginning of the year to inform them that their mentee is a program participant and sharing the program curriculum and expected timeline. Additionally, communication channels should be open throughout the duration of the program.

When asked whether the QE impacted their career goals, three mentees answered yes and 3 answered maybe. They expressed reconsidering whether they want to complete their PhD and whether they want to stay in academia. While the QE provides an opportunity for scientific development, the high rigor and scrutiny associated with the QE, compounded with a lack of structure and poorly stated expectations can have a negative effect on students' confidence and scientific identity. The QE, as it is employed in many programs, represents an antiquated way to measure student's success and preparation, as it seems that it is meant to filter students out rather than prepare them for an independent research project (222). Graduate academic programs should evaluate if their QE format is designed to achieve the stated goal of preparing students to undertake an independent research project (223). Additionally, if programs do not provide the tools and training required for student's QE success, each student is dependent on their advisor's and committee's involvement in their QE preparation, which does not ensure an equitable experience (224). Providing the proper training to encourage student QE success represents one of many necessary steps in creating more accessible and equitable environments in STEM PhD programs.

In conclusion, the Inclusive Excellence in Quals Prep program was successful in its first year. A small cohort size that spanned three departments provided the opportunity to pilot the program and collect sufficient feedback to inform our recommendations for future iterations of the program and for implementation by others. The IEQP program provided students with the necessary structure, skill development, and mentorship during their QE prep. We hope that with the establishment of this model program at UC Berkeley, graduate programs will incorporate similar training into their curriculum to support graduate students throughout their QE preparation and ensure their success in graduate school. Standardizing programs such as IEQP can have the indirect effect of signaling to students from minoritized groups that critical support structures are available, thus contributing as a recruitment and retention tool.

Table 4.1. Questions in the pre-program and end-of-program surveys.

No.	Category	Pre-program survey questions and answer choices	End-of-program survey questions and answer choices
1	Self-identification	Name	
2	Self-identification	Which category best describes you? (1) White; (2) Hispanic/Latinx; (3) Black or African; (4) Asian; (5) Native American or Alaska Native; (6) Middle Eastern or North African; (7) Pacific Islander; (8) Not listed above, please specify	
3	Self-identification	What is your gender identity? (1) Man; (2) Woman; (3) Trans man; (4) Trans woman; (5) Genderqueer/Nonbinary Gender; (6) Not listed above, please specify; (7) Prefer not to say	
4	Self-identification	Are you a member of the LGBTQ+ community?	
5	Self-identification	Are you a first-generation student?	
6	Self-identification	Are you a student with a disability?	
7	Self-identification	Are you an international student?	
8	Self-identification	What are your parents' or caretakers' highest educational degree? (1) PhD, (2) Professional degree, (3) Masters' degree, (4) Bachelor's degree, (5) High school diploma, (6) Some high school, (7) Did not attend high school	
9	QE preparedness	Have you written a research proposal before? (1) Yes, (2) No	Did you pass your qualifying exam? (1) Yes, (2) No, (3) I haven't taken my qualifying exam yet
For questions 10 - 26: Indicate agreement or disagreement with each of the following statements. (1) Strongly disagree, (2) Disagree, (3) Neutral, (4) Agree, (5) Strongly agree			
10	QE preparedness	I feel prepared to take my qualifying exam	At the time of taking my qualifying exam, I felt prepared

11	QE preparedness	I know what I need to do to prepare for my qualifying exam	I am confident in my writing and editing skills
12	QE preparedness	I am confident that I can pass my qualifying exam	I am confident in my science presentation skills
13	QE preparedness	I am confident in my writing and editing skills	I am confident in my ability to interpret research articles in my field
14	QE preparedness	I am confident in my science presentation skills	I am confident when communicating with faculty
15	QE preparedness	I am confident in my ability to interpret research articles in my field	
16	QE preparedness	I am confident when communicating with faculty	
17	Community support	I can discuss my proposal with my PI	I frequently discussed my proposal with my PI
18	Community support	My PI provides useful input regarding my proposal	My PI provided useful input regarding my proposal
19	Community support	I have agency over the content of my quals proposal	I had agency over the content of my quals proposal
20	Community support	I can discuss concerns about my quals preparation outside of the research proposal with my PI	I discussed concerns about my quals preparation outside of the research proposal with my PI
21	Community support	I can ask lab members for help or advice with my proposal	I asked lab members for help or advice with my proposal
22	Community support	I can discuss concerns about my quals preparation outside of the research proposal with members of my lab	I discussed concerns about my quals preparation outside of the research proposal with members of my lab
23	Community support	I can ask members of UC Berkeley for help or advice regarding my qualifying exam	I asked members of UC Berkeley for help or advice regarding my qualifying exam
24	Community support	I can ask members outside of UC Berkeley (friends, family) for help or advice regarding my qualifying exam	I frequently asked members outside of UC Berkeley (friends, family) for help or advice regarding my qualifying exam
25	Community support	Indicate what type of support you have received from your department. Select all that apply: (1) My department provided a detailed document for the qualifying exam, (2) My department provided a workshop/panel/seminar about the	I received training similar to that covered in IEQP workshops in my program's curriculum

		qualifying exam, (3) My department didn't provide any guidance	
26	Community support		I felt supported in my quals prep by the following people: (1) PI, (2) Lab members, (3) People at UC Berkeley outside my lab, (4) Friends/family, (5) My IEQP mentor, (6) IEQP faculty advisors, (7) IEQP directors
For questions 27-32: Indicate agreement or disagreement with each of the following statements. (1) Strongly disagree, (2) Disagree, (3) Neutral, (4) Agree, (5) Strongly agree			
27	Wellness, belonging, and stress management	I belong in graduate school	I belong in graduate school
28	Wellness, belonging, and stress management	I am stressed about taking my qualifying exam	I know techniques that help me deal with stress
29	Wellness, belonging, and stress management	I know techniques that help me deal with stress	I can overcome hardship
30	Wellness, belonging, and stress management	I can overcome hardship	
31	Career goals	I can complete my Ph.D.	I can complete my Ph.D.
32	Career goals	My degree is preparing me for the type of career that I want.	My degree is preparing me for the type of career that I want.
For questions 33-38: Select one, or complete with short answer			
33	Career goals	Indicate what type of career you wish to pursue (1) education: tenure track faculty, non-tenure track faculty, K-12 teacher; (2) non-faculty research: scientist in national lab, academic, industry setting; technical specialist; (3) other science: science communication, science outreach, science policy, patent lawyer, medical professional; (4) other: finance, accounting, marketing, sales,	Indicate what type of career you wish to pursue

		musician, self-employed; (5) undecided	
34	Career goals	What are your career goals? short answer	What are your career goals? short answer
35	Career goals		Did participating in IEQP impact your career goals? Yes, No, Maybe
36	Career goals		If you answered yes to the above question, explain why: short answer
37	Career goals		Did the quals process impact your career goals? Yes, No, Maybe
38	Career goals		If you answered yes to the above question, explain why: short answer
For questions 39-48: Select all that apply: (1) Workshops; (2) Practice exams; (3) Community social events; (4) Interacting with mentor one-on-one; (5) Meeting with Rachel and Diana; (6) Other			
39	Program evaluation		Which program elements were most effective in terms of preparation for the experience of the oral exam?
40	Program evaluation		Which program elements were least effective in terms of preparation for the experience of the oral exam?
41	Program evaluation		Which program elements were the most effective in terms of writing the inside proposal?
42	Program evaluation		Which program elements were the least effective in terms of writing the inside proposal?
43	Program evaluation		Which program elements were the most effective in terms of stress management?
44	Program evaluation		Which program elements were the least effective in terms of stress management?
45	Program evaluation		Did your interactions with your peer mentor add a lot to your program experience, and how could they improve? Short answer
46	Program evaluation		Did your interactions with faculty directors (Rachel and Diana) add a lot to your program experience, and

			how could they improve? Short answer
47	Program evaluation		What was the hardest part of quals prep for you and how could the program have helped you meet this challenge better? Short answer
48	Program evaluation		Is there anything you would change in the program going forward? Short answer

Acknowledgements
We thank all students for their participation as mentors or mentees, Lyn Rivera and Hannah Bloom for help recruiting mentees, and UC Berkeley's Office of Graduate Diversity for administrative support. Jeff Royal provided instrumental support with survey development. Additionally, we acknowledge Dr. Nancy Liu, Dr. Marla Feller, Dr. Amy Honigman, Dr. Sheila McCormick, Dr. Brett Mensh, and Dr. Britt Glaunsinger for facilitating workshops in their areas of expertise, and the graduate students and postdocs who participated as panelists and provided their support for practice exams. We thank Dr. Michi Taga for critical reading of the manuscript. The IEQP program was funded by UC Berkeley's Office of Graduate Diversity.

Concluding remarks

Since the discovery of vitamin B_{12} and its importance for human health in 1926 (8–10), and the subsequent discovery of its requirement in bacterial metabolism (225, 226), the field of corrinoid biology as it relates to microbes has continued to expand. To date, 17 corrinoids varying in their lower ligand structure have been found in many environments (14, 19, 34, 38). These nutrients are produced by bacteria and archaea, and they impact the metabolism and growth of many microbes (39). Further, corrinoids have been shown to be shared in a number of bacteria-bacteria and bacteria-algae co-cultures (24–33, 100). The application of corrinoids as a model shared nutrient presents a unique approach for advancing the study of microbial interactions in complex communities (7, Chapter 1).

My dissertation work is part of the first use of the model nutrient approach applied to soil microbial communities. Soil's physical and chemical complexity, coupled with its taxonomic diversity, makes it one of the hardest environments in which to elucidate the roles of individual microbes and their impact on community structure and function (145). Focusing on a model nutrient is particularly useful for soil microbiology research, as it allows us as scientists to "zoom in" and study interactions in a reductionist manner that allows for the mechanistic dissection of processes.

This research spanned the organismal scale described in Chapter 1, and expanded into binary and higher order interactions, the more granular components of the community scale. In sum, through the isolation of bacteria from a California grassland soil and evaluation of their interactions – both competitive and cooperative – this work begins to elucidate how environmental corrinoid availability impacts individual soil bacteria. Two components of the bacterial isolation in Chapter 2 enabled the interaction studies in Chapter 3. First, bacteria were isolated from the same environment, which highlights the possibility that these bacteria interact in their native environment. Second, isolating bacteria on one defined medium served as a great primer for interaction studies because it allows for easy removal or addition of specific compounds. Another highlight of this work is the refinement of the "producer," "dependent," and "independent" categories, as well as definition of the "provider" term. These categories, defined experimentally and genomically, have established a language that facilitates prediction and testing of corrinoid-based interactions in culture and from genomic data.

By evaluating corrinoid-based interactions among a set of isolates, we were able to test the predictability of competitive interactions between dependents and measure corrinoid sharing between producers and dependents. While some interactions were predictable from monoculture growth, others were not, possibly due to emergent properties arising in co- and tri-cultures. Genome analysis suggests that corrinoids are the main shared nutrient between the strains tested, meaning that the consortia established here represent an effective system in which to study additional aspects of corrinoid-based interactions.

Altogether, I built and characterized a collection of 161 bacterial isolates that I then leveraged to research interactions in a quantifiable and hypothesis-driven manner by applying the model nutrient approach. Future research may expand to more combinations of isolates or types of interactions and refine predictions by considering emergent properties that arise when isolates are combined. Genetic manipulation of a subset of bacteria in the collection may be achievable,

thus expanding the range of possible experiments and paving the way for an increasingly detailed understanding of interactions. Overarching questions can be addressed by researching the mechanisms of corrinoid sharing, symbiotic relationships that may arise when corrinoids are shared, and how pairwise interactions based on vitamin sharing influence overall community structure. Parallel research found that addition of corrinoids can impact the structure of soil-derived communities (112). Studying the effects of corrinoids on bacterial isolates can help us understand how individual strains may shape corrinoid effects at the community scale. Future work on the corrinoid model should also include environments with higher corrinoid diversity, such as the human gut, as soil contains 95% vitamin B_{12} (19). This makes it one of the least diverse environments in terms of corrinoid structures, in stark contrast to its taxonomic diversity.

Many ecological and mechanistic questions remain regarding interactions in complex communities. Detailed approaches such as the model nutrient approach, combined with large scale -omics and modeling approaches can help to answer many of these questions over the coming years. I look forward to future research that will uncover many secrets about how microbial communities are structured and how they contribute to our changing world.

References

1. D'Souza G, Waschina S, Pande S, Bohl K, Kaleta C, Kost C. 2014. Less is more: selective advantages can explain the prevalent loss of biosynthetic genes in bacteria. Evol Int J Org Evol 68:2559–2570.

2. Johnson WM, Alexander H, Bier RL, Miller DR, Muscarella ME, Pitz KJ, Smith H. 2020. Auxotrophic interactions: a stabilizing attribute of aquatic microbial communities? FEMS Microbiol Ecol 96:fiaa115.

3. Mee MT, Collins JJ, Church GM, Wang HH. 2014. Syntrophic exchange in synthetic microbial communities. Proc Natl Acad Sci U S A 111:E2149-2156.

4. Stubbendieck RM, Vargas-Bautista C, Straight PD. 2016. Bacterial communities: Interactions to scale. Front Microbiol 7:1234.

5. Abreu NA, Taga ME. 2016. Decoding molecular interactions in microbial communities. FEMS Microbiol Rev 40:648–663.

6. Seth EC, Taga ME. 2014. Nutrient cross-feeding in the microbial world. Front Microbiol 5:350.

7. Sokolovskaya OM, Shelton AN, Taga ME. 2020. Sharing vitamins: Cobamides unveil microbial interactions. Science 369:eaba0165.

8. Minot GR, Murphy WP. 1926. Treatment of pernicious anemia by a special diet. Yale J Biol Med 74:341–353.

9. Scott JM, Molloy AM. 2012. The discovery of vitamin B_{12}. Ann Nutr Metab 61:239–245.

10. Hodgkin DC, Kamper J, Mackay M, Pickworth J, Trueblood KN, White JG. 1956. Structure of Vitamin B_{12}. Nature 178:64–66.

11. Roth JR, Lawrence JG, Bobik TA. 1996. COBALAMIN (COENZYME B_{12}): Synthesis and biological significance. Annu Rev Microbiol 50:137–181.

12. Herbert V. 1988. Vitamin B_{12}: plant sources, requirements, and assay. Am J Clin Nutr 48:852–858.

13. Martens J-H, Barg H, Warren M, Jahn D. 2002. Microbial production of vitamin B_{12}. Appl Microbiol Biotechnol 58:275–285.

14. Renz P. 1999. Biosynthesis of the 5, 6-dimethylbenzimidazole moiety of cobalamin and of the other bases found in natural corrinoids., p. 960. *In* Banerjee, R (ed.), Chemistry and Biochemistry of B_{12}. John Wiley & Sons, Hoboken, NJ, USA.

15. Banerjee R. 2003. Radical Carbon Skeleton Rearrangements: Catalysis by coenzyme B_{12}-dependent mutases. Chem Rev 103:2083–2094.

16. Bridwell-Rabb J, Drennan CL. 2017. Vitamin B_{12} in the spotlight again. Curr Opin Chem Biol 37:63–70.

17. Bridwell-Rabb J, Li B, Drennan CL. 2022. Cobalamin-dependent radical S-adenosylmethionine enzymes: Capitalizing on old motifs for new functions. ACS Bio Med Chem Au 2:173–186.

18. Neil E, Marsh G. 1999. Coenzyme B_{12} (cobalamin)-dependent enzymes. Essays Biochem 34:139–154.

19. Hallberg ZF, Seth EC, Thevasundaram K, Taga ME. 2022. Comparative analysis of corrinoid profiles across host-associated and environmental samples. Biochemistry 61:2791–2796.

20. Mok KC, Hallberg ZF, Taga ME. 2022. Chapter Three - Purification and detection of vitamin B_{12} analogs, p. 61–85. *In* Marsh, ENG (ed.), Methods in Enzymology. Academic Press.

21. Gómez-Consarnau L, Sachdeva R, Gifford SM, Cutter LS, Fuhrman JA, Sañudo-Wilhelmy SA, Moran MA. 2018. Mosaic patterns of B-vitamin synthesis and utilization in a natural marine microbial community. Environ Microbiol 20:2809–2823.

22. Magnúsdóttir S, Ravcheev D, de Crécy-Lagard V, Thiele I. 2015. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. Front Genet 6:148.

23. Shelton AN, Seth EC, Mok KC, Han AW, Jackson SN, Haft DR, Taga ME. 2019. Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. ISME J 13:789–804.

24. Belzer C, Chia LW, Aalvink S, Chamlagain B, Piironen V, Knol J, de Vos WM. 2017. Microbial metabolic networks at the mucus layer lead to diet-independent butyrate and vitamin B_{12} production by intestinal symbionts. mBio 8:10.1128/mbio.00770-17.

25. Cooper MB, Kazamia E, Helliwell KE, Kudahl UJ, Sayer A, Wheeler GL, Smith AG. 2018. Cross-exchange of B-vitamins underpins a mutualistic interaction between *Ostreococcus tauri* and *Dinoroseobacter shibae*. ISME J 13:334–345.

26. Gude S, Pherribo GJ, Taga ME. 2022. A salvaging strategy enables stable metabolite provisioning among free-living bacteria. mSystems 7:e00288-22.

27. Iguchi H, Yurimoto H, Sakai Y. 2011. Stimulation of methanotrophic growth in cocultures by cobalamin excreted by Rhizobia. Appl Environ Microbiol 77:8509–8515.

28. Kazamia E, Czesnick H, Nguyen TTV, Croft MT, Sherwood E, Sasso S, Hodson SJ, Warren MJ, Smith AG. 2012. Mutualistic interactions between vitamin B_{12} -dependent algae and heterotrophic bacteria exhibit regulation. Environ Microbiol 14:1466–1476.

29. Ma AT, Beld J, Brahamsha B. 2017. An Amoebal Grazer of Cyanobacteria Requires Cobalamin Produced by Heterotrophic Bacteria. Appl Environ Microbiol 83:e00035-17.

30. Men Y, Seth EC, Yi S, Allen RH, Taga ME, Alvarez-Cohen L. 2014. Sustainable growth of *Dehalococcoides mccartyi* 195 by corrinoid salvaging and remodeling in defined lactate-fermenting consortia. Appl Environ Microbiol 80:2133–2141.

31. Swaney MH, Henriquez N, Campbell T, Handelsman J, Kalan LR. 2024. Skin-associated *Corynebacterium amycolatum* shares cobamides. bioRxiv https://doi.org/10.1101/2024.04.28.591522.

32. Wienhausen G, Moraru C, Bruns S, Tran DQ, Sultana S, Wilkes H, Dlugosch L, Azam F, Simon M. 2024. Ligand cross-feeding resolves bacterial vitamin B_{12} auxotrophies. Nature 629:886–892.

33. Yan J, Ritalahti KM, Wagner DD, Löffler FE. 2012. Unexpected specificity of interspecies Cobamide transfer from *Geobacter* spp. to organohalide-respiring *Dehalococcoides mccartyi* strains. Appl Environ Microbiol 78:6630–6636.

34. Ford JE, Hutner SH. 1955. Role of vitamin B_{12} in the metabolism of microorganisms, p. 101–136. *In* Harris, RS, Marrian, GF, Thimann, KV (eds.), Vitamins & Hormones. Academic Press.

35. Mok KC, Taga ME. 2013. Growth inhibition of *Sporomusa ovata* by incorporation of benzimidazole bases into cobamides. J Bacteriol 195:1902.

36. Sokolovskaya OM, Mok KC, Park JD, Tran JLA, Quanstrom KA, Taga ME. 2019. Cofactor selectivity in methylmalonyl coenzyme A mutase, a model cobamide-dependent enzyme. mBio 10:e01303-19.

37. Sokolovskaya OM, Plessl T, Bailey H, Mackinnon S, Baumgartner MR, Yue WW, Froese DS, Taga ME. 2021. Naturally occurring cobalamin (B_{12}) analogs can function as cofactors for human methylmalonyl-CoA mutase. Biochimie 183:35–43.

38. Yan J, Bi M, Bourdon AK, Farmer AT, Wang P-H, Molenda O, Quaile AT, Jiang N, Yang Y, Yin Y, Şimşir B, Campagna SR, Edwards EA, Löffler FE. 2018. Purinyl-cobamide is a native prosthetic group of reductive dehalogenases. Nat Chem Biol 14:8–14.

39. Mok KC, Sokolovskaya OM, Deutschbauer AM, Carlson HK, Taga ME. 2024. Microbes display broad diversity in cobamide preferences. bioRxiv https://doi.org/10.1101/2024.11.04.621602.

40. Mok KC, Sokolovskaya OM, Nicolas AM, Hallberg ZF, Deutschbauer A, Carlson HK, Taga ME. 2020. Identification of a novel cobamide remodeling enzyme in the beneficial human gut bacterium *Akkermansia muciniphila*. mBio 11:10.1128/mbio.02507-20.

41. Miles ZD, McCarty RM, Molnar G, Bandarian, V. 2011. Discovery of epoxyqueuosine (oQ) reductase reveals parallels between halorespiration and tRNA modification. Proc Natl Acad Sci 108:7368–7372.

42. Benjdia A, Berteau O. 2023. B_{12}-dependent radical SAM enzymes: Ever expanding structural and mechanistic diversity. Curr Opin Struct Biol 83:102725.

43. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE. 2001. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. Nucleic Acids Res 29:1097–1106.

44. Lengyel P, Mazumder R, Ochoa S. 1960. Mammalian methylmalonyl isomerase and vitamin B_{12} coenzymes*. Proc Natl Acad Sci 46:1312–1318.

45. Poppe L, Stupperich E, Hull WE, Buckel T, Rétey J. 1997. A base-off analogue of coenzyme-B_{12} with a modified nucleotide loop. Eur J Biochem 250:303–307.

46. Barker HA, Weissbach H, Smyth RD. 1958. A coenzyme containing pseudovitamin B_{12}*. Proc Natl Acad Sci 44:1093–1097.

47. Kennedy KJ, Widner FJ, Sokolovskaya OM, Innocent LV, Procknow RR, Mok KC, Taga ME. 2022. Cobalamin riboswitches are broadly sensitive to corrinoid cofactors to enable an efficient gene regulatory strategy. mBio 13:e01121-22.

48. Poppe L, Bothe H, Bröker G, Buckel W, Stupperich E, Rétey J. 2000. Elucidation of the coenzyme binding mode of further B_{12}-dependent enzymes using a base-off analogue of coenzyme B_{12}. J Mol Catal B Enzym 10:345–350.

49. Tanioka Y, Miyamoto E, Yabuta Y, Ohnishi K, Fujita T, Yamaji R, Misono H, Shigeoka S, Nakano Y, Inui H, Watanabe F. 2010. Methyladeninylcobamide functions as the cofactor of methionine synthase in a Cyanobacterium, *Spirulina platensis* NIES-39. FEBS Lett 584:3223–3226.

50. Warren MJ, Raux E, Schubert HL, Escalante-Semerena JC. 2002. The biosynthesis of adenosylcobalamin (vitamin B_{12}). Nat Prod Rep 19:390–412.

51. Brindley AA, Raux E, Leech HK, Schubert HL, Warren MJ. 2003. A story of chelatase evolution: Identification and characterization of a small 13-15kDa "ancestral" cobaltochelatase (CbiXS) in the archaea*. J Biol Chem 278:22388–22395.

52. DiMarco AA, Bobik TA, Wolfe RS. 1990. Unusual coenzymes of methanogenesis. Annu Rev Biochem 59:355–394.

53. Thomas MG, Escalante-Semerena JC. 2000. Identification of an alternative nucleoside triphosphate: 5′-deoxyadenosylcobinamide phosphate nucleotidyltransferase in *Methanobacterium thermoautotrophicum* ΔH. J Bacteriol 182:4227–4233.

54. Woodson JD, Escalante-Semerena JC. 2006. The *cbiS* gene of the archaeon *Methanopyrus kandleri* AV19 encodes a bifunctional enzyme with adenosylcobinamide amidohydrolase and α-ribazole-phosphate phosphatase activities. J Bacteriol 188:4227–4235.

55. Woodson JD, Peck RF, Krebs MP, Escalante-Semerena JC. 2003. The *cobY* gene of the archaeon *Halobacterium* sp. strain NRC-1 is required for *de novo* cobamide synthesis. J Bacteriol 185:311–316.

56. Campbell GRO, Taga ME, Mistry K, Lloret J, Anderson PJ, Roth JR, Walker GC. 2006. *Sinorhizobium meliloti bluB* is necessary for production of 5,6-dimethylbenzimidazole, the lower ligand of B_{12}. Proc Natl Acad Sci 103:4634–4639.

57. Gray MJ, Escalante-Semerena JC. 2007. Single-enzyme conversion of FMNH2 to 5,6-dimethylbenzimidazole, the lower ligand of B_{12}. Proc Natl Acad Sci U S A 104:2921–2926.

58. Taga ME, Larsen NA, Howard-Jones AR, Walsh CT, Walker GC. 2007. BluB cannibalizes flavin to form the lower ligand of vitamin B_{12}. Nature 446:449–453.

59. Hazra AB, Han AW, Mehta AP, Mok KC, Osadchiy V, Begley TP, Taga ME. 2015. Anaerobic biosynthesis of the lower ligand of vitamin B_{12}. Proc Natl Acad Sci 112:10792–10797.

60. Mehta AP, Abdelwahed SH, Fenwick MK, Hazra AB, Taga ME, Zhang Y, Ealick SE, Begley TP. 2015. Anaerobic 5-hydroxybenzimidazole formation from aminoimidazole ribotide: An unanticipated intersection of thiamin and vitamin B_{12} biosynthesis. J Am Chem Soc 137:10444–10447.

61. Villa EA, Escalante-Semerena JC. 2024. Corrinoid salvaging and cobamide remodeling in bacteria and archaea. J Bacteriol e00286-24.

62. Butzin NC, Secinaro MA, Swithers KS, Gogarten JP, Noll KM. 2013. *Thermotoga lettingae* can salvage cobinamide to synthesize vitamin B_{12}. Appl Environ Microbiol 79:7006–7012.

63. Di Girolamo PM, Bradbeer C. 1971. Transport of Vitamin B_{12} in *Escherichia coli*. J Bacteriol 106:745–750.

64. Yi S, Seth EC, Men Y-J, Stabler SP, Allen RH, Alvarez-Cohen L, Taga ME. 2012. Versatility in corrinoid salvaging and remodeling pathways supports corrinoid-dependent metabolism in *Dehalococcoides mccartyi*. Appl Environ Microbiol 78:7745–52.

65. Gray MJ, Escalante-Semerena JC. 2010. A new pathway for the synthesis of α-ribazole-phosphate in *Listeria innocua*. Mol Microbiol 77:1429–1438.

66. Allen RH, Stabler SP. 2008. Identification and quantitation of cobalamin and cobalamin analogues in human feces. Am J Clin Nutr 87:1324–1335.

67. Girard CL, Santschi DE, Stabler SP, Allen RH. 2009. Apparent ruminal synthesis and intestinal disappearance of vitamin B_{12} and its analogs in dairy cows. J Dairy Sci 92:4524–4529.

68. Degnan PH, Barry NA, Mok KC, Taga ME, Goodman AL. 2014. Human gut microbes use multiple transporters to distinguish vitamin B_{12} analogs and compete in the gut. Cell Host Microbe 15:47–57.

69. Cadieux N, Bradbeer C, Reeger-Schneider E, Köster W, Mohanty AK, Wiener MC, Kadner RJ. 2002. Identification of the periplasmic cobalamin-binding protein BtuF of *Escherichia coli*. J Bacteriol 184:706–717.

70. DeVeaux LC, Kadner RJ. 1985. Transport of vitamin B_{12} in *Escherichia coli*: Cloning of the *btuCD* region. J Bacteriol 162:888–896.

71. Putnam EE, Goodman AL. 2020. B vitamin acquisition by gut commensal bacteria. PLOS Pathog 16:e1008208.

72. Bradbeer C, Gudmundsdottir A. 1990. Interdependence of calcium and cobalamin binding by wild-type and mutant BtuB protein in the outer membrane of *Escherichia coli.* J Bacteriol 172:4919–4926.

73. Chimento DP, Kadner RJ, Wiener MC. 2003. The *Escherichia coli* outer membrane cobalamin transporter BtuB: Structural analysis of calcium and substrate binding, and identification of orthologous transporters by sequence/structure conservation. J Mol Biol 332:999–1014.

74. Santos JA, Rempel S, Mous ST, Pereira CT, ter Beek J, de Gier J-W, Guskov A, Slotboom DJ. 2018. Functional and structural characterization of an ECF-type ABC transporter for vitamin B_{12}. eLife 7:e35828.

75. Nijland M, Lefebvre SN, Thangaratnarajah C, Slotboom DJ. 2024. Bidirectional ATP-driven transport of cobalamin by the mycobacterial ABC transporter BacA. Nat Commun 15:2626.

76. Rempel S, Gati C, Nijland M, Thangaratnarajah C, Karyolaimos A, de Gier JW, Guskov A, Slotboom DJ. 2020. A mycobacterial ABC transporter mediates the uptake of hydrophilic compounds. Nature 580:409–412.

77. Abellon-Ruiz J, Jana K, Silale A, Frey AM, Baslé A, Trost M, Kleinekathöfer U, van den Berg B. 2023. BtuB TonB-dependent transporters and BtuG surface lipoproteins form stable complexes for vitamin B_{12} uptake in gut *Bacteroides*. Nat Commun 14:4714.

78. Wexler AG, Schofield WB, Degnan PH, Folta-Stogniew E, Barry NA, Goodman AL. 2018. Human gut *Bacteroides* capture vitamin B_{12} via cell surface-exposed lipoproteins. eLife 7:e37138.

79. Putnam EE, Abellon-Ruiz J, Killinger BJ, Rosnow JJ, Wexler AG, Folta-Stogniew E, Wright AT, van den Berg B, Goodman AL. 2022. Gut commensal Bacteroidetes encode a novel class of vitamin B_{12}-binding proteins. mBio 13:e02845-21.

80. Rempel S, Colucci E, de Gier JW, Guskov A, Slotboom DJ. 2018. Cysteine-mediated decyanation of vitamin B_{12} by the predicted membrane transporter BtuM. Nat Commun 9:3038.

81. Weissbach H, Toohey J, Barker HA. 1959. Isolation and properties of B_{12} coenzymes containing benzimidazole or dimethylbenzimidazole*. Proc Natl Acad Sci 45:521–525.

82. Lu X, Heal KR, Ingalls AE, Doxey AC, Neufeld JD. 2019. Metagenomic and chemical characterization of soil cobalamin production. ISME J 14:53–66.

83. Mera PE, Escalante-Semerena JC. 2010. Multiple roles of ATP:cob(I)alamin adenosyltransferases in the conversion of B_{12} to coenzyme B_{12}. Appl Microbiol Biotechnol 88:41–48.

84. Padovani D, Labunska T, Palfey BA, Ballou DP, Banerjee R. 2008. Adenosyltransferase tailors and delivers coenzyme B_{12}. Nat Chem Biol 4:194–196.

85. Mok KC, Hallberg ZF, Procknow RR, Taga ME. 2024. Laboratory evolution of *E. coli* with a natural vitamin B_{12} analog reveals roles for cobamide uptake and adenosylation in methionine synthase-dependent growth. bioRxiv https://doi.org/10.1101/2024.01.04.574217.

86. Gray MJ, Escalante-Semerena JC. 2009. The cobinamide amidohydrolase (cobyric acid-forming) CbiZ enzyme: a critical activity of the cobamide remodelling system of *Rhodobacter sphaeroides*. Mol Microbiol 74:1198–1210.

87. Woodson JD, Escalante-Semerena JC. 2004. CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B_{12} precursor cobinamide in archaea. Proc Natl Acad Sci 101:3591–3596.

88. Ma AT, Beld J. 2021. Direct cobamide remodeling via additional function of cobamide biosynthesis protein CobS from *Vibrio cholerae*. J Bacteriol 203:10.1128/jb.00172-21.

89. Ma AT, Tyrell B, Beld J. 2020. Specificity of cobamide remodeling, uptake and utilization in *Vibrio cholerae*. Mol Microbiol 113:89–102.

90. Helliwell KE, Lawrence AD, Holzer A, Kudahl UJ, Sasso S, Kräutler B, Scanlan DJ, Warren MJ, Smith AG. 2016. Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B_{12}. Curr Biol 26:999–1008.

91. Lennon SR, Batey RT. 2022. Regulation of gene expression through effector-dependent conformational switching by cobalamin riboswitches. J Mol Biol 434:167585.

92. McCown PJ, Corbino KA, Stav S, Sherlock ME, Breaker RR. 2017. Riboswitch diversity and distribution. RNA 23:995.

93. Johnson Jr JE, Reyes FE, Polaski JT, Batey RT. 2012. B_{12} cofactors directly stabilize an mRNA regulatory switch. Nature 492:133–137.

94. Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR. 2002. Genetic control by a metabolite binding mRNA. Chem Biol 9:1043.

95. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. 2003. Comparative genomics of the vitamin B_{12} metabolism and regulation in prokaryotes. J Biol Chem 278:41148–41159.

96. Chan CW, Mondragón A. 2020. Crystal structure of an atypical cobalamin riboswitch reveals RNA structural adaptability as basis for promiscuous ligand binding. Nucleic Acids Res 48:7569–7583.

97. Peselis A, Serganov A. 2012. Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch. Nat Struct Mol Biol 19:1182–1184.

98. Procknow RR, Kennedy KJ, Kluba M, Rodriguez LJ, Taga ME. 2023. Genetic dissection of regulation by a repressing and novel activating corrinoid riboswitch enables engineering of synthetic riboswitches. mBio 14:e01588-23.

99. Alvarez-Aponte ZI, Govindaraju AM, Hallberg ZF, Nicolas AM, Green MA, Mok KC, Fonseca-García C, Coleman-Derr D, Brodie EL, Carlson HK, Taga ME. 2024. Phylogenetic distribution and experimental characterization of corrinoid production and dependence in soil bacterial isolates. ISME J 18:wrae068.

100. Sultana S, Bruns S, Wilkes H, Simon M, Wienhausen G. 2023. Vitamin B_{12} is not shared by all marine prototrophic bacteria with their environment. ISME J 17:836–845.

101. Dorrell RG, Nef C, Altan-Ochir S, Bowler C, Smith AG. 2023. Presence of vitamin B_{12} metabolism in the last common ancestor of land plants. bioRxiv https://doi.org/10.1101/2023.06.07.544158.

102. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, Sterzenbach T, Tsolis RM, Roth JR, Bäumler AJ. 2011. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proc Natl Acad Sci 108:17480–17485.

103. Fontecave M. 1998. Ribonucleotide reductases and radical reactions. Cell Mol Life Sci CMLS 54:684–695.

104. Gonzalez JC, Banerjee RV, Huang S, Sumner JS, Matthews RG. 1992. Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*: Two solutions to the same chemical problem. Biochemistry 31.

105. Vetting MW, Al-Obaidi N, Zhao S, San Francisco B, Kim J, Wichelecki DJ, Bouvier JT, Solbiati JO, Vu H, Zhang X, Rodionov DA, Love JD, Hillerich BS, Seidel RD, Quinn RJ, Osterman AL, Cronan JE, Jacobson MP, Gerlt JA, Almo SC. 2015. Experimental strategies for functional annotation and metabolism discovery: Targeted screening of solute binding proteins and unbiased panning of metabolomes. Biochemistry 54:909–931.

106. Zallot R, Ross R, Chen W-H, Bruner SD, Limbach PA, de Crécy-Lagard V. 2017. Identification of a novel epoxyqueuosine reductase family by comparative genomics. ACS Chem Biol 12:844–851.

107. Helliwell KE, Collins S, Kazamia E, Purton S, Wheeler GL, Smith AG. 2015. Fundamental shift in vitamin B_{12} eco-physiology of a model alga demonstrated by experimental evolution. ISME J 9:1446–1455.

108. Hondorp ER, Matthews RG. 2004. Oxidative stress inactivates cobalamin-independent methionine synthase (MetE) in *Escherichia coli*. PLOS Biol 2:e336.

109. Mogk A, Tomoyasu T, Goloubinoff P, Rüdiger S, Röder D, Langen H, Bukau B. 1999. Identification of thermolabile *Escherichia coli* proteins: Prevention and reversion of aggregation by DnaK and ClpB. EMBO J 18:6934–6949.

110. Taga ME, Walker GC. 2010. *Sinorhizobium meliloti* requires a cobalamin-dependent ribonucleotide reductase for symbiosis with its plant host. Mol Plant-Microbe Interact MPMI 23:1643–1654.

111. Zhang C, Liu G, Huang M. 2014. Ribonucleotide reductase metallocofactor: Assembly, maintenance and inhibition. Front Biol 9:104.

112. Hallberg ZF, Nicolas AM, Alvarez-Aponte ZI, Mok KC, Sieradzki ET, Pett-Ridge J, Banfield JF, Carlson HK, Firestone MK, Taga ME. 2024. Soil microbial community response to corrinoids is shaped by a natural reservoir of vitamin B_{12}. ISME J 18:wrae094.

113. Höllriegl V, Lamm L, Rowold J, Hörig J, Renz P. 1982. Biosynthesis of vitamin B_{12}. Arch Microbiol 132:155–158.

114. Wang P, Wang Y, Su Z. 2012. Improvement of adenosylcobalamin production by metabolic control strategy in *Propionibacterium freudenreichii.* Appl Biochem Biotechnol 167:62–72.

115. Deptula P, Kylli P, Chamlagain B, Holm L, Kostiainen R, Piironen V, Savijoki K, Varmanen P. 2015. BluB/CobT2 fusion enzyme activity reveals mechanisms responsible for production of active form of vitamin B_{12} by *Propionibacterium freudenreichii*. Microb Cell Factories 14:186.

116. Stupperich E, Steiner I, Eisinger HJ. 1987. Substitution of Co alpha-(5-hydroxybenzimidazolyl)cobamide (factor III) by vitamin B_{12} in *Methanobacterium thermoautotrophicum*. J Bacteriol 169:3076–3081.

117. Anderson PJ, Lango J, Carkeet C, Britten A, Kräutler B, Hammock BD, Roth JR. 2008. One pathway can incorporate either adenine or dimethylbenzimidazole as an alpha-axial ligand of B_{12} cofactors in *Salmonella enterica*. J Bacteriol 190:1160–1171.

118. Crofts TS, Seth EC, Hazra AB, Taga ME. 2013. Cobamide structure depends on both lower ligand availability and CobT substrate specificity. Chem Biol 20:1265–1274.

119. Keller S, Kunze C, Bommer M, Paetz C, Menezes RC, Svatoš A, Dobbek H, Schubert T. 2018. Selective utilization of benzimidazolyl-norcobamides as cofactors by the tetrachloroethene reductive dehalogenase of *Sulfurospirillum multivorans*. J Bacteriol 200:e00584-17.

120. Keller S, Ruetz M, Kunze C, Kräutler B, Diekert G, Schubert T. 2014. Exogenous 5,6-dimethylbenzimidazole caused production of a non-functional tetrachloroethene reductive dehalogenase in *Sulfurospirillum multivorans*. Environ Microbiol 16:3361–3369.

121. Crofts TS, Men Y, Alvarez-Cohen L, Taga ME. 2014. A bioassay for the detection of benzimidazoles reveals their presence in a range of environmental samples. Front Microbiol 5.

122. Pherribo GJ, Taga ME. 2023. Bacteriophage-mediated lysis supports robust growth of amino acid auxotrophs. ISME J 17:1785–1788.

123. Gude S, Pherribo GJ, Taga ME. 2020. Emergence of metabolite provisioning as a by-product of evolved biological functions. mSystems 5:10.1128/msystems.00259-20.

124. Wang J, Zhu Y-G, Tiedje JM, Ge Y. 2024. Global biogeography and ecological implications of cobamide-producing prokaryotes. ISME J 18:wrae009.

125. Zhang Y, Rodionov DA, Gelfand MS, Gladyshev VN. 2009. Comparative genomic analyses of nickel, cobalt and vitamin B_{12} utilization. BMC Genomics 10:78.

126. Swaney MH, Sandstrom S, Kalan LR. 2022. Cobamide sharing is predicted in the human skin microbiome. mSystems 7:e00677-22.

127. Frye KA, Piamthai V, Hsiao A, Degnan PH. 2021. Mobilization of vitamin B_{12} transporters alters competitive dynamics in a human gut microbe. Cell Rep 37:110164.

128. Wienhausen G, Dlugosch L, Jarling R, Wilkes H, Giebel H-A, Simon M. 2022. Availability of vitamin B_{12} and its lower ligand intermediate α-ribazole impact prokaryotic and protist communities in oceanic systems. ISME J 16:2002–2014.

129. Kelly CJ, Alexeev EE, Farb L, Vickery TW, Zheng L, Eric L C, Kitzenberg DA, Battista KD, Kominsky DJ, Robertson CE, Frank DN, Stabler SP, Colgan SP. 2019. Oral vitamin B_{12} supplement is delivered to the distal gut, altering the corrinoid profile and selectively depleting *Bacteroides* in C57BL/6 mice. Gut Microbes 10:654–662.

130. Xu Y, Xiang S, Ye K, Zheng Y, Feng X, Zhu X, Chen J, Chen Y. 2018. Cobalamin (vitamin B_{12}) induced a shift in microbial composition and metabolic activity in an *in vitro* colon simulation. Front Microbiol 9:2780.

131. Zhu X, Xiang S, Feng X, Wang H, Tian S, Xu Y, Shi L, Yang L, Li M, Shen Y, Chen J, Chen Y, Han J. 2019. Impact of cyanocobalamin and methylcobalamin on inflammatory bowel disease and the intestinal microbiota composition. J Agric Food Chem 67:916–926.

132. Bunbury F, Deery E, Sayer AP, Bhardwaj V, Harrison EL, Warren MJ, Smith AG. 2022. Exploring the onset of B_{12}-based mutualisms using a recently evolved auxotroph and B_{12}-producing bacteria. Environ Microbiol 24:3134–3147.

133. Beauvais M, Schatt P, Montiel L, Logares R, Galand PE, Bouget F-Y. 2023. Functional redundancy of seasonal vitamin B_{12} biosynthesis pathways in coastal marine microbial communities. Environ Microbiol 25:3753–3770.

134. Heal KR, Qin W, Ribalet F, Bertagnolli AD, Coyote-Maestas W, Hmelo LR, Moffett JW, Devol AH, Armbrust EV, Stahl DA, Ingalls AE. 2017. Two distinct pools of B_{12} analogs reveal community interdependencies in the ocean. Proc Natl Acad Sci 114:364–369.

135. Morris JJ, Lenski RE, Zinser ER. 2012. The Black Queen Hypothesis: Evolution of dependencies through adaptive gene loss. mBio 3:10.1128/mbio.00036-12.

136. Taga ME, Ludington WB. 2023. Nutrient encryption and the diversity of cobamides, siderophores, and glycans. Trends Microbiol 31:115–119.

137. Wang M, Osborn LJ, Jain S, Meng X, Weakley A, Yan J, Massey WJ, Varadharajan V, Horak A, Banerjee R, Allende DS, Chan ER, Hajjar AM, Wang Z, Dimas A, Zhao A, Nagashima K, Cheng AG, Higginbottom S, Hazen SL, Brown JM, Fischbach MA. 2023. Strain dropouts reveal interactions that govern the metabolic output of the gut microbiome. Cell 186:2839-2852.e21.

138. Herren CM. 2020. Disruption of cross-feeding interactions by invading taxa can cause invasional meltdown in microbial communities. Proc Biol Sci 287:20192945.

139. Ramoneda J, Jensen TBN, Price MN, Casamayor EO, Fierer N. 2023. Taxonomic and environmental distribution of bacterial amino acid auxotrophies. 1. Nat Commun 14:7608.

140. Rodionov DA, Arzamasov AA, Khoroshkin MS, Iablokov SN, Leyn SA, Peterson SN, Novichkov PS, Osterman AL. 2019. Micronutrient requirements and sharing capabilities of the human gut microbiome. Front Microbiol 10:1316.

141. Sharma V, Rodionov DA, Leyn SA, Tran D, Iablokov SN, Ding H, Peterson DA, Osterman AL, Peterson SN. 2019. B-vitamin sharing promotes stability of gut microbial communities. Front Microbiol 10:1485.

142. Staley JT, Konopka A. 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu Rev Microbiol 39:321–346.

143. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Kosciolek T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Clauset A, Stevens RL, Shade A, Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. 7681. Nature 551:457–463.

144. Tecon R, Or D. 2017. Biophysical processes supporting the diversity of microbial life in soil. FEMS Microbiol Rev 41:599–623.

145. Howe JA, Smith AP. 2021. The soil habitat, p. 1–22. *In* Gentry, TJ, Fuhrman, JJ, Zuberer, DA (eds.), Principles and applications of soil microbiology. Third edition. Elsevier.

146. Gougoulias C, Clark JM, Shaw LJ. 2014. The role of soil microbes in the global carbon cycle: tracking the below-ground microbial processing of plant-derived carbon for manipulating carbon dynamics in agricultural systems. J Sci Food Agric 94:2362–2371.

147. Banerjee S, van der Heijden MGA. 2023. Soil microbiomes and one health. Nat Rev Microbiol 21:6–20.

148. Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. 2005. Algae acquire vitamin B_{12} through a symbiotic relationship with bacteria. Nature 438:90–93.

149. Men Y, Seth EC, Yi S, Crofts TS, Allen RH, Taga ME, Alvarez-Cohen L. 2015. Identification of specific corrinoids reveals corrinoid modification in dechlorinating microbial communities. Environ Microbiol 17:4873–4884.

150. Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G, Gordon JI. 2011. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proc Natl Acad Sci 108:6252–6257.

151. Nicolas AM, Sieradzki ET, Pett-Ridge J, Banfield JF, Taga ME, Firestone MK, Blazewicz SJ. 2023. A subset of viruses thrives following microbial resuscitation during rewetting of a seasonally dry California grassland soil. Nat Commun 14:5835.

152. Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, Maestre FT, Singh BK, Fierer N. 2018. A global atlas of the dominant bacteria found in soil. Science 359:320–325.

153. Starr EP, Shi S, Blazewicz SJ, Koch BJ, Probst AJ, Hungate BA, Pett-Ridge J, Firestone MK, Banfield JF. 2021. Stable-isotope-informed, genome-resolved metagenomics uncovers potential cross-kingdom interactions in rhizosphere soil. mSphere 6:10.1128/msphere.00085-21.

154. Kim M, Oh H-S, Park S-C, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351.

155. Lu X, Heal KR, Ingalls AE, Doxey AC, Neufeld JD. 2020. Metagenomic and chemical characterization of soil cobalamin production. ISME J 14:53–66.

156. Engel P, Dehio C. 2009. Genomics of host-restricted pathogens of the genus *Bartonella*. Genome Dyn 6:158–169.

157. Lochhead AG, Thexton RH. 1951. Vitamin B_{12} as a growth factor for soil bacteria. 4260. Nature 167:1034–1034.

158. Ford JE, Hutner SH. 1957. On the nature of the vitamin B_{12} requirement in soil bacteria isolated by Lochhead and his co-workers. Can J Microbiol 3:319–327.

159. Duda J, Malinska E, Pedziwilk Z. 1957. Relation between the vitamin B_{12} content and the microorganism count in soil. Acta Microbiol Pol 6:355–365.

160. Bradbeer C. 1999. Cobalamin transport in bacteria, p. 490. *In* Banerjee, R (ed.), Chemistry and biochemistry of B_{12}. John Wiley & Sons, Inc.

161. Shelton AN, Lyu X, Taga ME. 2020. Flexible Cobamide Metabolism in *Clostridioides (Clostridium) difficile* 630 Δerm. J Bacteriol 202:e00584-19.

162. Tang YZ, Koch F, Gobler CJ. 2010. Most harmful algal bloom species are vitamin B_1 and B_{12} auxotrophs. Proc Natl Acad Sci 107:20756–20761.

163. Tecon R, Ebrahimi A, Kleyer H, Erev Levi S, Or D. 2018. Cell-to-cell bacterial interactions promoted by drier conditions on soil surfaces. Proc Natl Acad Sci 115:9791–9796.

164. Raynaud X, Nunan N. 2014. Spatial ecology of bacteria at the microscale in soil. PLOS ONE 9:e87217.

165. Gregor R, Szabo RE, Vercelli GT, Gralka M, Reynolds R, Qu EB, Levine NM, Cordero OX. 2023. Widespread B-vitamin auxotrophy in marine particle-associated bacteria. bioRxiv https://doi.org/10.1101/2023.10.16.562604.

166. Escalas A, Hale L, Voordeckers JW, Yang Y, Firestone MK, Alvarez-Cohen L, Zhou J. 2019. Microbial functional diversity: From concepts to applications. Ecol Evol 9:12000–12016.

167. Walkup J, Dang C, Mau RL, Hayer M, Schwartz E, Stone BW, Hofmockel KS, Koch BJ, Purcell AM, Pett-Ridge J, Wang C, Hungate BA, Morrissey EM. 2023. The predictive power of phylogeny on growth rates in soil bacterial communities. 1. ISME Commun 3:1–8.

168. Martiny AC, Treseder K, Pusch G. 2013. Phylogenetic conservatism of functional traits in microorganisms. ISME J 7:830–838.

169. Placella SA, Brodie EL, Firestone MK. 2012. Rainfall-induced carbon dioxide pulses result from sequential resuscitation of phylogenetically clustered microbial groups. Proc Natl Acad Sci U S A 109:10931–10936.

170. Santos F, Vera JL, van der Heijden R, Valdez G, de Vos WM, Sesma F, Hugenholtz J. 2008. The complete coenzyme B_{12} biosynthesis gene cluster of *Lactobacillus reuteri* CRL1098. Microbiol Read Engl 154:81–93.

171. Zengler K, Zaramela LS. 2018. The social network of microorganisms — how auxotrophies shape complex communities. Nat Rev Microbiol 16:383–390.

172. Baran R, Brodie EL, Mayberry-Lewis J, Hummel E, Da Rocha UN, Chakraborty R, Bowen BP, Karaoz U, Cadillo-Quiroz H, Garcia-Pichel F, Northen TR. 2015. Exometabolite niche partitioning among sympatric soil bacteria. Nat Commun 6:8289.

173. Fossum C, Estera-Molina KY, Yuan M, Herman DJ, Chu-Jacoby I, Nico PS, Morrison KD, Pett-Ridge J, Firestone MK. 2022. Belowground allocation and dynamics of recently fixed plant carbon in a California annual grassland. Soil Biol Biochem 165:108519.

174. Sait M, Hugenholtz P, Janssen PH. 2002. Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. Environ Microbiol 4:654–666.

175. dos Santos HRM, Argolo CS, Argôlo-Filho RC, Loguercio LL. 2019. A 16S rDNA PCR-based theoretical to actual delta approach on culturable mock communities revealed severe losses of diversity information. BMC Microbiol 19:74.

176. Carlson HK, Piya D, Moore ML, Magar RT, Elisabeth NH, Deutschbauer AM, Arkin AP, Mutalik VK. 2023. Geochemical constraints on bacteriophage infectivity in terrestrial environments. ISME Commun 3:78.

177. Maidak BL, Cole JR, Lilburn TG, Parker CT, Saxman PR, Farris RJ, Garrity GM, Olsen GJ, Schmidt TM, Tiedje JM. 2001. The RDP-II (Ribosomal Database Project). Nucleic Acids Res 29:173–174.

178. Gould AL, Zhang V, Lamberti L, Jones EW, Obadia B, Korasidis N, Gavryushkin A, Carlson JM, Beerenwinkel N, Ludington WB. 2018. Microbiome interactions shape host fitness. Proc Natl Acad Sci U S A 115:E11951–E11960.

179. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, Cho H, Karaoz U, Loqué D, Bowen BP, Firestone MK, Northen TR, Brodie EL. 2018. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nat Microbiol 3:470–480.

180. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28:1823–1829.

181. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.

182. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2--approximately maximum-likelihood trees for large alignments. PloS One 5:e9490.

183. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296.

184. Tariq M, Liu Y, Rizwan A, Shoukat CA, Aftab Q, Lu J, Zhang Y. 2024. Impact of elevated CO_2 on soil microbiota: A meta-analytical review of carbon and nitrogen metabolism. Sci Total Environ 950:175354.

185. Wu H, Cui H, Fu C, Li R, Qi F, Liu Z, Yang G, Xiao K, Qiao M. 2024. Unveiling the crucial role of soil microorganisms in carbon cycling: A review. Sci Total Environ 909:168627.

186. Custódio V, Gonin M, Stabl G, Bakhoum N, Oliveira MM, Gutjahr C, Castrillo G. 2022. Sculpting the soil microbiota. Plant J Cell Mol Biol 109:508–522.

187. Goldford JE, 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 361:469–474.

188. Banerjee S, Schlaeppi K, van der Heijden MGA. 2018. Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol 16:567–576.

189. Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, Kemen EM. 2016. Microbial hub taxa link host and abiotic factors to plant microbiome variation. PLoS Biol 14:e1002352.

190. Dahlstrom KM, McRose DL, Newman DK. 2020. Keystone metabolites of crop rhizosphere microbiomes. Curr Biol CB 30:R1131–R1137.

191. Zimmer RK, Ferrer RP. 2007. Neuroecology, chemical defense, and the keystone species concept. Biol Bull 213:208–225.

192. Ferrer RP, Zimmer RK. 2013. Molecules of keystone significance: Crucial agents in ecology and resource management. BioScience 63:428–438.

193. Roy Chowdhury T, Lee J-Y, Bottos EM, Brislawn CJ, White RA, Bramer LM, Brown J, Zucker JD, Kim Y-M, Jumpponen A, Rice CW, Fansler SJ, Metz TO, McCue LA, Callister SJ, Song H-S, Jansson JK. 2019. Metaphenomic responses of a native prairie soil microbiome to moisture perturbations. mSystems 4:10.1128/msystems.00061-19.

194. Barberán A, Bates ST, Casamayor EO, Fierer N. 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J 6:343–351.

195. Shi S, Nuccio EE, Shi ZJ, He Z, Zhou J, Firestone MK. 2016. The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. Ecol Lett 19:926–936.

196. Morin MA, Morrison AJ, Harms MJ, Dutton RJ. 2022. Higher-order interactions shape microbial interactions as microbial community complexity increases. Sci Rep 12:22640.

197. Banerjee R, Ragsdale SW. 2003. The many faces of vitamin B_{12}: Catalysis by cobalamin-dependent enzymes. Annu Rev Biochem 72:209–247.

198. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114.

199. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927.

200. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2022. GTDB-Tk v2: Memory friendly classification with the Genome Taxonomy Database. Bioinformatics 38:5315–5316.

201. Escalante-Semerena JC. 2007. Conversion of cobinamide into adenosylcobamide in Bacteria and Archaea. J Bacteriol 189:4555–4560.

202. Molina Ortiz JP, Read MN, McClure DD, Holmes A, Dehghani F, Shanahan ER. 2022. High throughput genome scale modeling predicts microbial vitamin requirements contribute to gut microbiome community structure. Gut Microbes 14:2118831.

203. Noto Guillen M, Rosener B, Sayin S, Mitchell A. 2021. Assembling stable syntrophic *Escherichia coli* communities by comprehensively identifying beneficiaries of secreted goods. Cell Syst 12:1064-1078.e7.

204. Hessler T, Huddy RJ, Sachdeva R, Lei S, Harrison STL, Diamond S, Banfield JF. 2023. Vitamin interdependencies predicted by metagenomics-informed network analyses and validated in microbial community microcosms. Nat Commun 14:4768.

205. Titman D. 1976. Ecological competition between algae: Experimental confirmation of resource-based competition theory. Science 192:463–465.

206. Midani FS, Collins J, Britton RA. 2021. AMiGA: Software for automated analysis of microbial growth assays. mSystems 6:10.1128/msystems.00508-21.

207. Green MA, Alvarez-Aponte ZI, Trotter VV, Taga ME. 2023. Draft genome sequence of *Pedococcus* sp. strain 5OH_020, isolated from California grassland soil. Microbiol Resour Announc 12:e0002523.

208. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569.

209. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM, Liu P, Narrowe AB, Rodríguez-Ramos J, Bolduc B, Gazitúa MC, Daly RA, Smith GJ, Vik DR, Pope PB, Sullivan MB, Roux S, Wrighton KC. 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. Nucleic Acids Res 48:8883–8900.

210. Price MN, Deutschbauer AM, Arkin AP. 2020. GapMind: Automated annotation of amino acid biosynthesis. mSystems 5:e00291-20.

211. U.S. Census Bureau. 2019. U.S. Population Data.

212. U.S. Census Bureau. 2020. 2020 Census Results. https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-results.html. Retrieved 15 November 2024.

213. Liu S-N, Brown SEV, Sabat IE. 2019. Patching the “leaky pipeline”: Interventions for women of color faculty in STEM academia. Archives of Scientific Psychology 7:32–39.

214. Turner CSV. 2002. Women of color in academe: Living with multiple marginality. J High Educ 73:74–93.

215. Fisher R, Brock CH, Frahm T, Van Wig A, Gillis VR. 2019. Reflections on writing and identity: Exploring the role of qualifying exams in the sociocultural development of doctoral students. Stud Contin Educ 42:365–380.

216. Manus MB, Bowden MG, Dowd ET. 1992. The Purpose, philosophy, content, and structure of doctoral comprehensive/qualifying exams: A survey of counseling psychology training programs. Couns Psychol 20:677–688.

217. UC Davis Graduate Studies. Doctoral qualifying exam. https://grad.ucdavis.edu/doctoral-qualifying-exam. Retrieved 15 November 2024.

218. Doctoral Qualifying Exam/Candidacy | Graduate Student Resources | Graduate Support Office | USF College of Education. https://www.usf.edu/education/graduate/graduate-support-office/grad-stud-resources/qualifying-exam-candidacy.aspx. Retrieved 15 November 2024.

219. Lopez G. 2022. More than an exam. Science 375:582.

220. Calarco JM. 2020. A field guide to grad school: Uncovering the hidden curriculum. Princeton University Press.

221. Apple M, King N. 1977. Apple, Michael W. and Nancy King. “What do schools teach?” Curriculum Inquiry, 6 (No. 4, 1977), 341-358. Reprinted pp. 108-126 in Alex Molnar and John A. Zahorik, eds., Curriculum Theory, Washington, DC: ASCD, 1977. Reprinted pp. 29-47 in Richard H. Weller, ed., Humanistic Education: Visions and Realities, Berkeley, CA: McCutchan, 1977. Reprinted pp. 442-470 in George Willis, ed., Qualitative Evaluation, Berkeley, CA: McCutchan, 1978.*. CIRS Curric Inq Relat Stud Educ Res Searchable Bibliogr Sel Stud.

222. Burakgazi SG, Yildirim A. 2017. Journey into doctoral candidacy: A grounded theory study of doctoral qualification exam process. Croat J Educ - Hrvat Časopis Za Odgoj Obraz 19.

223. Yamamoto KR. 2023. Institutions’ role in trainee success. Science 379:1308.

224. Malcom SM, Parikh S. 2023. Students and postdocs deserve more. Science 379:519.

225. Shorb M. 1947. Unidentified essential growth factors for *Lactobacillus lactis* found in refined liver extracts and in certain natural materials. J Bacteriol 53:669–671.

226. Shorb MS. 1948. Activity of Vitamin B_{12} for the growth of *Lactobacillus lactis*. Science 107:397–398.