Bacterial Interspecies Interactions and Microbial Community Assembly

by

Anthony F. Ortiz Lopez

L.C.G., Universidad Nacional Autónoma de México (2016)

Submitted to the Microbiology Graduate Program
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author .....
Microbiology Graduate Program
August 30, 2021

Certified by .....
Jeff Gore
Associate Professor of Physics
Thesis Supervisor

Accepted by .....
Otto Cordero
Director, Microbiology Graduate Program

Thesis Committee Members

Otto Cordero, Ph.D.(Chair)
Associate Professor, Civil and Environmental Engineering
Massachusetts Institute of Technology

Nicole M. Vega, Ph.D.
Assistant Professor of Biology
Emory Univerisity

Tami Lieberman Ph.D.
Assistant Professor, IMES, Civil and Environmental Engineering
Massachusetts Institute of Technology

Jeff Gore, Ph.D.(Thesis Supervisor)
Associate Professor of Physics
Massachusetts Institute of Technology

Bacterial Interspecies Interactions and Microbial Community Assembly

by

Anthony F. Ortiz Lopez

Submitted to the Microbiology Graduate Program
on August 30, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Microbial communities play central roles in the development and maintenance of human health and in the functioning of the Earth's ecosystems. The microbial biodiversity of many environments has been thoroughly studied in recent years, yet the dominant processes shaping microbiota assembly remain unresolved. In this thesis, I leverage a bottom-up approach to experimentally build synthetic microbial communities and to test the prevalence of different ecological and evolutionary forces. High-throughput experiments in nanoliter droplets show a wide occurrence of bacterial interactions in the form of one species or consortium affecting the abundance and yield of another species. Positive and negative interactions appear unavoidable in bacterial co-cultures when growth is permitted, with growing bacteria typically facilitating non-growing bacteria. In this thesis, I also show that bacterial interspecies interactions in the *C. elegans* intestine are mostly competitive and hierarchical. Interestingly, simple two-species microbiotas can predict the composition of three-species and eight-species microbiotas in this nematode. Finally, we found that the importance of interspecies interactions is robust to bacterial strains with and without previous exposure to the *C. elegans* gut, and to worm mutants with different immune activities. These results show that constructing and characterizing synthetic microbial communities can elucidate fundamental principles for the control of microbial communities.

Thesis Supervisor: Jeff Gore
Title: Associate Professor of Physics

Acknowledgments

Special thanks to my Thesis Advisory Committee, Professors Nicole Vega, Otto Cordero, and Tami Lieberman, for their time and guidance on this thesis. And thanks to Audrey Saltzman and Daniel Rodriguez Amor for your thoughtful comments and words of encouragement.

I am most deeply grateful to my advisor, Prof. Jeff Gore, for giving me the great opportunity to work along his side. The last six years in the Gore Lab have been the most vibrant and gratifying experience I have ventured on, and I will be forever thankful to Jeff and all the amazing people at Physics of Living Systems for their wisdom, guidance, and friendship. To Daniel Rodriguez Amor, Nic Vega, Ty Ogun, Clare Abreu, Jonathan Friedman, Alfonso Pérez Escudero, Martina Dal Bello, Christoph Ratzke, Arolyn Conwill, Hyunseok Lee, Logan Higgins, Jiliang Hu, Shreyas Gokhale, Saurabh Gandhi, William Lopes, Simon Lax, Avihu Yona, Akshit Goyal, Sivan Pearl, Barrett Deris, Eugene Yurtsev, William Bloxham, Tommaso Biancalani, Manoshi Datta, Kevin Axelrod, Jinghui Liu, Hugh Higinbotham, Julien Barrere, Mikail Khona, Juan Díaz, Shayna Hilburg, Emiko Zumbro, Hejin Huang, Todd Gingrich, Melis Tekant, Alex Bacanu, and to Prof. Nikta Fakhri and Prof. Alfredo Alexander-Katz and their teams, it was my great honor to learn from and interact with all of you.

Thank you so much to Jared Kehe, Tony Kulesa, Nic Vega, Jonathan Friedman and Christoph Ratzke for the great years of collaborative work. The science in this thesis could not have occurred without you.

Thanks to MIT and its Microbiology Graduate Program. Thank you to my classmates Rachel Szabó, Liz Ward, Jai Padmakumar, and Sarah Schwartz, and to all other students in the program, including Mariana Matus, Bernardo Cervantes, Michaela Gold, and Joseph Elsherbini. Thank you to all the great faculty who taught me at MIT, including Jeff Gore, Serguei Saavedra, Otto Cordero, Martin Polz, Michael Laub, Alan Grossman, Paul Blainey, Scott Sheffield, Eric Grunwald, Tami Lieberman, Michael Follows, David Des Marais, Eric Alm, and Katharina Ribbeck. To

Otto, David and Katharina, thank you for your time and guidance on my qualifying exam. Special thanks to Jacquie Carota, Jacquin Niles, Kristala Prather, Martin Polz, and Bonnielee Whang for overseeing and managing our beloved program. Thank you to the Genomic Sciences program at UNAM and to Prof. Ayari Fuentes Hernandez and Prof. Rafael Peña Miller for all the energy and knowledge you gave me. Thank you to all my other teachers and professors throughout my two decades of education.

Thanks to my MIT Gymnastics team, my two-time-champion Fuerza Mexico soccer team, and my intramural waterpolo team. To Tathagata Srimani, Angie Cacciola, Avery Normandin, Marlis Denk-Lobnig, Matthew Sobiesk, Audrey Saltzman, Anna Platzek, Kwangjun Ahn, Sara Arganda, Andres Campero, Carolyn Fu, Fatima Hussain, Jordan Johnson, and all my other friends at MIT, Cambridge, Somerville, and Boston, thank you from the bottom of my heart! Thank you Rogelio Rodriguez Gonzalez, Diego Arroyo, Paloma Meneses, and Brian Ramirez for your invaluable long-distance friendship. Thanks to Stan Wulf, Steve Speredina, the FSILG Office, and all the Phi Delta Theta brothers for your camaraderie.

¡Gracias a toda mi familia! Thank you to my mother, Angélica Lopez Toriz, and my father, Fidel H. Ortiz Espinoza, for your enormous support over the years and the high moral values that you seeded within me. All my love to my siblings Atenea, Josue, Mitzi, and Edgar, and to my nieces Aruba and Nicté. Thanks to my aunt Elvia and all of my beloved family that through her become such, Pedro, Karina, Germán, Claudia, Fermín, Betsabé, Miguel, Adriana, and Daniela. Thanks to my great grandmothers, Ines Toriz Ríos and Trinidad Espinoza Mancilla. Thank you to all my extended family and ancestors from Morelos, Mexico City, Veracruz, and Guerrero. Finally, thank you to my brother Noe Ortiz Lopez—your departure changed it all, but your presence forever endures.

Thank you!

Contents

Chapter 1

Introduction

The biological sciences have grown immensely over the last few centuries. Biology in the 1500s referred mainly to medical descriptions of humans and luminous drawings of animals and plants. Nowadays, biology engulfs a multitude of fields ranging from ecology to epigenetics, from bioethics to astrobiology, and from marine biology to neurobiology. The concept of life is one of the most complex ones in our lexicon, and the scientific exploration of living things keeps adding evermore depth to our understanding of it. This thesis will guide you through my unique exploration of the biological sciences, emphasizing my research on microbial communities assembled in nanoliter droplets and in the *C. elegans* intestine and highlighting my findings regarding the interactions between bacterial species. But before diving into the details, let us first define microbiota assembly and let us connect this specific area of study to the microbiology field and the biological sciences at large.[1]

1.1 Microbiota assembly as part of the biological sciences

Microbiology is the scientific study of those living organisms smaller than what the naked eye can see. The development of early microscopes and the observations by

[1]Narrower and more concise introductions are also provided in Chapters 2 & 3.2.

Antoine van Leeuwenhoek, Robert Hooke, and Athanasius Kircher in the 1600s confirmed the existence of microscopic living creatures [1]. Such findings were instrumental for the development of the germ theory that G. Fracastoro seeded and that Louis Pasteur and Robert Koch completed in the 1800s, stating that disease spreads due to small and specific transmissible living particles. This germ theory had monumental implications for the human lifespan expectancy, which has nearly doubled in the last hundred years thanks in great part to hygiene improvement, which was pioneered by Ignaz Semmelweis and Joseph Lister, and due to vaccines [2], pioneered by Edward Jenner and Pasteur. These origins of microbiology show that the cohesiveness of this science does not come from its object of study, which has always been scattered and diverse, but rather from the employed techniques and its medical applications.

Nowadays microbiology extends well beyond the characterization of agents of disease (Figure 1-1). The fast life cycles of microbes makes them great model systems to study the cellular and molecular processes occurring within their walls. An exemplary molecular discovery in microbes is the CRISPR-Cas[2] system: originally a bacterial defense against viruses and now a revolutionizing tool in the biotechnology industry. Yeast microbes are also the industrial fermentation powerhouse behind many of the products that we love, from cheese and kombucha to biopolymers and biofuel. Most importantly, microbes are the quintessential degraders of organic matter. The extraordinary metabolic capabilities of microbes allows them to consume virtually any organic material and colonize any moist environment. The fields of cellular biology, biotechnology, and ecology (Figure 1-1) have come to grow in scope with a renewed attention to microbes. All in all, it has become clear that the majority of microorganisms are not inherently pathogenic, and that instead they hold a great beneficial potential for the human population.

In the late 1800s Serguei Winogradsky founded environmental microbiology with his studies on soil microbes [3]. Winogradsky observed that microbes modify the chemical states of nitrogen in the environment and provide a cycling of nutrients.

[2]CRISPR stands for clustered regularly interspaced short palindromic repeats. Cas are CRISPR associated cleaving enzymes.

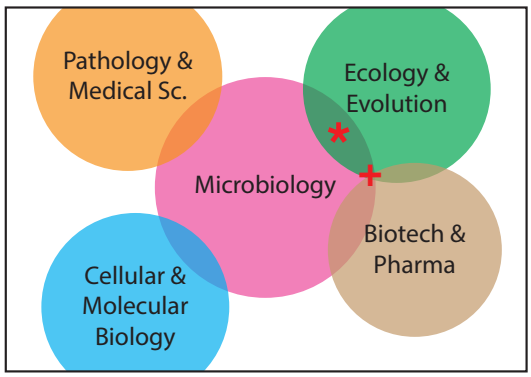

Figure 1-1: Diagram highlighting some biological sciences that synergize with microbiology and some of their intersections. *Microbial ecology* emerges at the intersection of *Microbiology* and *Ecology & Evolution*. Conceptual placement for my *C. elegans* and kChip projects are denoted with $\star$ and $+$, respectively.

Given how microbes interact between them to complete the stages of oxidation and reduction of nitrogen, Winogradsky's work is also considered the first ecological exploration of microbes [4]. Also in the late 1800s, Martinus Beijerinck discovered viral particles. Although it took decades to actually see such viruses, the field of virology had been seeded with Beijerinck's new techniques to enrich these novel microbes and grow them in batch culture. Winogradsky and Beijerinck's new approach to look at microbes in their environmental and community context created the field of microbial ecology.

Microbial ecology can be further broken down based on its multiple objects of study. At the single species level, or population level, microbial ecology is interested in:

1. the growth requirements and dynamics of individual species, and
2. defining the causes for a single bacterial type to stay homogeneous or to speciate into different types. Meanwhile, at the multi-species level, microbial community ecology is mainly interested in:
3. defining the species that are found in different natural environments and their functional role, and
4. understanding the assembly processes that lead a new set of microbes to form

a community.

The microbial speciation process (2) has direct connections to population genetics, while the biodiversity found in different microbial ecosystems (3) can be part of biogeography. In this thesis we will deal with (1) and (4), utilizing monoculture[3] measurements to understand microbiota assembly.

A microbial community, or microbiota, is a consortium of interacting microorganisms in space and time. Microbial communities refer most commonly to bacteria (prokaryotes), but single cell fungi and protists (eukaryotes), viruses, and archaea, are also very important members to consider. Some environments have very minimalistic microbiotas, harboring only a handful of species [5, 6]. Nevertheless, most natural environments are very complex, and their associated microbiotas are complex too. A recurrent theme in microbial biogeography is to assess if these complex distributions of microbial biodiversity follow the distributions of environmental variables, such as temperature and pH. Oftentimes, the microbial abundances do correlate with an environmental variable, and such correlations are informative of the broad-scale community assembly patterns [7, 8, 9]. The field of microbiota assembly complements microbial biogeography by means of testing such correlations and hypotheses with more controlled laboratory experiments. In the last two decades we've learned that the assembly of microbial communities depends on many factors, which I will explore in the rest of this introduction. Perhaps soon there will be a consolidation of such many factors into a more general theory on microbiota assembly.

Established fields of ecology and evolution jumpstarted the microbial ecology field. In 1916, Frederic Clements [10] summarized an ecological *succession theory*, which states that early colonizing grasses, bushes, and other species modify the environment such that middle and late colonizing species become more favorable—this succession continues in an orderly manner until a climax community is reached [11]. Originally, this theory resembled the ordered succession of plant communities to the highly orchestrated development of organisms [12], but nowadays the succession process is seen as much less deterministic. Successions in microbial communities have been observed

[3] A monoculture is a single species growing in isolation.

in solid compost, where yeast and some bacterial species decrease as temperature rise [13], in microbial mats [14], and in chitin particles floating in liquid, where early degraders are supplanted by non-degraders [15, 16].

The strong legacy of theoretical and experimental ecology has also helped advance our understanding of microbial communities, specially with the conceptualization of growth rates and ecological interactions. In 1932, Georgy Gause studied the simplified growth dynamics of two yeast species in liquid media. He found that each species has an effect in the abundance of the other species, with the production of alcohol being the probable cause of interaction. In stark contrast to Winogradsky and Beijerinck's search for novel microbial taxa performing specific environmental functions, Gause's work prompted the use of microbial communities with simple growth requirements as a form of testing ecological mathematical theory. Nowadays, the field of microbiota assembly makes extensive use of this experimental framework.

The assembly process of complex microbiotas is currently under intense scrutiny. From Leeuwenhoek's first observations of microbes and Pasteur's studies of pathogens in monoculture, to the first examinations of microbial communities and interactions by Winogradsky and Gause, great advancement has been made in the study of microorganisms. Microbial communities might be the most complex biological communities of them all, and to understand them we will need both to stand on the shoulders of giants and take gigantic, forward-thinking leaps.

1.2 Driving forces of microbial ecology and evolution

Microbial growth is the production of cellular mass that occurs due to favorable chemical reactions and stops due to saturation. This basic definition of cellular growth has a clear link to chemical reactions, which convert reactants into products based on the concentrations and the thermodynamics of the system. For example, breaking down the sugar glucose into pyruvate and further oxidizing pyruvate into carbon dioxide and water is a highly favorable reaction that produces as many as 38 molecules

of ATP[4]. However, when the concentration of glucose falls below a threshold, or when ATP or other products are overabundant, the reaction will come to a halt. Likewise, microbes stop dividing when nutrients are scarce or when the microbial mass exceeds the carrying capacity of the system. Microbial carrying capacities depend, among other things, on the level of crowdedness that a given microbe can withstand before its growth becomes inhibited by nearby kin. This growth inhibition is called intra-species interactions[5]. After reaching saturation, populations can experience long-term stability, oscillatory dynamics, or net negative growth rates (death rate). The concepts of growth and saturation are very simple yet powerful enough to explain in broad terms the ecological patterns of many microbial populations. In section 2.2.1 I will showcase how different bacterial species reach different population sizes in the intestine of *C. elegans*.

Due to the imperfect replication process of DNA, every time that a cell divides there is a small chance of introducing new mutations into a population. Although most new mutations don't get fixed in a population, some mutations do get inherited and become permanent, creating new alleles or genotypes. This change of allelic frequencies in a population over time is the modern meaning of evolution [17]. The deterministic part of evolution causes populations to adapt over generations to their specific circumstances[6]. This evolutionary force is called natural[7] selection. At the same time, evolution also occurs through more random, non-deterministic forces, such as drift, gene flow and mutation. I will expand further in the rest of this section on how these four forces occur not only in the evolution of microbial systems, but also in their ecology — Let's start by elaborating on the differences and the overlap of ecology and evolution.

Ecological dynamics are usually thought to occur faster than evolutionary dynamics. For example, a great novel mutation granting carriers an increase of 10% in offspring compared to the original genotype would take on the order of 100 genera-

[4]Overall oxidation of glucose: $6\,O_2 + C_6H_{12}O_6 + 38\,ADP \longrightarrow 38\,ATP + 6\,CO_2 + 6\,H_2O$

[5]Also referred as intraspecific and same-type interactions.

[6]Such match between species and environment can then be described as a habitat or niche.

[7]Selection can also occur in a less natural and more conscious way by human hand, such as in crops, livestock and pets.

tions to fix in a small population of 100 individuals. This theoretical calculation of extremely fast evolution is still slower than ecological events like a destructive wildfire or rising global temperatures, which can occur within one generation. Furthermore, all theoretical descriptions of evolution rely on a fitness variable that also occurs within a generation, which then makes evolution occur at lengthier timescales.

But here are two arguments against such simplifying distinction of timescales. First, some species, particularly microbes, have really fast generation times [18], which in theory could allow them to evolve faster than ecological dynamics[8]. Hyper-evolving microbial populations have indeed been observed [19] in long-term evolutionary experiments[9], and it's been proposed that the fast acquisition of mutations can override ecological constraints, leading to an "eco-evolutionary tunneling" [20]. Another feedback between population and evolutionary dynamics has been observed in a yeast system [21], highlighting that the study of different species oftentimes requires a recalibration of timescales [22]. Secondly, some ecological events and/or the time-step between them can occur very slowly. Events such as a chain of islands surging in the Pacific or the separation of Pangaea into continents occur more slowly than the slower examples of evolution. Given how these geological events can eventually dominate the development of an ecosystem [23, 24], they should also be taken into account in our ecological framework. The term historical contingency [25, 26] was coined by paleontologists to describe such geo-ecological forces. Overall, the ecological and evolutionary processes of natural systems are interlaced to an utmost degree.

Experimental microbial systems have made it possible to focus more closely on the shorter dynamics of a community by leaving chaotic environmental variables aside. Experiments can also be run fast and tight to avoid evolutionary dynamics. Microbiota assembly becomes then the core object of study in these experiments. One interesting and overarching observation in these systems is that communities converge to a stable state [27, 28]. Convergence means that starting from different pools of microbial species, an specific laboratory environment selects for similar commu-

[8]The larger population sizes of microbes could also lead to slower times to fixation, but larger populations allow for a faster exploration of the mutational landscape.

[9]Pioneered by Richard Lenski.

nity members, which tend to stabilize around an specific fractional abundance. Such intriguing observation makes us wonder, why do microbial communities converge? What determines the assembly of such stabilized microbiotas?

A framework used in traditional ecology to describe communities is that "species are added to communities via speciation and dispersal, and the relative abundances of these species are then shaped by drift and selection, as well as ongoing dispersal, to drive community dynamics" [29]. This framework has been widely applied to microbial systems too [30, 31]. These four categorical ecological forces are very similar to the forces that population genetics uses to explain evolution, with only two small changes: speciation instead of mutation, and dispersal instead of gene flow. It is then perhaps plausible to explain the ecology and evolution of microbiotas, as well as their assembly process, by studying these four conceptual forces in controlled experiments. In a recent study, Vega and Gore (2017) [32] explored the effects of bacterial migration into the *C. elegans* intestine, observing an stochastic colonization between equivalent bacteria (hence describing the dispersal and drift forces in this system). In Chapter 2, I will continue exploring the effects of bacterial competitive advantages (selection) and worm immune system defects (mutation) on the *C. elegans* microbiota assembly process. With this, we will complete a conceptual framework to describe the ecology and evolution of the microbial community in this system.

1.2.1 Natural selection and environmental filtering

Charles Darwin observed immense biodiversity on his travels to the Galapagos Islands and across the world. To explain the origin of so many different species, he and Alfred Wallace postulated in 1858 that individuals with certain maximized traits can be naturally selected to leave more descendants in a specific niche [33]. Such continued selection and increased offspring across many generations allows these novel type of organisms to separate from the original population and speciate. Among all the controversy that "The origin of the species" brought, special attention and criticism was given to the mechanisms of inheritance in nature. The work with pea plants by Gregor Mendel provided a more defined process by which the variation in a population

gets segregated across generations, and how a beneficial allele can be selected for. The work of Mendel and Darwin was amalgamated into a more defined evolutionary framework in the early 1900s, which is known today as the *modern synthesis*. Some great contributors to this stepping stone in evolutionary biology are Ronald Fisher, Julian Huxley, Ernst Mayr, and Theodosius Dobzhansky.

Aside from the mechanisms of inheritance, Darwin's thesis on natural selection relies on a "struggle for existence" or "survival of the fittest". This struggle is said to occur as both interactions within the species (intra-species) and between species (inter-specific or interspecies). Intra-species interactions are said to be the most severe form of competition, since it is the closest organisms that will compete for the most similar resources[10]. Such interactions connect to the concept of saturation that I introduced earlier. Given the relevance of interspecies interactions for my scientific results, I will explain these further in the next introductory section. The simple struggle for existence argument is that if an organism is at a disadvantage in an environment because the organisms around utilize the space and nutrients more efficiently, such organism will be filtered out of the community. This argument is known as the *competitive exclusion principle*[11], and it has dominated ecological thought and created long discussions over the last century [37, 38].

Baas Becking formulated in 1934 the also simple yet far-reaching statement that "every [microbial taxon] is everywhere, but the environment selects" [39]. Becking received partial instruction at the school founded by Beijerinck and was a big advocate of his enrichment techniques for culturing microbes.[12] Nowadays, many studies on microbial communities enlist this environmental selection or environmental filtering as one of the main forces shaping microbiota assembly [40, 41, 9]. In section 2.2.5, I will evaluate the importance of environmental filtering by comparing microbiotas

[10] Alle effects add nuance to the classical view that a species will inevitable compete most strongly with itself. Allee effects occur when, at low population sizes, growth rates become negative due to little cooperative or facilitative interactions and lead to a population collapse [34, 35, 36].

[11] Negative frequency dependent selection (NFDS) occurs when the per-capita growth rate of a species decreases as the population becomes more abundant, and it is usually observed in predator-prey systems. NFDS adds nuance to the competitive exclusion principle.

[12] To repeatedly see that specific environmental conditions can enrich for specific microbes regardless of the isolation origin was the probable motivation for Becking's influential statement.

assembled in the *C. elegans* intestine vs. the surrounding liquid environment.

The ideas of natural selection and environmental filtering are very similar. The first one puts emphasis on the struggle between organisms, while the second one highlights the abiotic environmental components. Nevertheless, the broad idea of natural selection does encompass both the biotic and abiotic components of the environment, and it recognizes the strong link between them. Overall, both ideas refer to the deterministic processes of ecology and evolution. Several other terms have been used to describe the concept of selection, such as species sorting [42], niche based selection, and stabilizing mechanisms [30].[13]

I will frame my experimental results on the *C. elegans* microbiota as an interplay between environmental filtering and interspecies interactions. I favor this framework because it acknowledges the importance of intra-species interactions by measuring carrying capacities, hence nodding at the "survival of the fittest." Then my framework quickly moves on to explore how the gut environment and the addition of other microbial species affect such initial population sizes.

1.2.2 Migration, dispersal, and invasion

In this subsection I will define the different forms of gene flow in ecological systems while overviewing recent literature on the topic[14]. Migration is the movement of organisms from one location to another. The movement of microbial cells can be a tightly regulated process, and it can result in complex dynamics in microbial communities. For example, *E. coli*[15] has a genetic regulatory network to decide how much energy to invest on the molecular motors (flagella) that enable its movement [43]. This bacterium senses its environment and moves towards higher nutritional gradients. Such directional motile behavior is called chemotaxis, and chemotactic

[13]Although the expansion of our scientific lexicon is a fine endeavor, giving many names to a similar concept can be detrimental to our clarity. Perhaps the current paradigm in microbial ecology is to break down the concept of natural selection into more specific processes leading to community assembly and evolution.

[14]This subsection won't introduce vital concepts for the understanding of this thesis, but it serves to complete my view on microbial ecology and community assembly. Narrower and more concise introductions are provided in Chapters 2 & 3.2.

[15]The Gammaproteobacteria *Escherichia coli* is the most studied microorganism of all time.

bacteria can have an advantage over non-chemotactic bacteria that are identical in every other regard [43]. Other recent studies of bacterial migration via motility have shown aggregation around lysed diatoms [44], optimal expansion speeds [45], larger populations due to mixing [46], and coexistence patterns through spatial niche partitioning [47, 48, 49].

Migration in microbial systems can also occur in the form of dispersal. This term highlights a more random, non-deterministic movement of cells. Dispersal can be easily implemented in the laboratory by mixing liquid cultures from different test tubes [50]. In these experiments, the array of test tubes models an irregular patchy landscape, such as islands [23] or scattered natural reserves [51]. In these experiments, dispersal has been found to increase the spread of adaptations [52] and modify the survival probability of a community [53, 54].

Migration and dispersal are also widely studied in more natural microbial systems. After sequencing a natural, semi-natural, or stabilized complex community, the level of similarity across biological replicates can be used to estimate the importance of dispersal in the assembly process [55]. It has been estimated that migration is important for microbiota assembly in wastewater treatment plants [56], and mineral microcosms like quartz [57]. However, migration might be less important in environments like bromeliads [58], and soils with low moisture [59]. The biogeography of soil microorganisms has also been studied in true islands [60], where it has been suggested that dispersal can regulate the biodiversity patterns of bacterial and fungi communities.

The invasion of an established community can follow after a migration event. Such invasions occur when an introduced species is capable of exploiting an underutilized resource [61] and/or when it has a competitive advantage over current community members. The "enemy release hypothesis" states that such competitive advantages tend to occur because there are fewer specifically adapted predators and competitors for the introduced species in a novel environment [62, 63]. Invasions can then be thought of as an interplay between migration and selection.

Several studies have shown the importance of microbial invasion for the assembly

of microbial communities and the close connections between invasion and interspecies interactions. For example, the direct growth inhibition from established microbial community members reduces the probability of invasion [64, 65]. In the *C. elegans* gut, *Salmonella* can establish an invader resistant population due to its ability to compete against other bacteria [66]. The expansion range of an invading species can also depend on the patchiness of the environment and the size of the invading population [35, 67]. Transient invaders that cannot persist in the community due to weak competitive abilities can still lead a community to switch across alternative stable states [68, 69]. The microbiota associated to seaweed is invaded by atypical microbes after exposure to antibiotics [70]. The transcription levels of an invader can be used to describe its niche [71]. And invasions are also widely implemented in theoretical models, where it is possible to test the stability of a community after many invasion events [61, 72, 73, 74]. Overall, directed migration, random dispersal, and invasion events are major forces that can shape the assembly of microbial communities.

The most relevant type of migration for my *C. elegans* work comes in the form of bacteria being ingested by animals [32, 75, 76]. I will expand on this form of migration in section 1.5.

1.2.3 Mutation, horizontal gene transfer, and speciation

Mutational change is another fundamental force driving the ecology and evolution of microbial communities [77]. Mutations can be thought as a tangible separation between ecology and evolution: mutations are oftentimes not needed for ecological dynamics to occur, but they are the fundamental generator of speciation and evolution. The most simple form of mutation comes in the form of single spelling errors introduced during the replication of the genetic code[16]. These simple genetic changes can accumulate [78, 79] and can cause DNA to evolve novel functions [80] or to repair regulatory mechanisms in the cell [81]. Single mutations can also cause diverging phenotypes that are capable of coexistence [48]. Mutation rates depend on mechanisms of DNA repair [82] and population dynamics [83], and fast mutation rates can lead

[16]Single nucleotide polymorphisms (SNPs)

to genetic hitchhiking [84] and novel evolutionary states [20]. Novel sequencing techniques detecting all mutations in a population have enabled better quantifications of natural selection [85], functions of genes [86, 87], and speciation thresholds [88].

Horizontal gene transfer (HGT) is another more impressive form of mutation. HGT occurs when a piece of DNA jumps from one organism to another non-descendent organism [89]. The three most recognized mechanisms of transfer within bacteria are conjugation [90], transformation [91], and transduction [92]. When HGT brings an advantage to the receiving organism, the novel genetic material can become permanent in the population. This was the case in the evolution of *Rhizobium*, a bacterial genus that forms close associations with plant roots because of a 500kb genomic island that has been shared through HGT [93, 94]. Enzymes utilized for chitin degradation have also been shared within the genus *Vibrio* [95]. HGT is more frequent between closely related organisms and between organisms occupying the same habitat [96]. Genetic plasmids can persist in populations through repeated HGT events [97]. HGT appears to be ubiquitous in microbial communities associated to humans, and of special importance is the spread of antibiotic resistance by this mechanism [98, 99, 96, 100]. SNPs and HGTs are two out of many other types of mutations that can lead microbes to accumulate changes in their genetic material.

Microorganisms traditionally binned into one single species can be widely different [101]. Such binning into one species is supported by close similarities in the ribosomal genes (16S), which are some of the oldest and most conserved genes [102]. While few mutations have occurred in core genes like these, bacteria have still diverged widely in other parts of their genetic material. These diverging genetic regions can be called the accessory, auxiliary, or flexible genome. Speciation in microbes can be thought of as an interplay of core genomes keeping species together and flexible genomes pulling them apart. The ecotype model proposes that some clusters in the large tight-knit bacterial diversity stay as cohesive units by means of HGT and gene selective sweeps [103, 104, 105]. The concept of a bacterial ecotype is similar to the more widespread concept of bacterial strain, but its value resides in allowing different species to be part of the same ecotype due to niche similarities.

Mutations occurring in higher organisms like animals can also influence host-associated microbial communities. The soil worm *C. elegans* was introduced as a model system in the 1960s by Sydney Brenner with the intention of creating mutant lines to study the function of specific genes [106, 107]. Worms that are deficient in their cellular signaling pathways can have higher [108] or lower susceptibility to pathogens [109]. Special attention has been paid to the TGFβ immune signaling pathway [110], which reduces the abundance of the bacterial family *Enterobacteriaceae* [111]. The DAF-2/IGF pathway also has a strong control of the microbial community size in the intestine of this worm [109, 112, 113]. Other more neutral mutations in *C. elegans* have also been described [114]. On the other side of the host-microbe associations, it has been found that bacterial mutants can be less virulent to worms through a decreased resistance to acid [115] and can also increase worm longevity through the expression of an exopolysaccharide called colanic acid [116]. Overall, mutations occurring in hosts and in microbes are an important driver of microbiota assembly.

1.2.4 Genetic and ecological drift

Our human minds are trained to search for the causes and consequences of things. However, many phenomena simply happen randomly. The directional forces of environmental filtering and natural selection shape to a large extent the ecology and evolution of microbiotas, and most attention is usually given to the description of such forces, but neutral processes are always present and can also be important for the assembly of microbiotas. In 1979, Stephen Gould and Richard Lewontin wrote about how non-adaptive processes can serve as better explanations for the evolution of some traits in biological organisms [117]. With the analogy of marvelous—yet not structurally relevant—paintings found on top of a church's columns, they drew attention to how some complex organismic traits could be the evolutionary byproduct of the more general evolutionary trajectory of the organism. This argument on the spandrels of San Marco remains relevant today [118] for the current developing discussion of how much a microbiota is co-adapted to the host where it is found [119]. My

research in section 2.2.4 argues against a strong host-microbe adaptation occurring in the *C. elegans* model system. In the meantime, here is a brief introduction to the large topic of stochasticity in biology.

Motoo Kimura developed the neutral theory for molecular evolution in the 1960s [120]. The central thesis of his work was to utilize Brownian motion and diffusion equations to mathematically model the frequencies of genotypes in a population. Genetic drift is then the random walk in the frequencies of alleles that can occur spontaneously when selection is small. The frequency of a mutation that brings neither harm nor benefit to carriers fluctuates randomly until it hits one of the absorbing boundaries and it becomes fixed or lost in a population. The fact that neutral mutations can fix in a population means that evolution can occur in a non-deterministic fashion. Similarly to this framework, Stephen Hubbell developed a neutral theory of biodiversity and biogeography in 2001 [24]. His central thesis was that in a neutral scenario, in which all organisms are equivalent (growth, dispersal, speciation, etc.), random fluctuations in the abundances of species can lead to rich biodiversity patterns that are comparable to the ones observed in nature. The neutral theories of Kimura and Hubbell don't disprove the concepts of natural selection and ecological niches. Instead, they are words of caution towards generalizing or overfitting deterministic ideas of adaptation to explain the existence of biological organisms and communities.

Review articles on microbial ecology and evolution have changed frameworks during the last 20 years. They have traversed from including little discussion on randomness [121] to including an stochastic niche theory [122] to acknowledging drift as a fundamental force shaping microbial communities [29, 30, 31]. Stochastic niche theory [61] serves as a bridge between the classical view of niches and trade-offs [123] and the more extreme view of neutrality in Hubbell's theory. One of the latest reviews on microbial ecology addresses the importance of integrating experimental data with mathematical models [124], which lies at the heart of the neutral theory framework. Earlier reviews on microbial biogeography framed the microbial distribution patterns as an interplay between randomness, environmental events, and historical events [125], or between speciation, dispersal, and extinction [126]. More recently,

microbial biogeography has also adopted the framework of selection, drift, migration, and speciation [127].

Drift and neutrality in microbial communities can be tested by comparing measured co-occurrence matrices and randomizations of themselves [128, 129, 130, 131, 132]. But a drawback of this framework is that it relies heavily on artificial thresholds to differentiate species. Other more complex methodologies utilize the average minimal phylogenetic distance between each taxon[17] in one community and the members of other communities [55, 57]. Other tests of "exact neutrality" have also been applied to human microbiota datasets [133]. Regardless of methodology, the assembly of microbial communities usually depends on both stochastic and deterministic forces [134, 58, 56].

Ecological drift in biological systems can occur in many forms [135]. Of great importance to my work, the bacterial colonization of the *C. elegans* gut can be stochastic due to the concentration of bacteria in the feeding substrate [32] and due to pharyngeal pumping bursts [136]. Also, the diversity of bacterial viruses can experience drift due to bottlenecks in the abundance of bacterial hosts within the mice intestine [137]. And similarly, functionally equivalent bacteria can get fixed or lost in a community in an stochastic manner [138]. But there are also other forms of biological noise, apart from traditional drift, that can play a role in bacterial population and community dynamics. For example, an isogenic population of *C. elegans* displays a variable onset in the decline of its tissues [139, 140]. The population growth rate of *E. coli* is larger than the mean growth rate of single cells due to noise [141]. Stochastic fluctuations in gene expression can be magnified by positive feedback loops and cause two distinct subpopulations within the same genotype [142]. Bet-hedging in bacteria refers to such phenotypic heterogeneities, which can be perpetuated in a population, and can also arise from epigenetic modifications and asymmetrical cell divisions [143]. Dynamical models exploring stochasticity in community assembly have also shown important roles for metabolic commensalism [144], positive density dependences [145], spatial

[17]Taxon (plular taxa) may refer to a taxonomic group of any rank, such as a species, family, or class, but it is mostly used to refer to a monoclonal population. Operational taxonomic unit (OTU) and single-nucleotide variant (SNV) are other more precise ways to refer to very similar bacteria.

structure [146], and other sources of noise extrinsic to the system [147].

As stated before, the ecology and evolution of microbial systems can be seen as an interplay of selection, mutation, migration, and drift. Out of these four forces, natural selection is perhaps the most complex, the most intriguing, and the most explored force of nature. My research projects and the rest of this thesis will focus mainly on bacterial interspecies interactions, which are part of natural selection. However, this previous introduction to mutation, migration, and drift hopefully places my research within a larger modern framework of microbial ecology and evolution. Testing all possible forces at play in microbiota assembly is experimentally unfeasible, but there is great value at integrating our specific results to the larger picture of microbial systems.

1.3 Bacterial interspecies interactions

Interactions occurring between organisms appeared briefly within the previous concepts of natural selection and invasion. In this section they will take the center stage. To explain bacterial interspecies interactions in more detail, first I will differentiate them from intra-species interactions and trophic interactions, which occur at lower and higher levels of phylogenetic distance, respectively. Secondly, I will elaborate on some important theoretical formulations surrounding interspecies interactions, such as the Lotka-Volterra model, intransitivity, and higher-order interactions. Afterwards, I will dive into the direct negative interactions that can occur between pairs of bacterial species. Positive interactions will follow after. Lastly, I will close this section with some nuances to the positive-negative framework.

Classical laboratory studies in the 1900s with yeast and ciliates showcased the existence of interspecies interactions in the microbial world [37, 148]. Such studies with microcosms laid the foundation for the current bottom-up exploration of microbiota assembly [149]. Novel experimental studies have exploited new techniques to decipher the molecular mechanisms stabilizing microbial communities [150] and have proposed testable community assembly rules [27]. My research will follow up the steps of these

and many other great pieces of science by measuring the prevalence of positive and negative interactions in static liquid cultures and in the *C. elegans* intestine.

1.3.1 From intra- to inter-species: a view on species boundaries

Organisms are said to be similar if they consume similar resources or if they are able to colonize similar environments. This so called niche overlap has caused great confusion over several decades. To make sense of the issue, let's first think on a *very specific niche* with well defined nutrients and spatial boundaries, and let's assume that organisms *A* and *A'* can colonize well this niche. Organisms *A* and *A'* are very similar in almost every regard, and there are only minor differences between them. The puzzle then becomes whether one type of organism will dominate in the community or if the two types will coexist. The competitive exclusion principle says that great similarity between organisms will limit their ability to coexist due to strong intra-species competition for space and nutrients. Gause and many other ecologists were big advocates of such exclusion principle based on laboratory experiments [151, 152]. However, how can different species—and hence interspecies interactions—be generated if one species drives all other emerging species to extinction?

The simple explanation is that different species can emerge due to ecological trade-offs and geographical separations (allopatric speciation) [153]. Very specific physiological niches can extend over broad spatial scales, at which point the properties of the adjacent environments can override the ecological competition that, at local scales, would lead to competitive exclusion [154]. The repeated observation of coexistence in nature, even at the most narrowly defined habitats and niches, led to the proposition that natural environments are kept far from equilibrium by external factors that change repeatedly over time [155]. Hence, whether organisms *A* and *A'* coexist and perhaps speciate or if they exclude each other is strongly dependent on the temporal and spatial scales of the niche. I will later show in my *C. elegans* project that, at the scale of multiple worms, bacterial competitive exclusion in the gut is unfrequent,

even between closely related species of the genus *Pseudomonas*.

DNA sequencing and phylogenetic relationships have come to greatly improve biological taxonomy and nomenclature [152, 134]. The great early work by Carl Linnaeus in the 1700s has been extended and adapted to include microbial diversity [156]. From closely similar to highly distant, it is still very useful to classify microorganisms into species, genera, families, orders, classes, phyla, and domains, but the similarity between organisms can now be measured more precisely with phylogenetic distances. Such distances have enabled more coherent and systematic ways to classify the large-scale diversity of the bacterial domain (perhaps into more than 90 different phyla [157]). At small-scale diversity, sequencing has complemented ecology to differentiate bacterial strains with as little as 100 SNPs along their genomes [158]. Interestingly, as phylogenetic distance increases, there appears to be a stepwise increase in the probability of inhibition between different bacterial isolates of the genus *Streptomyces* and *Vibrio* [159, 160]. However, at larger phylogenetic distances, the relationship gets inverted [161]. This stepwise increase followed by a gradual decrease in inhibitory interactions between different organisms is currently sculpting a better understanding of the central concept of species. I will later show that phylogenetic distances between bacteria also correlate with differences in fractional abundances in the *C. elegans* microbiota (Figure 2-4).

As phylogenetic distances increase, different trophic levels can appear, hence creating trophic interactions [162]. The idea of different trophic levels is very important in traditional ecology. A predator, for example, is said to be in a different level in the network of interactions than its prey [163], and the addition of a top predator can lengthen or shorten the food chain depending on the trophic levels that it preys upon [164]. Trophic levels can appear in the microbial world in many forms, such as a ciliate protozoan preying on Actinobacteria [165], flies, worms and other animals interacting with their microbiota[18] [166, 167, 168, 169], or plants interacting with microbes [58, 170]. The concept of environmental filtering is more widely used

[18]Nutrients provided by the host can be important for microbiome ecology, and microbes in turn can provide nutrients for the host.

to describe the interactions between hosts and microbes, but given how hosts are also biotic entities, it might also be appropriate to include host-microbe interactions within the framework of trophic interactions.

Where should we draw then the boundaries between intra-species interactions, interspecies interactions, and trophic interactions? Two simplifications can circumvent the complications around the concepts of species and trophic levels. First, any bacterial strain coming from a different isolation event (i.e. different colonies) can be considered to be a different species. Secondly, given that all bacteria occupy similar µm spatial scales, they could all be considered to be in the same trophic level[19]. After these two simplifications, all interactions occurring between non-clonal bacterial organisms would be considered bacterial interspecies interactions.

1.3.2 Theory, modeling, intransitivity, and higher-order interactions

The simple approach to mathematical modeling is to simplify a natural system into the most basic form, namely an equation or an algorithm [171], that still recapitulates the observations. This approach helps to identify the main variables playing a role in the dynamics of the system. The more nuanced approach to modeling is to also simplify the system but to get results different from traditional thought and experimental observations [172]. Here, the simplifying assumptions and the logical steps of the model attempt to challenge the paradigms in the field. Finally, there is the type of modeling that doesn't bind itself to experimental data and simply explores the dynamics of theoretical scenarios. Given the uncertainty associated to ecological data and the difficulty of collecting it, ecological theory and modeling oftentimes remain separate from field and experimental ecology. Experimental microbial systems have been useful at bridging such gap. My scientific projects themselves focus

[19]Gralka et al. (2020) [162] argue that microbes can be divided into more defined trophic levels, where secondary consumers feed on the byproducts of primary degraders. However, primary degraders and secondary consumers can interchange depending on the nutrient source. Hence, we favor a simplification of microbial communities into networks of species (fixed nodes) and interspecies interactions (variable links) rather than networks of trophic levels (variable nodes) and trophic interactions (conceptually fixed links).

on the experimental side of ecology, but they connect strongly to the quantitative and mathematical approach to interspecies interactions that theoretical ecology has championed.

Alfred Lotka and Vito Volterra proposed independently in the early 1900s a pair of first-order differential equations to model the dynamics of predator and prey populations[20]. The most important feature of the model is the explicit representation of the interspecies interactions as parameters connecting the abundances of the two species. Many studies have extended, simplified, and generalized the Lotka-Volterra model [173]. Some of these include explorations of quadratic forms for the interactions [174], lattice statistical mechanics [175], higher-order interaction parameters [176, 177], explicit death rates [178], non-symmetric interactions [146], logistic growth without interactions [147], and strong positive interactions [179]. Other ecological models also display similar interaction parameters connecting the species dynamics. The Beverton-Holt competition model can describe in discrete-time the population dynamics of plants, and it expresses competitive ability as a function of fecundity in addition to intra-species and interspecies interactions [180]. Another discrete-time model is the autoregressive Gompertz model, which has been used to predict grasslands' community composition with an intra-species interaction and a pooled interspecies interaction [181]. Overall, most of these equation-type ecological models include intra- and interspecies interactions, on top of a growth source, as plausible knobs that can regulate the abundances of species, while model extensions draw attention to other possible variables playing a role in ecological systems.

Other important ecological models include food webs [182], mutualistic positive networks [183, 184, 185], and other networks of interactions [186, 187][21]. Ecological models also come in the forms of flux balance analysis [191, 192, 193], cellular automata [194], resource-dependent invasions [61], stochastic metabolic commensalism [144], other consumer-resource models [195], and models with saturating functional

[20]The traditional LV model uses $\frac{dx}{dt} = \alpha x - \beta xy$ to represent the prey population, and $\frac{dy}{dt} = \delta xy - \gamma y$ for the predator population. The parameters β and δ connect the dynamics of the two equations.

[21]Networks are also widely used to model cellular organization and transcriptional regulation [188, 189, 190].

responses [196, 73]. The non-linear functional responses allow for a saturation of the interaction effects. Interestingly, it is possible to recapitulate the dynamics of multiple ecological models with only the mean, variance, and symmetry of the interactions [197][22]. Overall, the multitude of modeling approaches in ecology can perhaps be divided into dynamic equations, networks, and algorithms. A central theme in all of them is the exploration of biodiversity and stability as a function of interspecies interactions.

Intransitive interactions are one of the most widely proposed mechanisms leading to coexistence and biodiversity. In a system with three species, A, B, and C, if species A drives B extinct, B drives C extinct, and C drives A extinct, the pairwise interspecies interactions in the system, which resemble the rock-paper-scissors game, are defined as intransitive [198]. It has been observed that a cellular automata model implementing this intransitivity can lead to theoretical coexistence of species at low mobility regimes [199, 194, 200]. Intransitive networks that can add or cut interactions between the species can lead to long-living coexistence in a non-equilibrium quasi-steady state [201]. The abundance of intransitive interactions has been estimated to be almost 50% of all possible triads in plant communities [202]. However, contrary to these results in plant communities and the original theoretical proposition, studies with microbial systems have indicated lower levels of intransitive interactions [64, 203]. I will later show in my experiments that bacterial species tend to coexist in the *C. elegans* gut despite having hierarchical and transitive interactions.

Higher-order interactions (HOI) refer to the modification of a pairwise interaction by a third entity. HOI in ecology are similar in nature to the concept of epistasis in genetics, where a gene has an effect on how other genetic components interact [204]. HOI can also be defined as the non-additive effect of species towards a community trait, or as a failing null pairwise model [205]. The concept of "associational effects" is used in plant ecology to describe spatial structure and HOI between a plant and its biotic surroundings [206]. Three-way and four-way HOI have been shown to sta-

[22]Randomizing interaction networks, while keeping these parameters fixed, leads to similar predictions of biomass, fraction of survivors, diversity, and variance.

bilize biodiversity in simulations [207, 177]. On the experimental side, an *E. coli* that can invade a *Chlamydomonas* algae and a *Tetrahymena* ciliate cannot invade a community with those two species [65]. The mechanics of this HOI is that the algae inhibits aggregation of *E. coli*, hence leaving the bacteria vulnerable to ciliate predation. Another example of HOI is how pairwise interactions inferred from two-species microbiotas in zebrafish do not predict the high coexistence in five-species microbiotas [208].

The frequency and prevalence of HOI in microbial systems is currently under debate [209]. In 1969, John Vandermeer studied protozoan microorganisms and found that a Lotka-Volterra model without HOI could predict the composition of a four-species community [148]. This early finding has been complemented with novel studies on bacteria, which have found that the survival outcomes of pairwise experiments are enough to predict with high accuracy the survival of species in larger communities [27]. Part of my experiments in *C. elegans* push this last framework to show that fractional abundances of intestinal microbiotas can also be predicted with simple pairwise experiments. Theory and mathematical modeling have shown a large array of mechanisms that can lead to complex communities, such as intransitivity and HOI. Testing the occurrence of such mechanisms in real biological systems is one of the efforts in the field of microbiota assembly.

1.3.3 Bacterial competition is frequent based on low relative yields

Interactions occurring between pairs of bacterial species can be classified on a positive-to-negative axis. Positive and negative interactions are usually called facilitation and competition, respectively [210, 28]. Competition can also refer to bidirectional negative interactions [211], but apart from Chapter 3, I will use the term competition with its broader unidirectional sense. Competition in animal ecology has been broken down into exploitation, interference, and predation [163]. While exploitation occurs due to an overlap in the utilized resources, interference and predation refer to direct forms

of inhibition. A similar separation between direct and indirect negative interactions has been made for microorganisms, where exploitation for substrates or receptor sites is differentiated from interference or antagonism via antimicrobial substances [212].

Interference, antagonism, or inhibition between microbes are widely documented phenomena. In 1928, Alexander Fleming reported spots of no-bacterial-growth on a surface that should have been fully covered with bacteria. He discovered that an accidental fungal contaminant was producing a compound inhibiting the growth of the pathogen *Staphylococcus aureus*. The compound was named Penicillin, and its production and commercialization went on to save millions of lives. Fleming's accidental methodology has been widely improved over the last century. Modernized clearing lawn assays, Burkholder assays, and spent media assays have revealed a variety of compounds with antimicrobial properties [213]. Apart from its important medical applications, the observation of low bacterial growth or low productivity areas in co-culture experiments is the most evident type of bacterial negative interactions. These same methodologies have shown that the probability of direct inhibition between bacteria is a function of their phylogenetic separation [159, 160, 161, 5]. Clearing lawn assays have also shown that the activation of inhibitory mechanisms in bacteria can be subject to HOI among community members [200]. Overall, direct bacterial inhibition is the most thoroughly documented form of competition.

The broader concept of bacterial competition has been widely assessed over the last decade, and multiple review articles have listed several forms in which it takes place [210, 211, 214, 215]. One of these reviews, for example, highlights possible "buffer zones", where a high-diversity community neutralizes direct inhibition between species [210]. Following theoretical results, competition is predicted to induce a local reduction in diversity and an increase in ecological stability[23] in microbial communities. Another review on bacterial warfare highlights possible evolutionary paths that could have led to the diverse array of attack and defense mechanisms in bacteria [215].

The concepts of *negative interactions* and *competition* are usually understood in-

[23]Probability that a community will return to its previous state following a small perturbation.

tuitively in broad terms, but to agree upon the specifics of such concepts is a more difficult task [163]. Experimentally, microbial competition is measured in multiple ways. For example, in experiments with protists, a larger number of extinctions in a community than in all the monocultures combined was used as evidence of competition due to niche overlap and/or inhibitory effects [216]. Competition has been used to describe a longer and steadier attachment to surfaces that allows a newly differentiated bacterial species to reach higher densities than others [47]. Competitive exclusion has been observed in biofilms, where competition was detected through antibiotic stress [217]. A bacterial species can modify its surrounding physiological conditions and create a restrictive environment that disables the growth of other bacteria [218, 212]. And measuring the "growth rate when rare" of an introduced species has shown high levels of intransitive competition in plants [219]. These are some few examples of the several current experiment-based interpretations of competition.

Bacterial competition has also been reported through several inference methodologies. First of all, many studies parameterize Lotka-Volterra models with experimental data, and the recovery of negative interaction parameters is then used as evidence of competition [27, 148, 220, 221, 222, 223]. Based on genomic metabolic overlap, it has been suggested that cooccurring bacteria in human guts compete more strongly than non-cooccurring species [224]. Interspecies competition has also been inferred from the proteome overlap of *Pseudomonas aeruginosa* and *Candida albicans* [82]. Bacterial negative interactions have even been estimated with measurements and overlaps of proton nuclear magnetic resonance spectra [225]. Several other methods have been proposed to infer bacterial interactions from cooccurrence matrices obtained from sequencing data [226, 227], but these inferences don't come without challenges [228, 229]. Based on cooccurrence and genomic data, interspecies interactions in the *C. elegans* microbiota have been estimated at 34% negative, while the substrate around the worms had a much lower proportion of negative interactions [230, 231]. Although advanced methodologies to infer interspecies interactions have illustrated important ecological patterns in microbial communities, it has come at the cost of blurring the lines between the concepts of interspecies competition, negative

correlation, and trait overlap.

Relative yields are a well delineated methodology to define competition. Relative yields rely on comparing the abundance of each bacterial species in a community to its abundance in monoculture. A species reaching a lower abundance in a community than in isolation can be defined as receiving a net negative interaction from the members of the community. The yield or abundance of a species can be measured as its number of cells, colony forming units, biomass, florescence intensity, or other population size proxies. This simple calculation of relative yields can be easily adopted when individual species are capable of growing on their own. I will utilize relative yields throughout this thesis as the means of quantifying interspecies interactions in static liquid cultures and in the *C. elegans* intestine.

Two connected observations usually arise when comparing traits[24] at the level of the community and monocultures. First, the productivity of the community is much less than the additive null model of all monocultures [232, 233]. This simple result points towards a widespread niche overlap between microbial species. Secondly, the relative yields of the species are usually below one, meaning that the productivity of each species is hampered by the addition of other community members. My *C. elegans* microbiota experiments align with this view of competition. However, when the productivity of a species in monoculture is already very low, or even falls below the limit of detection, it would be hard for other community members to drive such low productivity even lower. My high-throughput experiments in liquid culture measure the productivity of bacterial species in environments where they cannot grow alone, and it is in these cases where we find an abundance of relative yields higher than one, which are indicative of positive interactions or facilitation.

1.3.4 Positive interactions, communication, and cross-feeding

Positive interactions, facilitation, or cooperation refers to an increase in productivity, fitness, and/or survival rate of a species due to the presence of another species. Allee

[24]Biomass is the most commonly studied microbial trait, but other interesting traits include antibiotic production and degradation, secretion of metabolic byproducts, etc.

effects, for example, can appear when intra-species positive interactions are needed by a population to increase in size [34, 67]. Likewise, interspecies facilitation can occur and shape the structure of microbial communities. For example, communication via quorum sensing [234] allows for a coordinated production of molecules, proteins, or biofilms that can provide a benefit to the group. Signaling molecules can also lead to higher antibiotic tolerance in the intestine of *C. elegans* [235, 236]. Conserved quorum sensing mechanisms have been found across different marine bacterial species [237], so facilitation by means of communication can be far-reaching.

Cross-feeding refers to the exchange of metabolites between organisms [238]. The exchange of primary metabolites such as vitamins, amino acids, nucleotides, or growth factors can facilitate the growth of bacterial neighbors [239]. Cross-fed nutrients can be essential for the interacting bacteria, in which case the coexistence or competitive exclusion of the species can depend on the external availability of the limiting nutrient [38]. Cross-feeding is regarded as an specific interaction that has been selected for, while the slightly different concept of overflow metabolism refers to an unspecific release of unconsumed nutrients or metabolic intermediates [240, 241, 242]. A recent important study on microbiota assembly found that four Gammaproteobacteria[25] are able to grow on the spent media of every other bacteria [28]. Such direct evidence of positive interactions was complemented with further observations of bacterial growth when the primary nutrient was depleted, and with mathematical models that incorporate by-product secretions. This study then highlights a widespread nonspecific cross-feeding, or overflow metabolism, as the main determinant of microbial community assembly.

Other forms of cooperation between microbial species have been observed apart from communication and cross-feeding, such as in the extracellular digestion of resources and protection against antibiotics [149]. For example, positive interspecies interactions have been inferred on the metabolic activity of a four-species community [245]. Spent media assays with bacteria from the urinary tract have shown a similar

[25]Gammaproteobacteria, or simply "Gammas", is a highly diverse and highly studied bacterial clade where several human pathogens are found [243, 244].

prevalence of positive and negative interactions, with 18% and 23% on each side of the spectrum [5]. Based on a graphical lasso[26] approach, interactions have been inferred as mostly positive among highly abundant and conserved species in the human microbiome [246]. RNA-sequencing of a marine bacteria invading a phytoplankton bloom has suggested that positive interactions can enlarge the realized niche of the bacteria beyond its fundamental niche [71]. These different studies on microbial positive interactions suggest that competition is not the rule across interspecies interactions. My reserach in Chapter 3 will continue to emphasize the commonality of bacterial positive interactions at larger scales.

The origin of cooperation and social traits in bacteria has recently received wide attention. The tragedy of the commons is a concept in economics, morality, and social evolution theory. "The tragedy is that as a group, individuals would benefit from cooperation, but cooperation is not stable because each individual can gain by selfishly pursuing their own short-term interests." [247]. However, it has been observed that a yeast strain investing on a cooperative behavior can coexist with its non-cooperative counterpart due to rare strategies outperforming common strategies [248]. It is worth noting that interactions between different guilds or different trophic levels are less constrained towards competition, and in this regard, many positive interactions have been described between plants and microbes [249, 250, 251]. It has been posited that positive interactions lead to more productive yet less stable microbiotas, so plant and animal hosts have incentives to balance the interspecies interactions within their microbiotas to balance their stability and productivity [119]. My research won't delve into the evolutionary origins of facilitation. Hence, it is important to highlight that positive interactions could be either specific and selected for (symbiosis), or unspecific and unavoidable (metabolic overflow).

[26] A sparse penalized maximum likelihood estimator for the concentration or precision matrix of a multivariate distribution.

1.3.5 Nuances in the positive-negative axis of interactions

Bacterial interspecies interactions are oftentimes more complex than a simple parameter on a positive-to-negative axis. Apart from higher-order-interactions, other added complexities to the concept of microbial interactions have emerged. For example, in the previously mentioned system of cooperative and cheater yeast strains [248], the cooperator strain was engineered to degrade sugars intracellularly to avoid the parasitic interaction. This mutant was capable of avoiding the decrease in relative yields when co-cultured with the cheater. However, this avoidance of competition came at the cost of reaching lower absolute abundances than the original cooperative strain. In other words, the public good generated from an extracellular degradation of sugars creates a niche for competition to emerge, but at the same time it allows for faster metabolic and productivity rates.

Another nuanced interspecies interaction similar to the concept of cheaters occurs with the motile oral Bacteroidetes *Capnocytophaga gingivalis*. Other non-motile bacteria can bind to *C. gingivalis* and can be transported over long distances as cargo [252]. Based on how the non-motile bacteria skip the cost of gliding, one could infer a parasitic interaction between the species; however, based on the correlated cooccurrence patterns of both species, one could estimate a mutualistic interaction; and based on the strong overlap in the consumed nutrients, one could estimate a bidirectional competitive interaction. A measurement of the relative yields of both species could be an accurate estimation of their interaction, but only if it takes into account the spatial scale at which the coupled migration occurs. The microscopic imaging of dental plaque shows that *C. gingivalis* forms repetitive spatial patterns with other non-motile bacteria, which is then suggestive of a mutualistic interaction [253].

Interactions can also change due to conditions extrinsic to the species. A transition between competition and cooperation due to the concentration of ammonia in the medium has been described in a phototrophic microbial community [254]. The initial inoculum ratio between a *Pseudomonas* and an *E. coli* species had an effect on the bacterial interactions in static cultures [255]. The competitive hierarchy of two *E.*

coli strains can be inverted due to active segregation and spatial exclusion [49]. The niche complementation and competition between the members of a seawater microbial community increase as taxonomic richness increases [233]. The overrepresentation of network motifs in plant communities has also suggested that competition between two species can arise from the presence of a third interacting species [256]. Finally, it has been estimated that different *C. elegans* mutants enrich for different bacterial interactions within the worm microbiota [113]. Overall, these studies highlight how interspecies interactions depend on many ecological parameters. My comprehensive quantification of bacterial interactions presented in Chapter 3 connects to this emerging view of interspecies interactions as a more responsive rather than permanent link between species.

Fast rates of evolution can also add nuance to bacterial interspecies interactions [20]. High rates of genetic exchange via HGT between spatially and phylogenetically associated microbes can decrease the stability of a competitive advantage [134, 90]. If a niche-determining gene is shared among microbiota members[27], the once deterministic coexistence or competitive exclusion of the neighbor species can shift to a neutral drift scenario. Similar to this, evolution of plants in response to interspecies (interspecific) competition resulted in more even population sizes of the competitors via evolved changes in competitive ability [258].

Although direct inhibition can be well tested in laboratory settings, some other interspecies interactions cannot be easily differentiated between direct or indirect. Accidental predation between animal species [163] and accidental generation of restrictive physiological conditions in bacterial systems [212, 218] are examples of these.

Other frameworks define interspecies interactions based on the productivity of the community [232, 233]. The highest of the monoculture yields can be thought as occupying a maximum niche, so an increase from this observed maximum could be called a positive interaction from the added species to the community productivity. Connecting to this framework, my experiments with *C. elegans* show that most co-

[27] Aquisition of mobile genes can be important for bacterial species to colonize specific human populations [257].

cultures reach lower community sizes than the higher of the monoculture population sizes, which would still be indicative of competition. However, other studies define competition as a community trait lower than the average of the monoculture traits [245], or lower than a calculated community trait based on fractional abundances [255]. These other frameworks may be of use when the community productivity is the object of study. Nevertheless, when the fractional abundances of the community, or community assembly, is the main object of study, the definition of positive and negative interactions makes most sense based on relative yields.

1.4 Massively parallel screening of synthetic microbial communities

Microbial communities exhibit emergent consortia-level functions that are vital to all ecosystems on Earth[28]. The complexity of microbial interactions and environmental dependencies can lead to unpredictable behaviors even in apparently simple communities, posing a challenge to consortia design [260, 261, 262, 263, 210, 264, 265, 266, 267]. High-throughput phenotypic screening has found widespread use as a discovery strategy for novel gene targets [268, 269] and drugs [270], but its adoption in microbial consortia discovery has been hindered by the logistical complexity of constructing strain combinations. Conventional liquid handling techniques and platforms (such as 96-well plates) may not be sufficient to adequately sample the combinatorial space in a single experiment [271]. For example, from a library of just 16 strains, generating all subsets of size 1 to 7 in a single medium would require, approximately, 160,000 liquid handling steps and three hundred 96-well plates (without replicates). As these combinations could not be prepared in advance and would have to be assembled on the timescale of cell division (one hour), generating even 10% of these combinations would likely be logistically impractical. Indeed, combinatorial studies conducted in liquid

[28]Kehe J, Kulesa A, Ortiz A, Ackerman CM, Thakku SG, Sellers D, Kuehn S, Gore J, Friedman J, Blainey PC. Massively parallel screening of synthetic microbial communities. Proceedings of the National Academy of Sciences. 2019 Jun 25;116(26):12804-9.[259]

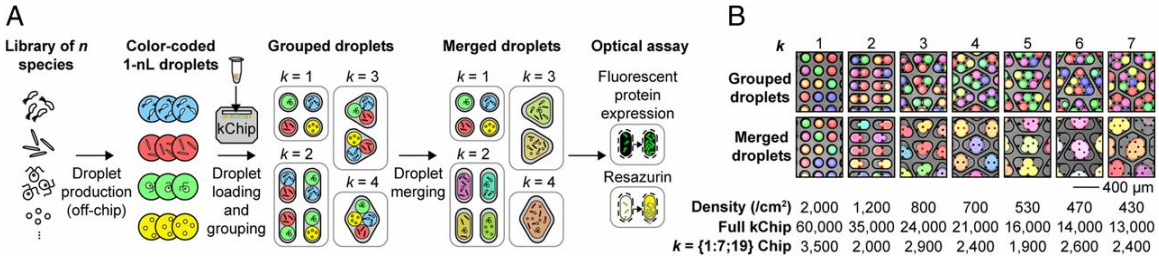

Figure 1-2: **kChip enables massively parallel construction of microbial communities.** (A) To run a kChip screen, 1-nL droplets are first produced. Each droplet contains a color code (a combination of three fluorescent dyes in a specific ratio) that maps to the corresponding input. Afterwards, droplets are pooled and loaded onto the kChip, where they randomly group into microwells. The microwells are designed to group precisely k droplets. The kChip is imaged to identify the contents of each microwell from the droplet color codes. Droplets are then merged within their respective microwells via exposure to an alternating-current electric field, generating parallel synthetic communities. Community phenotypes can be tracked via optical assays including fluorescent protein expression (e.g. GFP) and respiration-driven reduction of resazurin to the fluorescent product resorufin (Figure 5-7). (B) Example micrographs show grouping and merging of droplets for each microwell type. Microwells are densely packed on the kChip, with k dictating microwell density. A single microwell type can be arrayed across a kChip ("Full kChip"). For screening application reported in Figures 3-1, 3-2, and 5-8, we've generated a $k = \{1{:}7;19\}$ Chip that includes different microwell types arranged in parallel.

media typically construct $<10^3$ unique synthetic communities [27, 64, 232, 272, 273]. Some of the largest combinatorial studies [160, 161, 274] instead used the previously discussed Burkholder plate assay, whereby an array of n microbial species is laid on top of an agar plate previously inoculated with another species. Those studies have measured up to $\sim 10^3 - 10^4$ interactions, but they are typically restricted to binary compositions. Diffusion between colonies further places an upper bound on the density of the colony array and the throughput of the screen.

Kehe et al. (2019) [259], following from Kulesa et al. (2018) [275], introduce a platform called the kChip that addresses experimental scale and setup time requirements to assay microbial community function in high throughput. The kChip system (Figure 1-2) enables parallel construction and quantitative screening of up to 10^5 synthetic microbial communities per day and requires no robotic liquid handling. The platform screens n-multichoose-k combinations, meaning each parallel community is

composed of precisely k inputs (e.g. strains or media) randomly selected (with replacement) from a larger library of n, where both n and k are selected by the user. Drawing on advances in micro-compartmentalization [276, 277, 278, 279], the kChip platform generalizes a high-density microwell array approach that groups and merges sets of nanoliter droplets that each carry input components, an approach we previously demonstrated for pairwise combinatorial compound screening [275]. Further developments to the platform have enabled droplets to self-assemble randomly into groupings of $k = \{1, 2, \ldots, 7, 19\}$ dictated by microwell geometries (Figure 1-2B), greatly reducing the time and logistical complexity of combination assembly. As with other droplet microfluidic systems, the kChip platform is amenable to fluorescent assays and uses small assay volumes that conserve valuable inputs. Furthermore, the length scale of kChip microwells (100-1000 µm) is a natural ecological scale for interaction-driven microbial community assembly [15, 280, 281, 282].

1.4.1 The kChip constructs community sets of controlled size

To generate synthetic communities from a library of n input strains, the kChip groups a random subset of k inputs into each microwell. Each kChip contains tens of thousands of microwells where each microwell contains a random grouping of k species. Multiple kChips and/or values of k can be used in accordance with the desired size, number, and replication of combinatorial groupings.

The setup for kChip screening requires three steps: (1) droplet generation[29] and pooling (10 min), (2) droplet loading and grouping (20 min), and (3) droplet merging (10 seconds) (Figure 1-2). Prior to droplet generation, a "color code", or unique ratio of three fluorescent dyes, is mixed with each input. Each color code is therefore packaged with its input when droplets are initially generated and can be used to identify droplet contents [275]. Color-coded droplet sets are then pooled together to form a droplet library. The droplet library is loaded onto the kChip in a single pipetting step. Droplets spontaneously self-assemble into random groupings of k droplets determined by the size, shape, and internal design features of the microwells

[29]Bio-Rad QX200 Droplet Generator, which produces 20,000 1-nL droplets per 20-µL input

(Figure 1-2). The kChip is imaged (2X magnification) to identify contents of each microwell from the droplet color codes (Figure 1-2). The droplets in each microwell are subsequently merged to combine their contents via exposure to an alternating current electric field generating thousands of parallel n-choose-k synthetic communities.

Microwells that group different numbers of inputs can be combined on a kChip in any organizational pattern chosen by the user. Microwell density on a kChip decreases as k increases (Figure 1-2), so the total number of assay points on a given kChip depends on the microwell layout (∼13,000 if all microwells are k = 7; ∼60,000 if all microwells are k = 1). Our working kChip, used for screening applications described below, has microwells that accept up to 7 or 19 droplets ($k = \{1:7; 19\}$) with roughly even representation of each microwell type to enable simultaneous construction and assessment of communities of different richnesses.

1.5 Intestinal and host-associated microbiotas

The appearance of animals and plants on Earth occurred within an already highly diverse and sophisticated microbial world [283]. Long before eukaryotes appeared, major ecological and evolutionary forces had generated tight microbial communities capable of growing and colonizing every viable environment. For example, photosynthetic Cyanobacteria can associate with purple sulfur bacteria (Gammaproteobacteria) to build stratified microbial mats that resemble ancient Precambrian[30] stromatolites [284, 14]. Over the last billion years, as eukaryotes acquired organelles [285] and developed novel tissues and surfaces, microbial communities also developed to colonize these new environments. Two centuries of microscopy and two decades of next-generation sequencing have ascertained that microbial communities abound in nature, and they specially associate to animal and plant hosts. The currently emerging microbiome sciences strive to explain these host-associated microbiotas from the points of view of biochemistry, genetics, biodiversity, and assembly. My research

[30]The Precambrian spans from the Earth's formation, about 4.6 billion years ago, to the beginning of the Cambrian Period, about 541 million years ago. Stromatolite formation peaked about 1.25 billion years ago.

will delve specifically on the assembly process of simple intestinal microbiotas in *C. elegans*, but in this section I will give you a more encompassing overview of the microbiome sciences.

The development, nutrition, and immunity of animals are strongly dependent on their gut microbiotas [283]. Although we recently developed a clear picture of the microbial species present in healthy gut communities [7], we still lack understanding on how do they assemble and how do they break down. An early review on the human microbiota assembly recognizes the importance of the deterministic, stochastic, and history-dependent processes [122] that I introduced earlier. At broad phylogenetic levels, a developed microbiota resembles conceptually the *climax community* of succession theory [10, 12], but at finer phylogenentic levels, there exists a large inter-individual and temporal variability [122, 138]. The observation of similar healthy microbiomes across individuals yet variable dysbiotic communities can be explained as a transition from stable to unstable community states, and this phenomenon has been named the "Anna Karenina principle"[31] [286, 287]. A multi-omics approach has been suggested and embraced as the means to explain in more detail our microbial commensals' function and mechanism of action [288]. Sequencing technologies have expanded our view on microbial diversity by revealing vast amounts of unculturable microbes, or microbial dark matter [289]. Microbial species that have never been isolated can now be modeled with their metagenome-assembled genomes (MAGs) [290] and correlations between these MAGs can be used to infer microbial interspecies interactions at large scale [226].

Animal model systems have emerged to bridge the divide between sequencing-based and culturing-based approaches in microbiome studies [291]. Holistic sequencing-based approaches can *infer* interactions between all microbiome members, while experimental culturing-based approaches can *quantify* interactions between selected microbiota members. Useful model systems like mice, zebrafish, and insects have emerged as alternatives to study host-microbiota interactions and bridge this sequencing-

[31]Paralleling Leo Tolstoy's dictum that "all happy families look alike; each unhappy family is unhappy in its own way."

culturing divide [166]. Given the lower anatomical and spatial complexity of small animals, simple yet semi-complete microbiotas can be constructed within them to test hypothesis experimentally. The assessment of the microbiome of many different insects has shown Proteobacteria and Firmicutes as their most abundant phyla [292]. Interestingly, Proteobacteria is also one of the most unstable phyla in the human gut, and its overabundance is associated to obesity and other dysbiotic states [244]. Although Proteobacteria is found in the mammalian gut at abundances below 2%, its overabundance could be a trigger unbalancing the Firmicutes-to-Bacteroidetes ratio that is characteristic of dysbiotic states [293, 294]. Rigorous testing of such an hypothesis in humans is unfeasible, but mice and other model systems offer practical experimental opportunities to dissect hypothesis like this and explore microbiota assembly [295].

Intestinal communities have well defined spatial and temporal boundaries, which are dictated by the host life history. It is interesting how this defined boundaries are opposite to the view in traditional ecology of there being "no single correct scales at which to view ecosystems" and communities being "just an arbitrary subdivision of a continuous gradation of local species assemblages" [22]. The study of bacterial colonization of the digestive tract has shifted over the last twenty years from a genetics and culturing approach [212] to a sequencing and inference approach. The tractable intestinal microbiotas of animal model systems have opened new avenues to combine these frameworks and explore cause and consequence between microbiota status and host health.

1.5.1 Human microbiota assembly

The composition of the human gut microbiota can change dramatically based on host diet, and such microbiota changes have been linked to a wide variety of diseases and disorders[32]. One influential study linking human diet and microbiome composition utilized 16S sequences of the microbiota of young children from multiple countries

[32]Including Parkinson's disease, autism, depression, obesity, diabetes, ulcerative colitis, cardiovascular disease, and multiple cancer types [296]

to suggest that a cereals-heavy diet might increase Bacteroidetes' abundance and induce healthier and more diverse microbiomes [297]. Human adults can also display changes in their microbiome after strong diet shifts; for instance, a bacterium from the genus *Prevotella* associated to fiber degradation decreases in abundance when vegetarians switch to meat diets [298]. Diet shifts can also induce prompt changes in the abundances of Firmicutes and Bacteoidetes in mice [294]. The production and availability of short-chain fatty acids like propionate [299] and butyrate [300] have been proposed as mechanisms by which the gut microbiota influences host health.

Apart from diet, the human microbiota has also been shown to vary due to successional and inhibitory dynamics between the microbial community members. Time-series metagenomic data of adult diarrhea patients shows an initial decrease of Bacteroidetes and Firmicutes, followed by a transient increase of phages [301]. Viral sequencing data from the gut (virome) of newborns has shown a stepwise assembly modulated by breastfeeding, where initial bacterial phages (prophages) are succeeded by human phages during the first months after birth [302]. Some microbiota members have also been shown to interact with the host and other microbial species and lead to healthier human states. For example, asthmatic children containing a *Corynebacterium* in their upper respiratory tract have lower probabilities of asthma attacks [303]. Also, a *Lactobacillus* isolated from the human nose has been shown to inhibit the growth of pathogenic bacteria from the genera *Staphilococcus*, *Haemophilus*, and *Moraxella* through in vitro and in vivo co-culturing experiments [304]. Other synthetic microbiotas have been isolated from the human gut [273] and human urinary infections [5], and their build-up in laboratory settings has revealed important roles for interspecies interactions in their ecological dynamics.

Intergenerational transmission of microbes has lead to an adaptation between the human host and its microbiota [305], but defining which microbiota members have coevolved with the host is currently under debate [119]. Strong vertical transmission of microbes is known to occur through the birth canal [306] and breastfeeding [302]. Some evidence also suggests that the human intestinal microbiota is seeded before birth [307]. Although microbes can be found in the tar-like substance lining

the intestines of fetuses and newborns (meconium), the mechanism and timing of bacterial colonization are still unclear. Within-person adaptive evolution has also been described for populations of *Bacteroides fragilis*, which accumulate SNPs and HGTs across years of inhabiting their human hosts [308].

The mouse model system has been thoroughly used as an experimental bridge to understand the complex mammalian gut. For example, microbiotas from diverse habitats, including soils, humans, termites, and microbial mats have been fed to mice to build compounded foreign microbiotas [309]. These xenomicrobiotas have shown a gut environmental filtering enriching for Firmicutes from the Clostridia class. When mice are fed with unhealthy fiber-free diets, bacteria capable of degrading mucus increase in abundance, and the mucus barrier decreases in thickness [310]. Laxatives can cause mice to shed their gut mucus barrier and loose certain bacterial families [311]. It has been reported that an oral pathogen reaches higher abundances when co-infecting mice abscesses with a microbe from a different environment (allopatric), likely due to an expansion of the metabolic capacity of the bacterial pair [312]. Overall, great advancement has been made towards understanding the microbial colonization of the complex human and murine microcosms.

1.5.2 *C. elegans* as a tractable model system

C. elegans and other animal model systems display simpler physiological characteristics than mammals, which allow for easier tractable microbiota studies [313]. *C. elegans* is a non-pathogenic transparent roundworm from the phylum Nematoda of about 1 mm in lenght, and it is usually found in temperate soils enriched with organic matter [314]. *C. elegans* develops from eggs to adulthood through four larval stages, it is either hermaphrodite or male, and each worm lays an average of 300 eggs in the short time frame of 60 to 120 hours old [315]. The formation and dynamics of the *C. elegans* intestine and pharynx organs have been thoroughly characterized at the molecular level [316]. For example, *C. elegans* creates an acidic intestinal environment through the action of proton pumping V-ATPases [317], and its minute-long half-maximal-volume defecation cycles [318] generate strong pH oscillations [319].

The functioning of the *C. elegans* rectal canal [320] and the bi-nucleation of its epithelial cells [321] have been linked to its microbiota. The *C. elegans* physiological aspects of most relevance to microbiota assembly include its bacterial (bacteriovore) diet, the grinding action of its pharynx, and its innate immunity [112, 32]. Review articles have described the links between the worm microbiota and its health and aging processes [322], as well as host-microbe and microbe-microbe interactions [323].

Over the last two decades, several studies have shown the killing of *C. elegans* by many important microbial pathogens. These include *Pseudomonas aeruginosa* [324], Serratia marcescens [325], enterotoxigenic *E. coli* [326, 327], different *Salmonella* serovars [328, 115], gram-positive bacteria like *Staphylococcus aureus* [329] and *Microbacterium* [330], yeasts like *Cryptococcus neoformans* [331] and *Candida albicans* [332, 333], and the intracellular pathogen *Microsporidia* [334, 335], among many others [336, 323]. These studies have revealed infection mechanisms of great relevance for medical applications. The p38 mitogen-activated protein kinase (MAPK) pathway [108, 337, 338], the insulin-like signaling pathway [109, 339, 113], and the transforming growth factor beta (TGFβ) [110, 340, 111] have been shown to activate the *C. elegans'* immune response that counteracts intestinal bacterial proliferation. Several components of these immune pathways are conserved between *C. elegans* and mammals, which opens a path for translational research[33] [341]. The antimicrobial effector molecules under the regulation of these signaling cascades include enzymes to break bacterial cell wall, lipids, and sugars, such as lysozymes, lipases, and lectins [342, 343, 344, 345, 346]. My results in section 2.2.6 argue that although the *C. elegans'* p38 MAPK pathway reduces bacterial loads, it has little effect on the microbiota composition.

Apart from its immune system, *C. elegans* interacts with its microbiota through other mechanisms. *C. elegans* can learn behaviors from repetitive chemical and environmental cues [347], and it has been shown that larvae are capable of choosing bacteria on agar plates based on the food quality [348] and the pathogenicity of the bacteria [349]. Such preferential feeding behavior has not been reported in well-mixed

[33]Translation of basic scientific findings in a laboratory setting into potential disease treatments.

conditions. Other *C. elegans* behaviors that could influence its microbiota are social aggregation during feeding [350], the choosing of starvation in exchange for mates [351], and the neuron-dependent slowdown in movement when food is available [352].

Gel-forming mucins are known to cover the intestinal epithelium of mammals, where they act as a dynamic barrier to bacterial colonization [353]. Proteins resembling mammalian mucins are also found in *C. elegans* [354]. Interestingly, a recent study has shown that the mucin MUL-1 is exploited by *P. aeruginosa* to infect the worm [355], which expands our view on how mucins play a role in host colonization.

In the last ten years, an ecology and evolution point of view has been adopted to study the *C. elegans* microbiota. First, experiments with two microbial colonizers showed that *Salmonella* can establish a population that is resistant to invasions by *E. coli* [66]. It was also shown that *Lactobacillus acidophilus* [356] and *Acinetobacter* [332] can ameliorate other pathogens' infections. These co-culturing studies hinted at the importance of interspecies interactions for the worm microbiota assembly. Afterwards, studies co-propagating (co-passaging) worms and microbes over multiple generations showcased evolutionary transitions capable of producing more competitive bacteria [66], or also more pathogenic [357], more protective [358], and more defensive bacterial species [359] depending on the selection regime.

Adaptations between worms and microbes can occur under tight laboratory selection regimes, but do adaptations between worms and microbes happen as well in natural settings? With this question in mind, a last wave of studies directly[34] isolated microbes from *C. elegans'* guts after worms had been newly isolated from compost heaps [361] or European topsoil [362]. It was then reported that a *Pseudomonas* and a *Bacillus* isolated from worms can prevent infection by *P. aeruginosa* [361, 363]; natural bacterial isolates can also influence the development of *C. elegans* [364]; and a sysnthetic microbiome with 14 natural isolates increases the larval output of *C. elegans* compared to the typical *E. coli* OP50 microbiota [365]. Further exploration of the biodiversity in the *C. elegans* natural or native microbiota have taken a biogeog-

[34]The recollected worms have been grown on a complex mixture of microbes over several generations in the laboratory before isolating bacterial strains [360].

raphy approach to describe the differences and similarities to the feeding substrate [41, 230, 231]. In section 2.2.4, I will further explore the competitive ability of some of these native bacterial isolates.

This is the end of this thesis' introduction. First, I gave a layout of the main drivers in ecology and evolution. Secondly, I explained bacterial interspecies interactions from a conceptual and an experimental point of view. And third, I introduced a novel microfluidic device and an animal model system to explore bacterial interspecies interactions at large scale. In the next two Chapters, I will outline my findings on the prevalence of interspecies interactions and the effects they have on microbiota assembly.

Chapter 2

Interspecies bacterial competition regulates community assembly in the *C. elegans* intestine

In collaboration with Nicole M. Vega and Christoph Ratzke[1]

From insects to mammals, a large variety of animals hold in their intestines complex bacterial communities that play an important role in health and disease. To further our understanding of how intestinal bacterial communities assemble and function, we study the *C. elegans* microbiota with a bottom-up approach by feeding this nematode with bacterial monocultures as well as mixtures of two to eight bacterial species. We find that bacteria colonizing well in monoculture don't always do well in co-cultures due to interspecies bacterial interactions. Moreover, as community diversity increases, the ability to colonize the worm gut in monoculture becomes less important than interspecies interactions for determining community assembly. To explore the role of host-microbe adaptation, we compare bacteria isolated from *C. elegans* intestines and non-native isolates, and we find that the success of colonization is determined more by a species' taxonomy than by the isolation source. Lastly,

[1]Ortiz A, Vega NM, Ratzke C, Gore J. Interspecies bacterial competition regulates community assembly in the C. elegans intestine. The ISME Journal. 2021 Feb 15:1-5.[366]

by comparing the assembled microbiotas in two *C. elegans* mutants, we find that innate immunity via the p38 MAPK pathway decreases bacterial abundances yet has little influence on microbiota composition. These results highlight that bacterial interspecies interactions, more so than host-microbe adaptation or gut environmental filtering, play a dominant role in the assembly of the *C. elegans* microbiota.

2.1 Introduction

Bacterial communities are found almost everywhere in nature [367]. Among the many ecosystems in which bacterial communities play a fundamental role, the animal digestive tract is one of remarkable importance [368, 283],. These large, complex, and highly organized bacterial consortia [253] can degrade food and deliver nutrients to their host [369], protect against invading pathogens [370, 363], and even produce neurotransmitters that affect host behavior [371].

Despite considerable efforts toward elucidating the composition and function of these intestinal bacterial communities [372, 7], the rules that govern their assembly are still not fully understood [124]. Recent studies have taken advantage of animal model systems, such as mice [309], zebrafish [373, 208], honey bees [6], flies [374, 168], and worms [360], to experimentally address the composition and assembly of simpler gut microbiotas [166]. A recurrent explanation for the assembly of these communities is that the gut can strongly filter the bacterial colonizers and select for a core microbiota [375, 292]. If such environmental filtering is sustained over evolutionary timescales, an adaptation between hosts and microbes can occur and a symbiosis can develop [376], but not all associations between hosts and microbes are indicative of adaptation or co-evolution [377]. The competition assays via co-culturing microbes [6] and the bottom-up assembly of microbiotas [149] that are possible in animal model systems provide an opportunity to test which forces influence microbiota assembly (Figure 2-1A).

The nematode *C. elegans* is a good model system to study gut microbiota assembly [323]. Multiple human pathogens are also pathogens for *C. elegans* [323, 322], and the

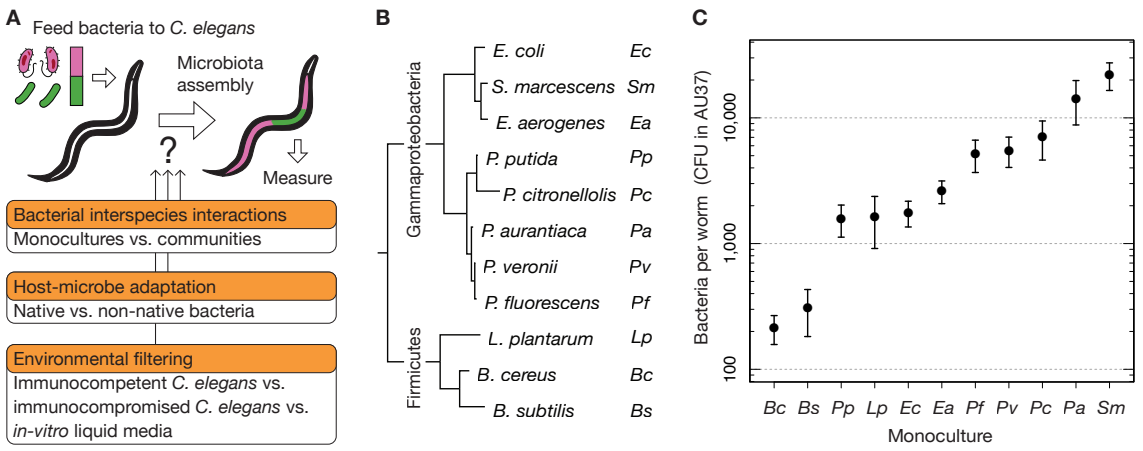

Figure 2-1: **Different bacterial species reach widely different population sizes in *C. elegans* gut.**(A) Diagram of the *C. elegans* microbiota assembly and the three biological forces (orange) that might influence this process and that we study in this article. To construct and measure simple microbiotas in *C. elegans*, a defined number of bacterial species are fed in liquid culture to a same-age adult population of *C. elegans* previously sterilized with antibiotics. The liquid feeding substrate is restored every day to maintain equal bacterial concentrations during the four days of colonization. Afterwards, worms are mechanically disrupted in batches of ~ 20, and counts of colony forming units (CFU) are used to determine bacterial population sizes in the worm gut. (B) Phylogenetic tree from full-length 16S rRNA gene sequences of the 11 non-native bacterial species used to colonize the gut of *C. elegans*. (C) Bacterial population sizes in monoculture colonization of immunocompromised *C. elegans* (AU37) span two orders of magnitude. These population sizes reflect the inherent abilities of bacteria to colonize the worm intestine environment. Points are the average of 8 or more biological replicates, and error bars are the standard error of the mean (s.e.m.).

longevity [116] and reproduction [378, 364], of this worm are linked to its microbiota. Furthermore, the *C. elegans* gut environment filters the larger bacterial pool found in natural feeding substrates, leading to the assembly of a core microbiota [41, 365]. Such environmental filtering can occur via behavioral food-avoidance [349], ingestion rates [379, 32], and host mucins [355], among other factors [363, 323]. Recent reports have suggested that the worm's innate immunity [113, 111] and microbe-microbe interactions [230] play a dominant role in the *C. elegans* microbiota assembly, but an experimental comparison of the many forces in play is still lacking.

In this study, we colonized *C. elegans* with simple microbiotas to determine the effect that bacterial interspecies interactions, host-microbe adaptation, and environ-

mental filtering have on the underlying assembly process (Figure 2-1A). We found that the ability of a bacterial species to colonize the worm gut in monoculture was often inadequate for predicting the relative abundances of two-, three-, and eight-species microbiotas. Additionally, in experiments with bacteria not isolated from *C. elegans* (non-native), we found that the fractional abundance of two-species microbiotas can be used to predict the composition of three-species microbiotas, indicating that assembly rules based on pairwise interactions [27] can provide insight into the composition of gut microbiota communities. Finally, *C. elegans* and its feeding substrate can reach different stable states, and the acidic pH of the worm gut may be a component of the environmental filtering by this host during community assembly. With this, we advance our understanding of the polymicrobial colonization of the *C. elegans* gut and provide insight into bacterial community assembly within a host.

2.2 Results

2.2.1 Monocultures differ significantly in their ability to colonize the *C. elegans* intestine

To investigate community assembly in the gut of *C. elegans*, we fed germ-free synchronized adult worms with different bacterial species, in monoculture or in mixture, over four days in a well-mixed rich liquid medium (Methods 2.4, Figure 5-1A[2]). The majority of worms survived the four-day period of feeding and colonization, after which we allowed live worms to feed briefly on heat-killed *E. coli* OP50 to remove transient colonizers [32, 361]. We then cleaned the surface of the worms with consecutive washes, and measured the intestinal bacterial densities by grinding batches of worms, plating, and counting colony forming units (CFU, Figure 5-1B) with distinct morphologies [380]. The supernatant of each sample was plated to verify that CFU counts came from the worm digestions instead of the background media (Methods 2.4). This protocol allowed us to construct and quantify simple microbiotas in *C.*

[2]Supplementary information like this can be found in Chapter 5

elegans.

We began by feeding *C. elegans* in monoculture to quantify the ability of a range of bacterial species to colonize and grow in the worm intestine. As a starting point, we first utilized an immunocompromised *C. elegans* mutant (AU37) and a set of eleven non-native bacterial species (Figure 2-1B), representing the phyla Firmicutes (gram-positive) and Proteobacteria (gram-negative). We found that all bacterial species colonize (i.e. accumulate with or without active growth) the *C. elegans* intestine, with mean population sizes (Figure 2-1C and 5-1C) ranging from 200 CFU per worm in the case of *B. cereus*, up to 20 000 CFU/worm in the case of *S. marcescens*. Our three Firmicutes reach low population sizes in the worm gut and low carrying capacities in the liquid media (Figure 5-1E), but the carrying capacities in the liquid media don't explain the variation in monoculture colonization (Figure 5-1F-G). These results indicate that different non-native bacterial species have a wide range of abilities to colonize the *C. elegans* intestine in monoculture.

2.2.2 Composition of two-species microbiotas are influenced by competitive and hierarchical bacterial interspecies interactions

To assess the compositional trends of the *C. elegans* microbiota, we constructed the simplest intestinal communities in this worm by feeding it with all possible two-species mixtures from the same eleven non-native bacteria as before (55 pairs, Figure 2-2A and 5-2A). We fed worms with both bacteria present at similar concentrations ($\sim 10^7$ CFU/mL, Methods 2.4) to normalize the rate of ingestion. We found that a majority (41 out of 55, $\sim$75%) of pairs displayed coexistence, with both species present above the detection limit of 2%, whereas the remainder (14 out of 55, $\sim$25%) led to competitive exclusion of a species (Figure 2-2B and 5-2B). These results show that bacteria with no prior conditioning for the *C. elegans* gut commonly reach coexistence in two-species microbiotas.

The interactions between bacterial species in a microbiota can be classified as pos-

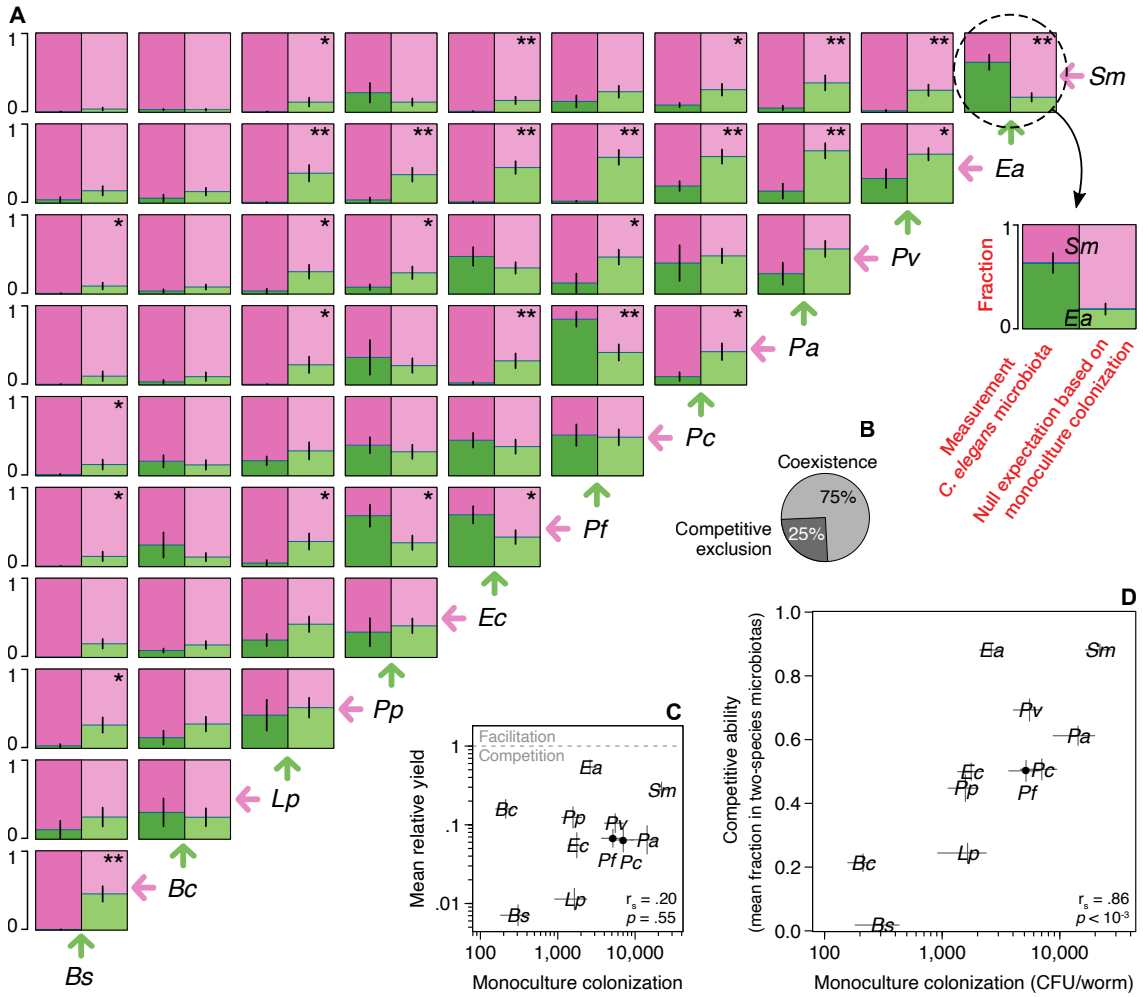

Figure 2-2: **Monoculture colonization of the worm intestine often fails to predict composition of two-species microbiotas.** (A) LEFT panels: Fractional abundances of 55 co-culture experiments in *C. elegans* intestine (AU37); error bars are the s.e.m. of 2 to 8 biological replicates (Figure 5-2). Bacterial species are ordered from left to right by their mean fraction across all co-cultures. RIGHT panels: Null expectation for the fractional abundances based on a non-interacting model where each bacterial species reaches its population size in monoculture; error bars are the s.e.m. from bootstrapping over the monoculture data. * and ** represent a statistically significant difference between the two panels at p-values of .05 and .01, respectively (Welch's T-test). (B) Coexistence of two species is more common than competitive exclusion in the worm intestine. (C) Low yields in two species microbiotas –relative to monocultures– are indicative of competitive interactions (Figure 2-2); error bars on X-axis are the s.e.m. and on Y-axis the s.e.m. from bootstrapping over monoculture and pairwise data simultaneously. (D) Competitive ability, defined as the mean fractional abundance in co-culture experiments, relates to monoculture population size, but there are significant deviations; error bars on Y-axis are the propagated error from the s.e.m. of the co-culture experiments.

itive, negative, or neutral based on the yields of the bacteria relative to their monoculture population sizes. To classify the interactions in our two-species microbiotas, we calculated the relative yield of species 'i' with species 'j', $RY_{i|j}$, as its population size in co-culture, $N_{i|j}$, divided by its population size in monoculture, N_i ($RY_{i|j} = \frac{N_{i|j}}{N_i}$, see Methods 2.4 for detailed implementation). We found that most species cannot reach their monoculture population size in co-culture experiments, $RY < 1$ (Figure 2-2C, 5-2D-E), suggesting that interactions between species are largely competitive [208, 232, 210]. From our 110 RY measurements, only *Ea* co-cultured with *Pa* reached a RY significantly greater than 1. Furthermore, most co-cultures reach lower community sizes than the higher population size of the monocultures (Figure 5-2D), indicating that the observed low relative yields are not simply due to competition for fixed space within the worm gut. Interestingly, the mean relative yield of each species does not correlate with its monoculture colonization (Spearman correlation r_s=.20, p=.55, Figure 2-2C), indicating that a large population size does not protect a bacterial species from being harmed by competition in co-culture experiments.

To explore the influence of monoculture colonization ability, which does not depend on interspecies interactions, on the assembly of the *C. elegans* microbiota, we further compared the monoculture population sizes and the two-species microbiotas. We first simplified the worm pairwise outcomes into a summary-metric by calculating the mean fractional abundance of each species in all two-species microbiotas, $\langle F_i \rangle_{\forall j} = \langle \frac{N_{i|j}}{N_{i|j}+N_{j|i}} \rangle_{\forall j}$. We found that this competitive ability score correlates to the population size reached in monoculture colonization (Figure 2-2D, r_s=.86, p<10-3). This positive relationship indicates that a bacterial species persists in two-species microbiotas due to similar properties to those favoring its monoculture colonization of the gut. This, together with the previous result, indicates that the uneven harm caused by competition doesn't dramatically alter the mean fractional abundance expected from monocultures.

Despite this correlation between monoculture colonization and competitive ability, some species in co-culture performed differently than would be expected based simply on population sizes. For example, *Ea* tied with *Sm* for being the strongest competitor

despite being only the sixth highest colonizer in monoculture (Figure 2-2D). We therefore sought to determine to what extent interactions between microbes are important to predict the pairwise outcomes in the host. We calculated a null expectation for the two-species microbiotas assuming that each species is able to reach the carrying capacity that was measured in monoculture colonization (Figure 2-2A, right panels). By comparing this null expectation with the experimentally measured fractions obtained from pairwise colonization (Figure 2-2A, left panels), we are able to identify the cases in which interspecies interactions play a dominant role in determining the composition of the gut microbiota [232, 167]. In 28 out of 55 cases this deviation is large enough to reject the null model ($p<.05$, Figure 2-2A), many more than the 2.75 cases expected by chance at this significance level (16 cases with $p_{FDR}<.05$). These results indicate that a null expectation, where each species' abundance in pairwise colonization is determined by its monoculture fitness to the worm gut environment, rather than by interactions between bacteria, fails to predict a significant number of two-species microbiotas.

To further characterize the structure of the competition network, we quantified its degree of hierarchy [203], which estimates how frequently a highly ranked competitor will dominate a lower-rank adversary. The hierarchy score of this network, 0.82, is significantly larger than the hierarchy score found in random matrices with the same distribution of fractional abundances ($p<10^5$, Supplementary), suggesting that there is an approximate ordering of the competitive abilities of these bacterial species in the worm gut. Consistent with this ordering, we do not observe any cases of intransitive competition, in which the pairwise interactions of three bacterial species would be analogous to the rock-paper-scissors game and no absolute winner would exist [198, 381]. This intransitivity has been proposed as a major mechanism inducing coexistence in natural populations [214, 209, 53, 382, 383], but we do not observe it in the pairwise interactions of any of our 165 hypothetical trios. With a more relaxed definition of intransitivity, in which a species wins a competition by being more abundant than the competitor instead of needing to fully exclude it, we find two candidate trios with a rock-paper-scissors-like structure: *Ec-Pf-Pa* and *Pp-Pf-Pa*

(although the dominance of some competitors is not statistically significant).

Collectively, we find that the monoculture colonization ability of a bacteria correlates with its mean abundance in two-species microbiotas. Additionally, the interspecies interactions, which are mostly competitive and hierarchical, alter the composition of at least half of the individual two-species microbiotas (Figure 2-2).

2.2.3 Three-species microbiotas are predicted by pairwise outcomes, not by monocultures

In the two-species microbiotas, we observed frequent coexistence of non-native bacteria in the *C. elegans* gut, but it remains to be tested if coexistence is also the norm in gut communities initialized with a larger number of species. We therefore constructed 20 three-species microbiotas in *C. elegans* to extend our analysis. From our eleven non-native bacterial species, we selected a set of six (*Bc, Lp, Pf, Pv, Ea,* and *Sm*) that span the range of competitive abilities observed, and we fed them in all possible trios to *C. elegans* (Figure 5-3A). In the trio *E. aerogenes-P. fluorescens-S. marcescens*, for example, we observed coexistence of the three species, in which *Ea* is the majority, *Sm* is close second, and *Pf* is a minority (55±7%, 42±5%, 3±3%, respectively; mean ± s.e.m.). These fractional abundances can be represented as a stack of bars (Figure 2-3A) or as a point in a simplex (Figure 2-3B), where the point moves closer to a vertex when that given species increases in abundance. By plotting the measurements of all 20 trios into one simplex (Figure 2-3C), we observe that most of the cases have one species as highly abundant, yet full exclusion is rare and only accounts for 3 out of the 20 trios tested (Figure 2-3D). Our trio feeding experiments therefore display a range of different outcomes, with frequent coexistence of the three species leading to multispecies gut microbiotas.

To test if monoculture colonization contains the information necessary to estimate the assembly of three-species microbiotas, we made quantitative predictions of the trio outcomes based on monocultures. We extended the null expectation described earlier by assuming that all species will reach their population sizes in monoculture

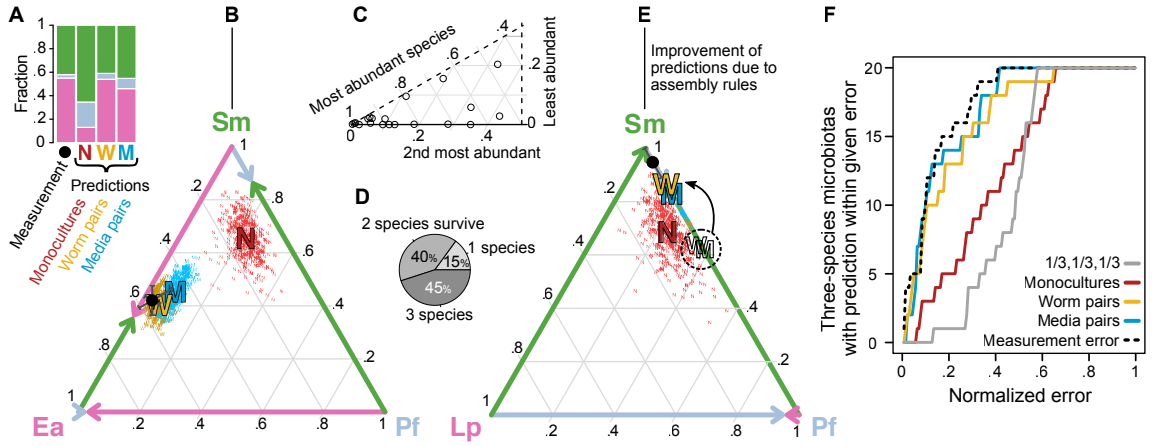

Figure 2-3: **Fractional abundances in three-species microbiotas are well predicted by pairwise outcomes.** (A) Outcome of trio *Ea-Pf-Sm* in *C. elegans* (AU37) intestine, together with predictions based on monoculture population sizes , two-species microbiotas, or pairwise outcomes in vitro liquid media (normalized arithmetic mean). (B) Simplex representation of trio outcome and predictions in (A), with the edges of the triangle depicting the two-species microbiotas in *C. elegans*. The error bars on measurement are the s.e.m. of 4 biological replicates, and the clouds of points around predictions are 400 bootstrap replicates ('N's sampling the monoculture data, and 'W's and 'M's sampling the pairwise data in worm and media, respectively). (C) 20 trio outcomes represented in one sixth of a simplex. (D) 3, 8, and 9 out of the 20 trios show full competitive exclusion, two- and three-species coexistence, respectively. (E) Assembly rules help the quantitative prediction of the trio outcomes based on pairwise outcomes when one of the pairs is competitive exclusion. (F) Cumulative distribution of error of predictions. Error calculated as the linear distance between prediction and measurement in the simplex. The distances are normalized by the maximal distance, $\sqrt{2}$. The dashed line is the mean distance between the measured mean and the 4 biological replicates of each trio, and serves as a lower bound for the error of the predictions.

colonization ('N' Figure 2-3 and 5-3). This null expectation achieves poor results at predicting trio outcomes, for its mean error of 35.7% is just somewhat better than the 43.8% mean error of an uninformed '$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$' prediction (Figure 2-3F). Hence, monocultures are inadequate at predicting three-species microbiotas, highlighting that as community diversity increases, the properties favoring monoculture colonization of the gut are less important than the interspecies interactions for determining community composition.

To determine the role of bacterial pairwise interactions on the assembly of the *C. elegans* microbiota, we made predictions of all three-species microbiotas based on two-species microbiotas. We first calculated a simple linear prediction for each trio by taking the arithmetic mean of each species' fraction in the co-culture experiments against the other two species (a normalization factor of 2/3 is needed for the fractions of the three species to add to one). This normalized arithmetic mean prediction, applied over the two-species microbiotas in the worm ('W' Figure 3A-B and S3A), quantitatively predicts some trios with high accuracy, and exhibits a mean error of 26% (Figure 5-3C). However, this prediction is prone to error (hollow 'W' Figure 2-3E) when one of the two-species microbiotas is competitive exclusion. A recently proposed assembly rule [27] is capable of adjusting these cases by simply removing a bacterial species from the trio prediction when it cannot survive both constituent co-culture experiments (solid 'W' Figure 2-3E). After the application of this assembly rule, the mean error of the predictions based on two-species microbiotas, 18.7%, comes close to the expected biological noise in three-species microbiotas, 13.3% (Figure 2-3F). The fact that two-species microbiotas can properly predict three-species microbiotas indicates that interactions between pairs of bacterial species are an important force in determining microbiota assembly and suggests that indirect higher-order interactions are uncommon or weak.

Our results diverge from three recent findings, where higher-order interactions played a role in the zebrafish microbiota assembly [208], changed an in vitro community amylolytic function [267], and prevented invasion of an algae-bacteria-ciliate community [65]. The discrepancy may be due to the interaction-estimation approach

used there and the simpler rules-based approach [27] used here, and/or due to biological differences in the model systems. However, consistent with what we report here, recent experiments have found that two-species microbiotas are able to predict the fitness traits of flies with multispecies microbiotas [168].

Next, we asked whether it would be possible to make predictions of the 20 three-species microbiotas based on the pairwise outcomes from a different environment, such as the in vitro liquid media used as feeding substrate. Thus, we performed all possible pairwise co-culture experiments in liquid media without worms and measured the equilibrated bacterial fractions after seven cycles of 100-fold daily dilution (Methods 2.4, Figure 5-4, further explored in section below). After applying the assembly rules, the mean error of the predictions based on media pairwise outcomes, 15.7%, also comes close to the expected biological noise of the three-species microbiotas (Figure 2-3F and 5-3A). Since the equilibrium fractional abundances in the liquid media are dependent on the dilution regime [178], our results highlight that our chosen daily dilution of 100-fold can resemble the pairwise outcomes in the worm gut. These results show that three-species microbiotas in *C. elegans* with non-native bacteria can still be predicted with the pairwise outcomes measured in a different environment, suggesting that the environmental filtering of the worm intestine is not the main determinant of community assembly.

2.2.4 Bacterial relative abundance in *C. elegans* microbiota is dependent on phylogeny rather than isolation origin

Thus far we have explored the assembly of the *C. elegans* microbiota with non-native bacteria (Figure 12-1B). These laboratory species were not isolated from any worm gut, and they hadn't been conditioned to grow in the *C. elegans* gut environment before the beginning of each of our experiments. Using these non-native bacteria helped us reduce the complexity associated with microbiota assembly by setting aside the selection force that *C. elegans* could have exerted on its microbiota during evolutionary timescales. Within our non-native species we observed strong and weak colonizers,

strong and weak competitors, and an imperfect relation between these two metrics because of interspecies interactions. It remains to be seen if these observations hold true with bacteria native to the *C. elegans* gut.

Previous studies have assessed the natural *C. elegans* microbiota [360] by isolating and cultivating nematodes in decaying organic matter, and have isolated native bacterial strains by grinding the naturally colonized worms [364, 41, 365, 361]. We isolated new bacterial strains from worms with a similar protocol (CR collection, Methods 2.4) and also utilized the MYb collection from Dirksen et al. (2016) [365, 365], which contains some of the bacteria persistently found in the *C. elegans* gut. From these two collections, we selected a phylogenetically diverse set of bacterial isolates, spanning 4 different phyla, to study the colonization patterns of native bacteria (Figure 2-4A, 12 MYb, 15 CR).

To characterize the monoculture colonization ability of these 27 native bacterial isolates, we fed them in monoculture to wild type *C. elegans* (N2, from where they were isolated) along with 8 of the previous non-native bacteria (Figure 2-4B). We found that the native and non-native bacteria colonize in a similar fashion. Regardless of the origin of the isolate, the Firmicutes colonized poorly and the Gammaproteobacteria often colonized well, resulting in monoculture population sizes that in some cases exceeded 10 000 CFU/worm. By comparing 16S phylogenetic similarity between each pair of bacteria (Figure 5-5A) against their fold-difference in monoculture colonization (Figure 5-5B), we observed that similar bacteria colonize similarly, and as phylogenetic distance increases, the difference in colonization ability tends to increase as well (r_s=.39, p=.003, Mantel test). This positive correlation is true regardless of the native/non-native dichotomy, and can be observed within the genera *Bacillus, Microbacterium, Sphingobacterium, Serratia, Ochrobactrum,* etc. (Figure 2-4A-B). Overall, these results show that monoculture colonization ability is similar between native and non-native bacterial strains, and suggest that this metric depends on evolutionary history rather than isolation origin.

The possible host-microbe adaptation of the native bacterial isolates to the worm gut could have selected for stronger bacterial competitors, besides the selection for

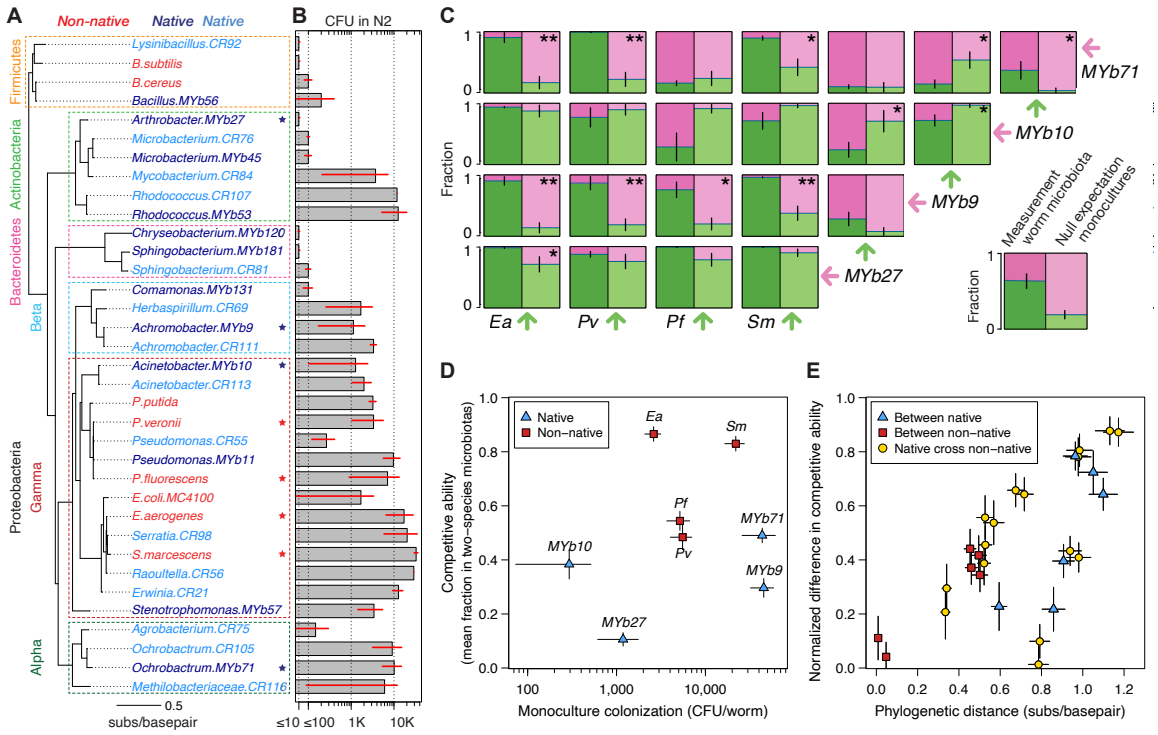

Figure 2-4: **Experimental colonization of *C. elegans* by a wide range of native and non-native bacteria reveals that phylogeny rather than isolation origin determines abundance in the gut microbiota.** (A) Phylogenetic classification of previously shown laboratory species (non-native) and bacterial strains isolated from *C. elegans* intestines (native; dark and light blue from MYb and CR collections, respectively; Methods 2.4). Phylogenetic tree built with maximum likelihood estimate utilizing alignment of full-length 16S gene sequences. The phylogenetic tree is sorted at each internal node to have the higher monoculture colonizers at the bottom. High level phylogenetic classification is given on the left side of the tree for ease of interpretation. Stars indicate bacteria used in follow-up two-species microbiotas. (B) Bacterial population sizes in monoculture colonization of wild type *C. elegans* (N2); error bars are s.e.m. of two to three replicates. (C) Left panels: Fractional abundances in two-species microbiotas with native and non-native bacteria in *C. elegans* intestine (AU37). Right panels: Null expectation for the fractional abundances based on monoculture population sizes. '*' and '**' represent a statistically significant difference between measurement and null expectation at p-values of .05 and .01, respectively (Welch's T-test). (D) Although 2 native strains can reach substantial colonization of the worm intestine in monoculture, these strains reach low fractional abundances in two-species microbiotas. (E) Differences in competitive ability correlate with phylogenetic distances regardless of the isolation origin of the bacteria. Phylogenetic distances are the horizontal distances in the phylogenetic tree. Differences in competitive ability are normalized by the maximum competitive ability of the pair (i.e. competitive abilities 0.8 and 0.4 are as different as 0.2 and 0.1).

health-promoting bacteria that has been previously reported [358]. To investigate how the competitive ability of native bacteria compares to that of non-native bacteria, we performed further feeding experiments in *C. elegans* with all monocultures and pairwise combinations of four native isolates and four non-native species (stars in Figure 2-4A; Natives selected to cover the range of monoculture colonization, expand our phylogenetic diversity, and display distinct colony morphologies: *Achromobacter MYb9, Acinetobacter MYb10, Arthrobacter MYb27,* and *Ochrobactrum MYb71*; Non-native: *Ea, Pv, Pf,* and *Sm*). The measured fractional abundances in immunocompromised AU37 worms showed that the composition of two-species microbiotas once again often deviates from a null expectation based on monoculture colonization (Figure 2-4C, 4/6 cases for between native pairs, and 8/16 for native cross non-native pairs with $p<.05$). In this new set of experiments involving native isolates, more interactions appear to be positive (in the form of parasitisms), with *Ea, Sm,* and MYb10 being facilitated the most (Figure 5-5C-D). Moreover, we found that non-natives reach higher mean fractional abundances than native bacteria in the two-species microbiotas (Figure 2-4D). *Ochrobactrum MYb71 and Achromobacter MYb9* had the largest monoculture population sizes yet low fractional abundances in two-species microbiotas, indicating a low competitive ability, while *Acinetobacter MYb10* showed the opposite characteristics. Further comparison of native and non-native bacteria is warranted, but our two-species microbiotas indicate that native bacteria also interact in the digestive tract of *C. elegans* to structure the microbiota composition (Figure 2-4C). These native isolates lack a clear competitive advantage over non-native bacteria, particularly when co-cultured with strong competitors (*Ea, Sm*; Figure 2-4D).

To test if phylogenetic differences are responsible for the observed differences in competitive ability, we compared these two metrics using our set of native and non-native bacteria (Figure 2-4E). We observed that the competitive abilities of a pair of bacteria differ more as the bacteria diverge phylogenetically (16S gene), regardless of the isolation origin. Controlling for phylogenetic distance, a bacterial strain native to the *C. elegans* gut appears equally different from another native strain or a non-native

strain (Figure 2-4E). Our data therefore suggest that the composition of two-species microbiotas in *C. elegans*, as well as the more basic monoculture colonization, is determined more by a species' phylogenetic classification than by whether the species was isolated from the worm microbiota.

2.2.5 Environmental filtering by *C. elegans* gut can alter pairwise outcomes

Previous detection of bacterial genera enriched in the *C. elegans* intestine compared to the substrate where the worms were grown, such as Ochrobactrum, has suggested an important role for environmental filtering by the *C. elegans* gut during microbiota assembly [231]. Nevertheless, our results, showing that three-species microbiotas are well predicted by in vitro pairwise outcomes ('M' Figure 2-3), suggests that environmental filtering by the worm intestine is not the main determinant of community assembly for this set of bacteria. To further characterize the effect of the *C. elegans* gut environment on bacterial community assembly, we directly compared the outcomes of 55 co-culture experiments between in vivo worm gut and in vitro liquid media (Figure 2-5A and S4A, Methods 2.4). We found that competitive ability (Figure 2-5A, r=.84, p=.0005) and mean relative yield (Figure 5-4B-C, r=.81, p=.0018) in the worm gut and liquid media are correlated, which suggests that the environmental filtering that *C. elegans* provides is not strong enough to alter the hierarchical ordering of these eleven bacterial strains. Although competitive ability is similar between the worm gut and in vitro liquid media, we found that 19 out of 55 bacterial pairs displayed a significantly different outcome in these two environments ($p<.05$, Figure 5-4A,D). From these 19 bacterial pairs, we found nine displaying coexistence in the worm yet competitive exclusion in the liquid media, while zero pairs displayed the opposite trend, indicating that the worm intestine allows for more coexistence. Our data therefore indicates that the *C. elegans* intestine doesn't alter the competitive hierarchy of its bacterial colonizers, but is capable of altering specific pairwise outcomes.

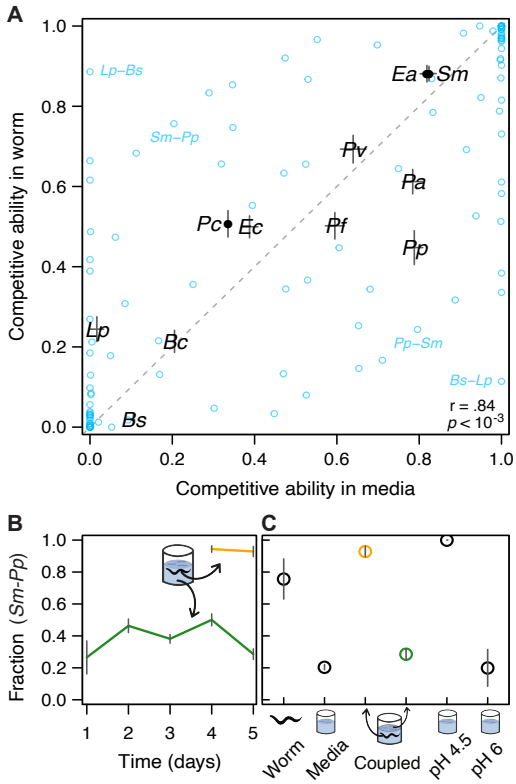

Figure 2-5: **Bacterial interspecies interactions are similar between the in vitro and in vivo environments, with some differences caused by the acidity of the worm gut.** (A) Black points are the mean fractional abundance in co-culture experiments in *C. elegans* intestine and liquid media (1% AXN); error bars are the propagated error from the s.e.m. of the underlying co-culture experiments. Blue points are the outcomes of individual co-culture experiments in worms and media. (B) *S. marcescens* and *P. putida* reach different fractional abundances in vivo worm gut and in vitro liquid media on a coupled experiment, where worms and liquid media from the same test tube are tested. (C) An acidic version of the media resembling the average pH of the worm intestine (4.5) shifts back the pairwise outcome to a worm-like state; error bars are the s.e.m. of at least 4 replicates.

In order to find the features that differentiate the worm gut and the liquid media, we set off to investigate more closely one bacterial pair, *S. marcescens-P. putida*, in these two environments. *Sm* is a majority of the in vitro community (76±12%) yet is a minority in the worm gut (20±2%, Figure 2-5A). We first fed a population of worms with a mixture of *Sm-Pp*, but instead of restoring the liquid feeding substrate every day to maintain equal proportions of both species, we allowed the bacteria to be carried over with the dilution step. In this coupled co-culture experiment,

where strong migration occurs between worms and media, we still observed different compositional states in the different environments (Figure 2-5B). This result shows that the environmental filtering imposed by the worm intestine can be strong enough to keep an internal bacterial community different from its surrounding environment.

Given the importance of pH to microbial growth and competition [367, 9, 218], we tested whether the low pH of the worm gut could cause the persistent difference across environments for the *Sm-Pp* bacterial pair. We therefore repeated the co-culture experiment in liquid media at its normal pH 6, and at lower pH 4.5, where the latter approximates the conditions within the nematode intestine [317, 319]. For this pair of species, lowering the pH of the media was sufficient to alter the pairwise outcome in vitro, resulting in a community very similar to that observed in the worm intestine (Figure 2-5C). Similar results were observed in the pair *Lp-Bs* (Figure 5-4E-F). We conclude that *C. elegans* and its feeding substrate can reach different stable states, and the acidic pH of the worm gut can be an important component of the environmental filtering by this host during community assembly.

2.2.6 Innate immunity of *C. elegans* reduces bacterial loads, but has little effect on microbiota composition

To further our understanding of the effect that *C. elegans* has in the assembly of its microbiota via its immune system, we set off to compare equivalent one- and two-species microbiotas in two *C. elegans* strains with different immunity levels (Figure 2-6). The *C. elegans* strain AU37 is susceptible to high bacterial colonization due to its deleted *sek-1* gene, which encodes for a kinase part of a signaling cascade homologous to the human p38 MAPK (mitogen-activated protein kinase) pathway [108, 384]. The *C. elegans* p38 MAPK pathway activates the production of immune effector molecules, such as lysozymes *lys-2* and *lys-8* [343]. The *C. elegans* strain SS104 has the same *glp-4* mutation as AU37 that leads to sterility at room temperature, allowing us to work with same-size, synchronized worms. While the strain SS104 is not immunologically wild-type—this mutant shows up-regulation of the DAF-2/IGF

pathway, presumably as a by-product of its reproductive sterility [361, 385, 386]—it is the immunocompetent counterpart of AU37 in that it has its wild-type *sek-1* gene and therefore intact signaling through its p38 MAPK pathway. Comparison of the gut microbiota communities in the worm strains SS104 and AU37 therefore allows us to directly study the role of the worms' p38 innate immunity pathway in structuring its gut microbiota.

We first explored the monoculture colonization of the immunocompetent *C. elegans* SS104 by feeding it with the 11 non-native and 4 native bacterial strains shown before. We found that all bacterial species colonize the SS104 intestine, with mean population sizes ranging from 30 CFU/worm in the case of *Bs*, to more than 10 000 CFU/worm for *Pa* (Figure 2-6A and S1C). On average, the bacterial population sizes in the p38-immunocompetent *C. elegans* were approximately 4 times lower than in the immunocompromised AU37 strain, but this reduction in bacterial load was uneven across species. For example, *Ea*, *Pa*, and *Acinetobacter MYb10* seemed to reach similar carrying capacities regardless of p38 MAPK pathway activity, while the population sizes of *Bs*, *Pv*, and *Achromobacter MYb9* were reduced by more than 90% when this pathway was active. These results show that the innate immune system of *C. elegans* via its p38 MAPK pathway reduces population sizes of most bacterial strains, but some colonizers are less affected than others.

In order to test the effect of the *C. elegans* immune system via p38 MAPK on the composition of two-species microbiotas, we fed immunocompetent SS104 with all bacterial pairwise combinations previously studied in immunocompromised AU37 (55 pairs from 11 non-native bacteria, 6 pairs from 4 native bacteria, and 16 pairs from 4 native vs. 4 non-native bacteria; 77 pairs in total; Figure 5-6A-B). We found a strong correlation between the mean fractional abundances in the two worm strains (Figure 2-6B), with *Ea* and *Sm* as the best competitors, and *Bs* and *Arthrobacter MYb27* as the worst competitors in both worm mutants. We also observe a majority of the two-species microbiomes displaying similar abundances in the two worm strains (Figure 5-6A-C). These results show that the *C. elegans* immunity via its p38 MAPK pathway has little effect on the composition of two-species microbiotas.

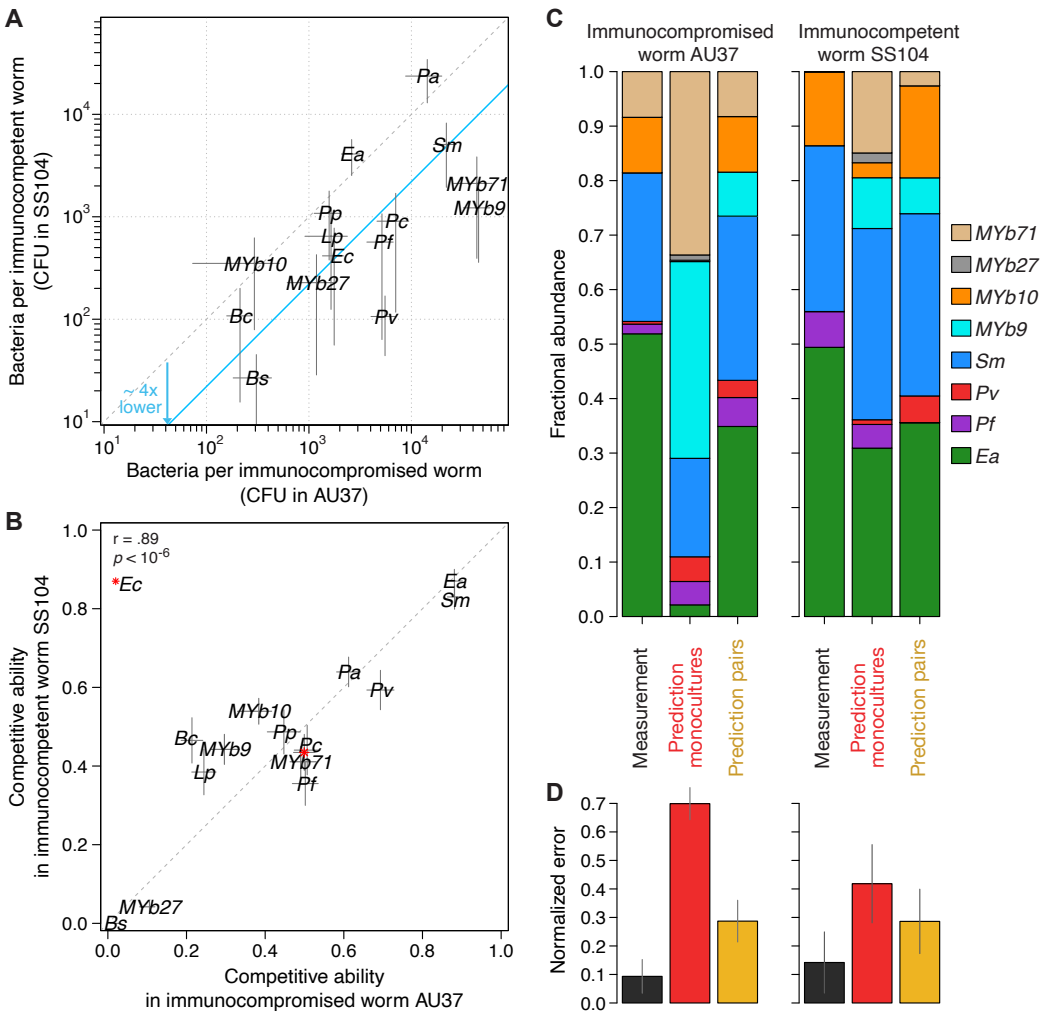

Figure 2-6: **Innate immunity of *C. elegans* via the p38 MAPK pathway reduces bacterial population sizes, but has little influence on the composition of the two- and eight-species microbiotas.** (A) Immune system of *C. elegans* reduces bacterial monoculture population sizes unevenly for different bacteria. Immunocompromised *C. elegans* (AU37) has larger bacterial population sizes in its intestine than immunocompetent *C. elegans* (GLP4). (B) The mean fractional abundances in co-cultures are similar between the two worm strains with different immune activity. (C) Composition of an intestinal microbiota in immunocompromised *C. elegans* AU37 and immunocompetent SS104, together with predictions based on monoculture colonization and pairwise outcomes in the same worm strains. Three or more batches of ~ 20 worms digested for each measurement. (D) Errors as the L1 norm (Manhattan distance) between measurement replicates and predictions. The variability across different batches of digested worms generates a measurement error of 9.3% and 14.2% for AU37 and SS104, respectively. The errors are normalized by 2, the maximum error. Confidence intervals of the prediction errors were calculated by bootstrapping over the corresponding data.

Moving into more complex microbiotas, we built, measured, and compared an eight-species microbiota in our two worm strains AU37 and SS104. We utilized the same set of native and non-native strains explored in pairwise combinations before (Figure 2-4D). After four days of feeding, we measured the bacterial abundances in the ensuing microbiotas (Figure 2-6C). The three most abundant bacterial species were the same in both worm strains: *Ea, Sm*, and *Acinetobacter MYb10*. In accordance with the results of two-species microbiotas, the two *C. elegans* strains with different immunity levels displayed very similar average microbiota compositions (L1 norm distance=1.4%, lower than the L1 norm measurement error for AU37=9.3% and SS104=14.2%, Figure 2-6D). These results show that the *C. elegans* immunity via its p38 MAPK pathway also has little effect in the composition of microbiotas with higher richness.

To test how predictable this eight-species microbiota is, we made predictions of it based on monocultures and pairwise outcomes (Figure 2-6C-D). As before, we first calculated a null expectation assuming that each species is able to reach the carrying capacity that was measured in monoculture colonization. This prediction based solely on monocultures reached a mean error of 70±5% in immunocompromised AU37, and 42±13% in the immunocompetent SS104. These high errors show once again that information on monoculture colonization is insufficient to predict the assembly of microbiotas, which highlights the importance in interspecies interactions. We then utilized the two-species microbiotas to predict the eight-species microbiota in the two *C. elegans* strains (Figure 2-6C-D). Only the bacterial strains that were not competitively excluded in co-cultures should survive in the larger microbiota (assembly rules), and then we predicted the abundance of each survivor as its mean fractional abundance in co-culture experiments against the other survivors (arithmetic mean prediction). This prediction based on pairs reached an error of 28±7% in immunocompromised AU37, and 28±11% in immunocompetent GLP4, smaller errors than the predictions based on monocultures. Collectively, these data suggest that pairwise interactions are useful estimators of larger microbiotas.

2.3 Discussion

Here we characterized the bacterial colonization of the *C. elegans* intestine by native and non-native strains. We observed across three different environments (p38-immunocompetent *C. elegans* SS104, p38-immunocompromised AU37, and in vitro liquid media) a similar ordering of all the bacterial strains based on their competitive ability, indicating that the environmental filtering by the *C. elegans* intestine and by its immune system modify only a subset of the bacterial communities tested. Overall, our results show that bacterial interspecies interactions strongly influence the composition of 2-, 3-, and 8-species microbiotas, while monoculture colonization, isolation origin of bacteria, environmental filtering by the worm intestine, and immune system of the worm play secondary roles in the assembly process.

The worm gut environment is capable of enriching for certain species from the surrounding environment [365], and bacteria can evolve higher competitive abilities after in vivo passages within *C. elegans* [358, 66] (similar results in zebrafish [387] and tomato plants [388]), but our results suggest that strains isolated from the gut of natural *C. elegans* are not adapted to form bigger populations or to be better competitors than strains isolated elsewhere. A possibility for this discrepancy is that the adaptation of the microbes is rather specific to the exact hosts from where they were isolated, as it has been suggested for the bee microbiota [6, 389]. Other possibilities are that host-microbe adaptation hasn't occurred yet with the MYb bacteria that we probe, or that the adaptation that occurred selected for a different trait instead of stronger colonizers, such as healthy microbes for *C. elegans* [363, 361]. Further experiments are warranted, but our results showing non-native *E. aerogenes* and *S. marcescens* as the fittest bacteria for the *C. elegans* intestine, which align with previous results showing Enterobacteriaceae and Pseudomonadaceae as the most abundant bacterial clades in the natural *C. elegans* microbiota [364, 41, 365, 111], suggest that Gammaproteobacteria are intrinsically good at colonizing the worm intestine.

The use of non-native bacteria in the prediction of three-species microbiotas leaves open the question of whether communities with native species can be predicted with in

vitro outcomes. Recent work has indicated that native bacteria are potentially functionally important for *C. elegans* [231] and induce specific transcriptional responses in this host [346], providing grounds on which natural selection for host association could occur. It is plausible that greater divergence between in vitro and in vivo community assembly would be seen in a co-evolved community of microbes. Future work comparing bacteria native to *C. elegans* [390] and non-native bacteria will be useful to elucidate the possible host-microbe adaptation occurring in hosts with large flexible microbiotas [7, 360].

We found that most pairwise interactions among non-native species are competitive, but some facilitative interactions appeared in the two-species microbiotas with native species. Further studies should test if native species are indeed more prone to facilitate each other instead of competing. Importantly, the ratio of positive and negative interactions depends on the nutrients in the environment [231, 391], with more nutrients allowing for more bacteria to grow in monoculture, which then leads to more competition [259]. Due to the bacterivore diet of *C. elegans*, a complex mixture of nutrients is perhaps the norm in the *C. elegans* gut, so a rich medium like the one we used is perhaps a good starting point to investigate the *C. elegans* microbiota. Interestingly, we do not observe any cases of strictly non-transitive pairwise interactions (rock-paper-scissors) questioning once more [203] the practical significance of this mechanism at stimulating coexistence and diversity in multispecies communities [383].

The low pH of the gut environment is thought to be a critical factor in host-microbe interactions, and recent work has explicitly demonstrated the importance of pH in modulating the interactions between microbes and determining the structure of synthetic and natural communities [218, 392]. Consistent with these results, we observed that reducing the pH of a liquid medium to simulate the host intestine could alter the outcomes of competition between species and in some cases substantially reduce the difference between in vitro media and in vivo gut.

The antimicrobial defenses activated via p38 MAPK in *C. elegans* SS104, which include lipases and saposin-like proteins, didn't substantially shift the microbial com-

position of two- and eight-species microbiotas in our experiments, which comes in accordance with previous findings [111]. Other signaling pathways in the worm, such as the TGFβ homologue [111] and insulin-like signaling pathways [113], may be more important in determining microbiota composition, directly and/or by the combined action of multiple pathways.

The work presented here focuses on population averages rather than the composition of bacterial communities within individual worms. Recent results from our group have demonstrated that variation between individuals can be informative—when the feeding densities are lower than the ones used here, there can be an extreme bottleneck during colonization of the worm gut that leads to marked heterogeneity across worms [32]. Similar results have recently been found in Drosophila [393], indicating that stochastic effects during colonization may be important in a wide range of host species. At the level of worm populations' averages, the competitive exclusion observed could be the result of resource competition, including competition for the limited space available within the *C. elegans* gut, or there could be more explicit forms of antagonism such as toxin production [210]. In cases of coexistence, which were the majority in our experiments, spatial partitioning within the host could play an important role, such as was recently found for monoculture colonization of the zebrafish gut [394, 395]. It will be important for future studies to determine the role of stochasticity, priority effects, and spatial dynamics during assembly of multispecies communities within *C. elegans*.

In this study, we have fed worms with defined combinations of bacterial species to elucidate the role of interspecies interactions in the assembly of host-associated microbial communities. These results add to our understanding of how interactions between pairs of bacterial species shape more complex bacterial communities. Our results show that experimental bottom-up microbial ecology is a tool for understanding the dynamics of bacterial gut communities in a simple model organism, providing insight into the forces that shape and control the structure of microbiotas.

2.4 Materials and methods

Nematode culture

Nematodes were grown, maintained, and manipulated with standard techniques [32, 396]. We utilized the *C. elegans* strains N2 (wild type), AU37 [*glp-4(bn2)* I; *sek-1(km4)* X], and SS104 [*glp-4(bn2)* I]. Worm strains were provided by the Caenorhabditis Genetic Center, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Synchronized *C. elegans* cultures were obtained using standard protocols [32, 396], which rely on the reducing agent DTT and the detergent SDS to break the *C. elegans* cuticle and harvest the unhatched eggs [397].

Bacteria

Non-native bacteria were obtained from ATCC: *Bacillus subtilis* (ATCC 23857) (*Bs*), *Enterobacter aerogenes* (ATCC 13048) (*Ea*), *Lactobacillus plantarum* (ATCC 8014) (*Lp*), *Pseudomonas aurantiaca* (*Pseudomonas chlororaphis* subsp. *aurantiaca*) (ATCC 33663) (*Pa*), *Pseudomonas citronellolis* (ATCC 13674) (*Pc*), *Pseudomonas fluorescens* (ATCC 13525) (*Pf*), *Pseudomonas putida* (ATCC 12633) (*Pp*), *Pseudomonas veronii* (ATCC 700474) (*Pv*), Serratia marcescens (ATCC 13880) (*Sm*). *Bacillus cereus* (*Bc*) was obtained from Ward's Scientific Catalog. *Escherichia coli* MC4100 (CGSC #6152) (*Ec*) was obtained from the *E. coli* Genetic Stock Center.

The microbiota strains native to *C. elegans* were isolated by growing *C. elegans* N2 for one week on individual types of rotten organic material (apples, celery, almonds, and parsnip), followed by washing and sterilizing the worms on the outside, grinding the worms, and plating the resulting bacterial suspension on agar plates (Supplementary). The species identity was analyzed by 16S Sanger sequencing (Genewiz, South Plainfield, NJ, USA).

Bacterial colonization of *C. elegans*

To construct cultures to feed *C. elegans*, bacterial strains were grown to saturation (24hrs, 30C, 2 mL LB [Difco]), and then washed in S medium [396] and resuspended in 1% v/v Axenic Medium diluted in S medium (1%AXN). Undiluted Axenic Medium was prepared by autoclaving 3g yeast extract and 3g soy peptone (Bacto) in 90 ml

water, and subsequently adding with sterile technique 1g dextrose, 200µl cholesterol (5 mg/ml in ethanol), and 10 ml of .5% w/v hemoglobin (Sigma-Aldrich) in 1 mM NaOH. The bacterial cultures were standardized to $\sim 10^8$ CFU/ml based on CFU counts.

Germ-free adult worms were re-suspended in 1%AXN to a concentration of ∼1000 worms/mL. Aliquots of 120 µL (∼100 worms) were transferred into 96-deep-well culture plates (1 mL well volume, Eppendorf). Bacterial suspensions were added to reach a concentration of $\sim 10^7$ CFU/mL per bacterial strain in all feeding experiments. Plates were covered with Breathe-Easy (Sigma-Aldrich) sealing membranes and incubated for four days with shaking at 400 RPM at 25ºC. Every day the worm samples were washed and the bacteria were replenished. Samples were washed with a liquid handler (VIAFLO 96, Integra) by adding 500µl of M9 Buffer [396] + 0.1% v/v Triton X-100 (Tx), pipetting 10 times, and removing the supernatant after worms precipitated. The worms were then transferred to new 96-deep-well plates to leave behind possible biofilms, and then washed in the same way two more times with 1%AXN. Fresh bacterial cultures were added as previously described.

Mechanical disruption of worms and quantification of bacteria

The worm samples were washed to remove most external bacteria and then incubated in 100 µL S medium + 2X heat-killed OP50 at 25ºC for one hour to allow the worms to evacuate any non-adhered bacterial cells from the intestine. Worms were then rinsed twice with M9 Buffer, cooled down 15 min at 4C to stop peristalsis, and bleached for 6 minutes at 4C with 100 µL M9 Buffer + 0.2% v/v bleach (Clorox). Worms were then rinsed three times with cold M9 Buffer + 0.1%Tx to remove the bleach.

Manual disruption with a motorized pestle and 96-deep-well plate disruption with silicon carbide grit followed previously described protocols [32]. To guarantee the background media was fully clean, the supernatant in each sample was also collected, serially diluted, and plated onto Nutrient Agar (3g yeast extract, 5g peptone, and 15g of agar [Bacto] in one liter of water). The plates were incubated at room temperature

for two days to allow distinct colony morphologies to develop, and then the colonies were counted with the aid of a stereo microscope (Leica MZ10 F).

Co-culture experiments in vitro

Utilizing the same 1%AXN medium, 96-deep-well culture plates, and 150µl volume per sample used to colonize *C. elegans*, pairs of bacterial species were mixed at a concentration of 10^5 CFU/ml each. This inoculum concentration is lower than in the worm colonization experiments to allow growth for all bacteria (Figure 5-1E). We allowed the bacterial relative abundances to equilibrate with seven growth-dilution cycles, where the bacteria are diluted 100-fold into fresh media each day. Bacterial abundances were quantified by plating onto agar and distinguishing colony morphologies. To lower the pH of the S medium + 1%AXN, NaOH 1M was added while measuring continuously the pH of the media with a microelectrode (N6000BNC, SI Analytics). The media was filtered (Millex-SV 0.2 µm, MerckMillipore) afterwards.

Data analysis

Pairwise and trio outcomes were categorized as coexistence if the rare species was present at an average abundance of more than 2%. This threshold is just above our usual limit of detection of ∼1%, which is inversely proportional to the number of colonies counted. The pair *Pf-Ea* (1.7%-98.3%) was defined as coexisting since we could reassure the presence of *Pf* with more than one biological replicate. Given that we fed the worms with one initial composition, we cannot detect the possibility of bistability, in which the final fractional abundance depends upon the starting ratio of the two species.

The mean relative yield of a species in a co-culture, $RY_{i|j}$, was calculated by bootstrapping (sampling with replacement) simultaneously over the pairwise and monoculture CFU/worm data to obtain vectors $N_{i|j}$ and N_i, respectively, and then calculating $log(RY_{i|j}) = \langle log(\frac{N_{i|j}+1}{\langle Ni+1\rangle})\rangle$. We utilized logarithmic scales to have comparable calculations regardless of numerator/denominator choice, and we added 1 to avoid ±infinite. The null expectation based on monocultures is obtained by averaging the fractional abundances of all possible combinations of the species' monoculture information. The

standard error of the mean (s.e.m.) is calculated with n as the least number of monoculture replicates. These mean and s.e.m. can also be obtained by bootstrapping over the monoculture data. The hierarchy score of the fractional abundances' matrix was calculated as previously described [203].

For phylogeny reconstruction, sequences of the full 16S rRNA gene were obtained from NCBI. Sulfolobus solfataricus, a thermophilic archaea, was used as an outgroup species to root the tree. Clustal X with default parameters was used to align the sequences [398]. PhyML-SMS with default parameters was used to select GTR+G+I as the best model and to infer the tree [399]. The phylogenetic distances were calculated directly from the phylogenetic tree.

Chapter 3

Positive interactions are common among culturable bacteria

In collaboration with Jared Kehe, Anthony Kulesa, Jonathan Friedman, and Paul C. Blainey[1]

Interspecies interactions shape the structure and function of microbial communities [186, 210]. In particular, positive, growth-promoting interactions can significantly affect the diversity and productivity of natural and engineered communities [401, 262]. However, the prevalence of positive interactions and the conditions in which they occur are not well understood. To address this knowledge gap, we used kChip [259], an ultra-high throughput coculture platform, to measure 180,408 interactions among 20 soil bacteria across 40 carbon environments. We find that positive interactions, often described to be rare, occur commonly, primarily as parasitisms between strains that differ in their carbon consumption profiles. Notably, non-growing strains are almost always promoted by strongly growing strains (85%), suggesting a simple positive interaction-mediated approach for cultivation, microbiome engineering, and microbial consortium design.

[1]Kehe J, Ortiz A, Kulesa A, Gore J, Blainey P, Friedman J. Positive interactions are common among culturable bacteria. bioRxiv. 2020 Jan 1.[400]

3.1 Background

3.1.1 kChip screening identifies compositions that robustly promote growth of *Herbaspirillum frisingense*

For our screen, we measured the yield of a GFP-expressing strain of *Herbaspirillum frisingense* (Hf-GFP), an Alphaproteobacteria that can colonize the interior of cereal plants' roots and is regarded as a model plant symbiont [402, 403]. We collected a diverse set of soil bacterial isolates (Figure 5-7) and measured the yield of Hf-GFP across isolate combinations. These combinations were constructed across a media library that included carbohydrate oligomers (sucrose, lactose, and raffinose) and their monomeric constituents (glucose, galactose, and fructose). Hf-GFP grew in monoculture to various extents on each of these carbon sources, with growth on sucrose being indistinguishable from background (Figure 3-1).

Most of the isolate compositions had a significant effect on Hf-GFP yield, showing both suppressive (decrease in yield) and facilitative (increase in yield) effects. We measured Hf-GFP yield at 24, 48, and 72 hrs and rank-ordered yield at 72 hr, the time point when yields were highest [259]. On carbon sources where Hf-GFP monocultures achieved high yield by 72 hours (glucose, galactose), the addition of other isolates almost always attenuated its growth. By contrast, facilitative compositions were common on carbon sources where Hf-GFP growth was low (fructose, raffinose, lactose) and ubiquitous when undetectably low (sucrose).

For the majority of compositions, however, the facilitative effect did not persist across different carbon sources or community contexts. For example, the composition [Bacillus sp. I + Rahnella sp.] ([BaT + Ra]) (measured at k = 2 on each kChip) facilitated HF-GFP growth on fructose, sucrose, and raffinose, but suppressed its growth on the other carbon sources (Figure 5-8). Similarly, the facilitation imparted by the composition [Enterobacter mori + Dyella sp.] ([En + Dy]) (measured at k = 2) in a medium containing galactose was not robust to community context [En + Dy + $\geq$1 additional unique isolates] (an s = 2 composition among all k $\geq$ 3 communities

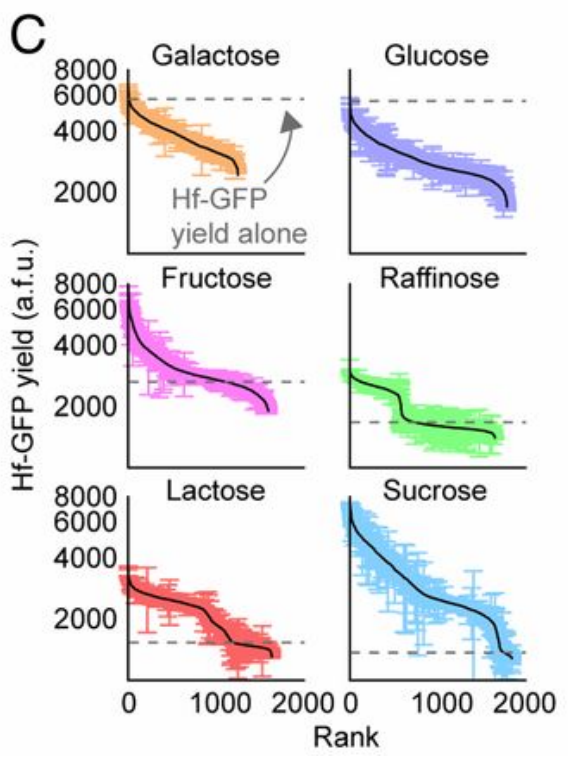

Figure 3-1: **Ranked Hf-GFP yield at 72 hr for all communities constructed.** Median calculated when community replicated >1 time; error bar = standard error of the mean; dotted line represents the yield of Hf-GFP in monoculture.

on the same kChip) (Figure 5-8).

We sought to determine the facilitative compositions that were most robust to both carbon source and community context. We scored each composition in two ways: First, we computed the median Hf-GFP yield at 72 hr in co-culture with the composition across all carbon sources ("Hf-GFP yield"). As a second score, we computed the 10th percentile of Hf-GFP yield in co-culture with all communities containing the given composition to detect whether the composition's effect on Hf-GFP was appreciably diminished by additional isolates across all carbon sources ("Hf-GFP robustness"). Based on noise estimates of Hf-GFP growth data, we restricted the analysis to instances where a given composition was represented $\geq$5 times on average on a kChip (or $\geq$30 times in total), which occurred for about half of k = 3 compositions, a value consistent with a probabilistic model of combinatorial space sampling [259].

We uncovered two isolate compositions that were strongly facilitative and robust to both carbon source and community context (Figure 5-8). While most facilitative compositions showed robustness to community context for a given carbon source, few showed robustness to both carbon source and community context. Interestingly, we identified a particular isolate, Burkholderia sp. I (BuC), consistently present among compositions where Hf-GFP yield was high across carbon source and community context. We further identified the isolate composition [Bacillus sp. II + Rahnella sp.] ([BaL + Ra]) that also enabled strong Hf-GFP yield across carbon source and community context. We validated these two compositions' facilitative effects on Hf-GFP with the different carbon sources in 96-well plate bulk co-culture experiments [259].

3.1.2 Facilitation increases with community richness and is driven by a small number of strains

The datasets that kChip screening generates can be used to detect the ecological trends that underpin particularly facilitative or robust compositions like BuC or [BaL + Ra]. Broadly, we found that Hf-GFP yield either increased or declined with community richness depending on its baseline growth on each carbon source in monoculture. In raffinose, lactose, and sucrose, the three carbon sources where it grew most poorly, Hf-GFP yield increased and then plateaued with community richness (Figure 3-2A). We observed a similar but weaker trend in fructose, where Hf-GFP grew to a limited extent alone, and an isolate-agnostic suppressive effect in glucose and galactose where Hf-GFP grew well in monoculture. Without restricting the analysis to communities consisting of all unique isolates, we observe the same trend for k = 1–7 inputs and a roughly equivalent yield of Hf-GFP for 7 and 19 inputs [259].

In co-culture with a single isolate (k = 1 microwells), we identified "primary facilitator" strains that facilitated Hf-GFP growth in a given carbon source (Figure 3-2D). We further found that the inclusion of $\geq$1 primary facilitator strains was necessary and almost always sufficient to facilitate Hf-GFP growth regardless of the number of

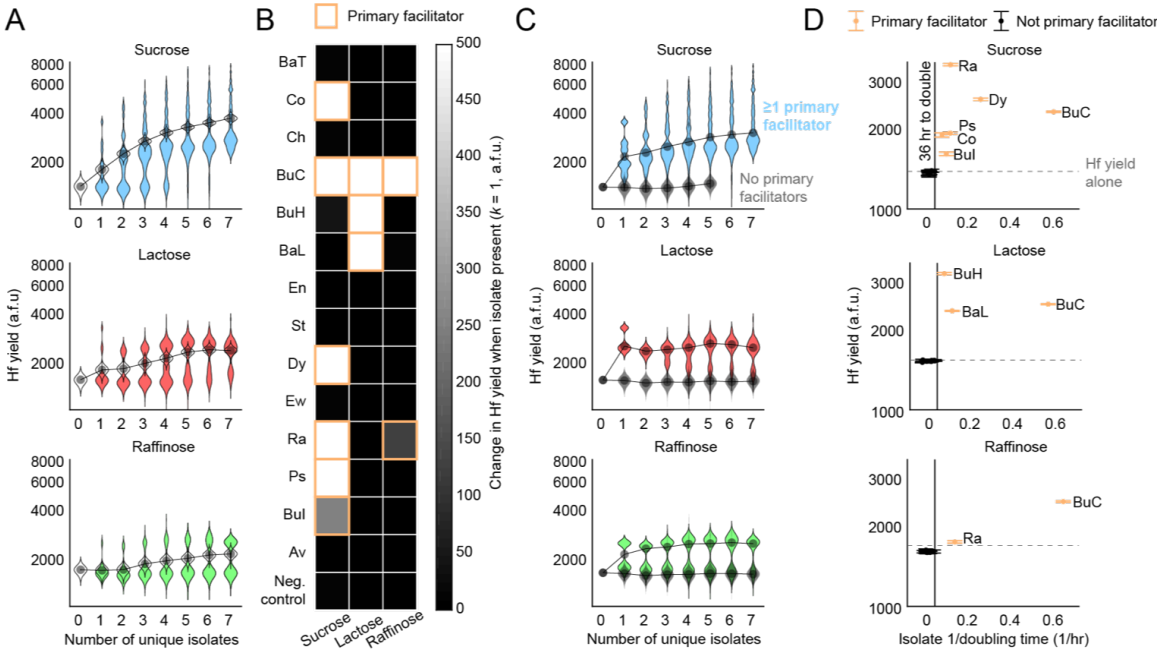

Figure 3-2: **Facilitation increases with community richness and is driven by a few species.** (A) In a medium containing sucrose, lactose, or raffinose, Hf-GFP yield increased with community richness. Colored distributions = median Hf-GFP yields for all unique compositions at given k (i.e. all droplets in a given combination contain different strains). Gray data point = median of distribution. Outlined distributions = medians of 100 bootstrap-resampled datasets at each k, whereby the Hf-GFP yield dataset for each k was resampled with replacement (with resampling sample size equal to the actual sampling size), and a median of the resampled data was calculated each time. (B) Primary facilitators (outlined in orange) were classified as the isolates increasing Hf-GFP yield by >100 GFP counts (a.f.u.). (C) The presence of ≥ 11 primary facilitator was necessary and typically sufficient to enable Hf growth and drive facilitative effect when additional isolates were present. Colored distributions = Hf-GFP growth in communities possessing ≥ 1 primary facilitators. Gray distributions = Hf-GFP yield in communities with no primary facilitators. (D) A resazurin assay was conducted on a separate k = 2 Chip in parallel with the screen to measure growth rate of each isolate. The subset of isolates that grew on a given carbon source (defined at ≥ 11 doubling of resorufin fluorescence by 36 hr) corresponded with the subsets of isolates identified as primary facilitators.

other strains present (calculated from k > 1 microwells) (Figure 3-2C). In conjunction with the screen, we also assayed growth rates of the isolates on the different carbon sources via the resazurin assay on a k = 2 Chip (Figure 5-7). For the raffinose, lactose, and sucrose conditions, we found that the subset of isolates that could grow ($\geq$1 doubling by 36 hr) matched the subset of primary facilitators (Figure 3-2D). We concluded that, for these conditions, isolate growth was sufficient for Hf-GFP facilitation. To investigate why facilitation increased with community richness beyond the presence of one primary facilitator, we first considered the k = 2 level. We found many instances where Hf-GFP yield in the presence of two isolates was greater than its yield with either isolate individually, particularly when the carbon source was sucrose [259]. With a single primary facilitator present, we observed that the largest Hf-GFP yield increases were imparted by the addition of a second primary facilitator (Figure 3-2C). We also observe that modest increases may be imparted by the addition of a non-[primary facilitator] when a primary facilitator is present.

These data highlight general design principles useful in constructing facilitative consortia. Based on carbon source utilization as a criterion for primary facilitation (Figure 3-2D), we might expect that certain "core" compositions of isolates facilitate Hf-GFP across all carbon sources if at least one isolate within the composition is able to grow on each carbon source. Further, based on the increase in Hf-GFP yield we see with community richness (Figure 3-2A), we might expect improvements to the facilitative effect size or its robustness with the incorporation of specific isolates to the composition. Indeed, our two top-scoring compositions, BuC and [BaL + Ra], abide by these principles [259].

3.2 Introduction

Microbial communities are composed of multiple species that interact with one another in a variety of ways as part of the "struggle for existence" [151]. Interactions between species can be negative, where a species inhibits another species' growth through nutrient exploitation and chemical warfare [210]; or positive, where a species

promotes another species' growth by increasing nutrient availability and creating new niches [401]. The bidirectional interaction between two species is determined by the two one-way interactions between the species. For example, a mutualism occurs when two species positively affect each other. The overall distribution of positive and negative interactions within a microbial community profoundly impacts the community's structure, stability and productivity [186, 262, 167]. These properties, in turn, shape a community's ability to perform vital functions for the environment [404, 405, 406, 407] and for host organisms [408, 409, 410, 411]. Despite the importance of the distribution of microbial interactions, the relative prevalence of positive and negative interactions in nature remains largely unknown.

Positive interactions are generally thought to be rare. Experimental evidence from coculture studies points to a dominance of negative interactions [210]. For example, evidence of positive interactions was found in $<10\%$ of pairs of bacteria isolated from tree holes [232]. However, these results are subject to strong experimental biases, such as the use of a single environment—even though microbial interactions can differ dramatically across environmental conditions [254, 412, 413] —and the use of strains that each grow individually in the environmental conditions being tested. Metabolic modeling, which can simulate millions of interactions across myriad environments, as well as limited experimental evidence, suggest that positive interactions emerge via environment-dependent secretions and can facilitate otherwise non-growing species [254, 414, 415, 28]. Additionally, evolutionary theories like the Black Queen Hypothesis argue that such secretion-mediated positive interactions are selected for [416]. Taken together, these findings suggest that positive interactions among microbes may be common and play a significant role in shaping microbial communities, but such theories have not yet been thoroughly tested experimentally.

Quantifying the prevalence of positive interactions and determining the conditions in which they occur could significantly improve our ability to predict and control the ecology of microbial communities [417, 73]. Positive interactions are predicted to enhance a community's diversity and productivity but decrease its stability [186, 418, 419]. Therefore, a better understanding of these interactions would

enhance our ability to manipulate and manage communities, with widespread applicability in environmental conservation [420], crop health [421], and human health [422]. Nevertheless, the data required for quantifying the distribution of interactions across environments is still lacking due to methodological limitations that frustrate comprehensive sampling of interactions under many conditions [271]. Inferring interactions from metagenomic sequencing remains an outstanding challenge [423, 424], and directly measuring interactions at scale is difficult to perform using existing experimental paradigms.

To gain a broad understanding of how species interact across a wide range of environments, we used a combinatorial screening platform called kChip [259, 275, 425] to measure >150,000 bidirectional bacterial cocultures among 20 different soil bacterial strains across 40 environments with differing carbon source identity or concentration. The kChip generates cocultures at an ultra-high throughput scale by rapidly and randomly combining droplets containing microbial cultures and/or medium components within microwells (Figure 3-3A-D). Here, we paired unlabeled (wild-type) and GFP-labeled versions of 20 Gammaproteobacteria isolated from soil (6 Enterobacterales and 14 Pseudomonadales, Figure 5-9 and 5-10). We selected these bacterial strains from our larger pool of fluorescently labeled soil isolates maximizing for phylogenetic diversity (Table S1, Materials and Methods). We cocultured each strain pair in each of 33 single carbon sources (0.5% w/v), a mix of these, five of the 33 at reduced concentration (0.05% w/v), and a no-carbon control. Carbon sources were drawn from several biochemical classes including carbohydrates, amino acids, and carboxylates (Table S2). We measured the effect of each unlabeled strain on the growth of each labeled strain in each carbon source, giving a total of 17,600 combinations (20 labeled strains, [20 + 2 control] unlabeled strains, and [39 + 1 control] carbon sources). Each combination appeared >10 times on average and only 3% appeared <3 times and were excluded from further analysis [400]. Measured one-way interactions were used to classify each bidirectional interaction qualitatively (Figure 3-3E) and quantitatively (Figure 3-3F).

Our data provide direct experimental evidence that positive interactions are indeed

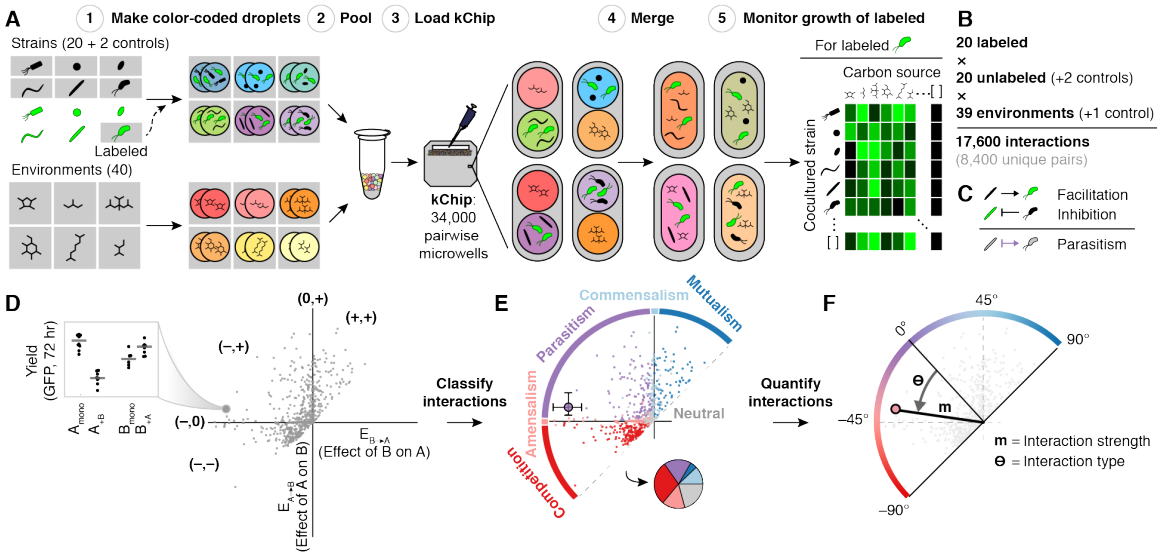

Figure 3-3: **High-throughput interaction assay and analysis.** (A) Steps to assay the effect of multiple unlabeled species on a single label species across carbon sources on each kChip. Color-coded droplets, each containing either a labeled + unlabeled coculture or a single carbon source, were generated (Step 1), pooled together (Step 2) and loaded onto a kChip (Step 3). Each kChip contained an array of microwells that randomly paired coculture droplets with carbon source droplets. After imaging the color codes to identify the inputs per microwell, droplet pairs were merged via exposure to an electric field (Step 4), and the growth of the labeled strain was measured at 0, 24, and 72 hrs (Step 5). (B) Overall size of the kChip screen. (C) Using data across kChips, bidirectional interactions were deduced by combining data where each strain within a given pair was the labeled strain. D-F. Framework for kChip data analysis. Each bidirectional interaction was described qualitatively (interaction classification) and quantitatively (interaction strength, m and interaction type, Θ).

common, and occur primarily as parasitisms, where one species' growth is improved at the expense of the other's. More broadly, we found that interactions strongly depend on the environment via differences in the carbon consumption preferences of the interacting strains. Notably, we discovered that strongly growing partners consistently enabled the growth of strains that were unable to grow in monoculture (85%), suggesting a simple strategy for cultivation, microbiome engineering, and design of microbial consortia.

3.3 Results

3.3.1 Positive interactions occur frequently between twenty Gammaproteobacteria isolates

The effect of each unlabeled strain on each fluorescently labeled strain was classified as positive (yield increase compared to monoculture), negative (yield decrease compared to monoculture), or 0 if there was no evidence for an effect (Figure 3-3C, D, S5) (see Materials and Methods). The size of an effect was calculated as log2 of the ratio of growth (fluorescence) in coculture to monoculture. Bidirectional interactions were classified as mutualism (+,+), commensalism (0,+), parasitism (–,+), amensalism (–,0), competition (–,–), or neutralism (0,0) (Figure 3-3, 5-12). We focused our analysis on the 72 hr time point by which cultures reached saturation. Thus the measured interactions reflect changes in strains' overall growth yield, rather than growth rates.

Positive interactions were common overall (Figure 3-4A). Excluding cases when neither strain within a pair grew detectably in monoculture, >40% of cocultures contained at least one (≥1) positive one-way interaction, with over half of these occurring within a parasitism (22% parasitisms, 14% commensalisms, and 5% mutualisms). There were also many cocultures that contained only negative interactions (35% competitions and 18% amensalisms). Relatively few interactions were neutralisms (6%), though neither strain grew in monoculture for 21% of pairs.

The extent to which interacting pairs influenced each other's growth also varied quantitatively across strain and carbon environments. For example, in a parasitism, the facilitated strain may have increased in yield more than the inhibited strain decreased or vice versa. To capture these quantitative differences, we measured the magnitude m (strength) and angle Θ (type) of each bidirectional interaction in polar coordinates (Figure 3-6-1F). Θ represented the relative effect of cocultured strains on each other (with Θ = -90° indicating that both strains inhibited each other equally, Θ = 0° indicating a balanced parasitism with equal and opposite effects, and Θ = 90°

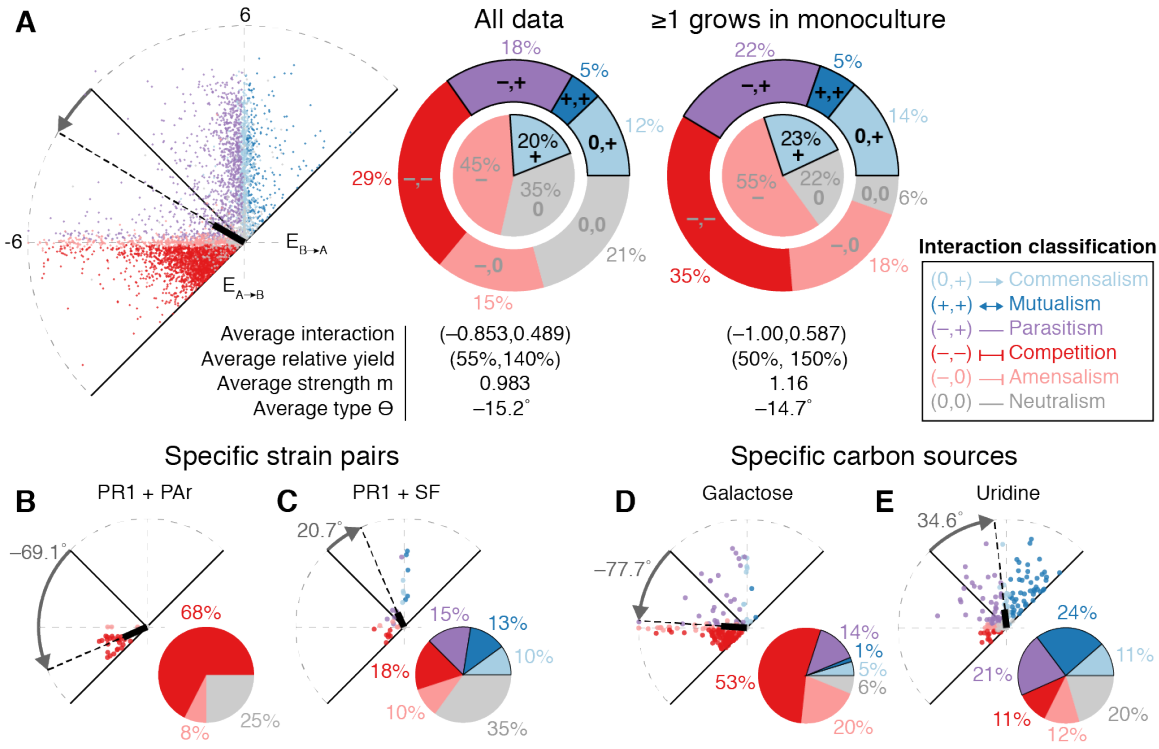

Figure 3-4: **Positive interactions occurred commonly and depended on strain pair and carbon source.** (A) (Left) All bidirectional interactions on 33 distinct carbon sources. (Middle) Interaction classification of all data. The inner circle represents one-way interactions classifications; the outer ring represents two-way interaction classifications. (Right) Interaction classification excluding cases in which both strains comprising a coculture showed no detectable growth as monocultures on a given carbon source. B-C. Interactions of two example cocultures on all carbon sources. D-E. Interactions of all cocultures on two example carbon sources. Colors indicate interaction classification. All data at 72 hr.

indicating both strains facilitated each other equally; no Θ is assigned to neutralisms).

The average interaction was a parasitism (Θ = -14.7° and m = 1.16) in which one strain was facilitated at expense of the other strain whose growth was inhibited (Figure 3-4A). On average, the inhibited strain was hurt ($\sim$ 50% growth compared to monoculture) more than the facilitated strain was helped (150% growth compared to monoculture). The results at 24 hr were similar and included even more positive interactions, with 54% of cocultures containing $\geq$1 one-way positive interaction (20% parasitisms, 25% commensalisms, and 9% mutualisms), although the average interaction, also a parasitism, favored the facilitated strain ($\sim$30% growth for one strain and 170% for the other as compared to monocultures) (Figure 5-14).

We have validated the kChip interaction assay by measuring a subset of the interactions in microtiter plates ([7 strains + 1 control] and 4 carbon sources, generating 256 combinations, or <2% of the total combination constructed on kChip). Monoculture fold-growth based on GFP and OD600 measurements in microtiter plates were strongly correlated (Figure 3-5A). In addition, there was strong agreement between monoculture growth based on GFP measurements in microtiter plates and kChip (Figure 3-5B), as well as between one-way effects and interaction types calculated from these measurements (Figure 3-5C-D). Finally, as expected, intra-strain interactions measured on kChip (i.e. those between labeled strains and their unlabeled counterparts) were typically competitions (90% when excluding non-growing strains) that reduced the yield of the labeled strain by 50% on average [400].

3.3.2 The occurrence of positive interactions differs among strain pairs and among carbon sources

The prevalence of positive interactions differed significantly among strain pairs: Positive interactions never occurred for some pairs, while for others they occurred in a large fraction of the tested environments (Figure 3-4B-C). Many pairs were capable of producing all six interaction types across the carbon source library (Figure 3-4C), indicating that it can be challenging to predict the interaction of a particular pair in a given environment based on the pair's interaction in a different environment.

Positive interactions were also more common on certain carbon sources than others. In galactose, for example, a majority of bidirectional interactions were competitions (53%), with only 20% containing ≥1 positive interaction (Figure 3-4D). In uridine, however, competitions were relatively rare (11%), with 56% of bidirectional interactions containing ≥1 positive interaction (Figure 3-4E). Several carbon sources were capable of producing all interaction classes across our strain set [400].

The strong dependence of the interaction type on strain identity and carbon source indicated a need to examine our dataset broadly in order to detect patterns governing the occurrence of positive interactions. We therefore tested whether characteristics

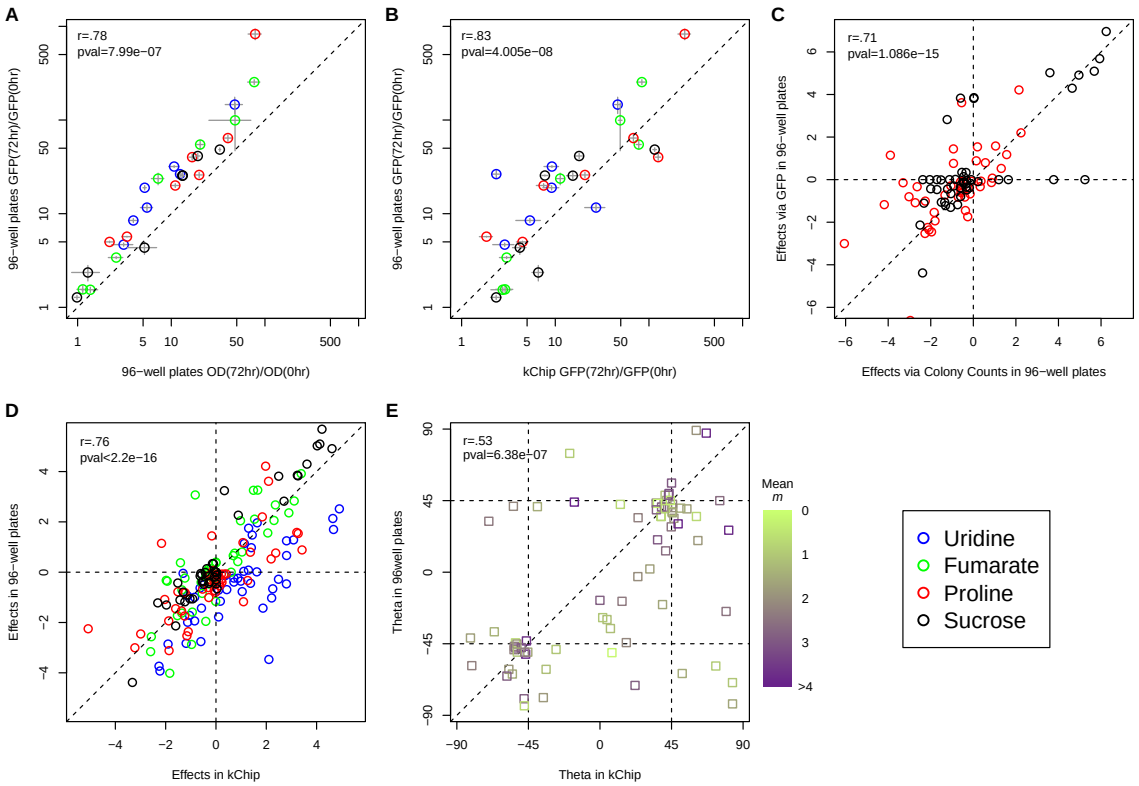

Figure 3-5: **Growth and interaction measurements in 96-well plates validate kChip measurements.** (A) Fold-growth of seven bacterial species in monoculture (RP1, KA, PAg2, LA, PP, PR2, and PK) calculated with optical density (OD600) and fluorescence intensity (GFP) in 96-well plates. Points are the mean of 6 replicates, and error bars are the standard error of the mean (s.e.m.). The colors denote the four different carbon sources (Uridine, Fumarate, Proline, and Sucrose). (B) Monoculture fold-growth in 96-well plates and on kChip. kChip GFP yields were normalized by subtracting the 20th percentile yield at 0 hrs. (C) Unidirectional pairwise effects among seven bacterial species in 96-well plates. The effects via colony counts and GFP measurements are calculated from two and three coculture replicates, respectively. (D) Unidirectional effects calculated from GFP data from kChip and 96-well plates. (E) Interaction type (Θ) calculated from kChip and 96-well plates data. The color of the points is the mean magnitude effect from both platforms. All Pearson correlation tests were calculated with a two-sided alternative hypothesis.

of the environmental conditions or the strain pairs could explain the interactions we observed.

3.3.3 Positive interactions increase with strain dissimilarity, yet they hardly correlate with carbon source biochemical class

Broad environmental characteristics, including carbon source biochemical class and number of carbon atoms, did not appear to dictate the occurrence of positive interactions. We found that the variation in interaction classification between carbon sources within a class was not significantly different than variation between classes (Figure 3-6A) (pairwise PERMANOVA, $p > 0.05$). We hierarchically clustered all interactions for all strain pairs and carbon sources and found that carbon sources did not group by biochemical class (though TCA cycle carboxylate ions grouped more adjacently) [400]. We further found that the occurrence of positive interactions was only weakly correlated to the number of carbon atoms each carbon source possessed (Figure 3-6B) (Spearman correlation = 0.32, $p = 0.07$). For example, serine, a compound with only 3 carbon atoms, enabled more positive interactions than all other carbon sources except uridine. We found that positive interactions were more common among taxonomically dissimilar strains. We compared interactions within and between the two taxonomic orders represented in our set, Enterobacterales (En) and Pseudomonadales (Ps), since pairwise phylogenetic distances, calculated from full-length 16S rRNA gene sequences, distributed bimodally (Figure 5-11). More positive interactions occurred among inter-order pairs (En+Ps) than intra-order pairs (En+En and Ps+Ps) (Figure 3-6C): 26% of En+En pairs and 20% of Ps+Ps pairs contained $\geq$1 positive interaction, compared to 47% for En+Ps pairs. The average interaction type was negative for intra-order cocultures (-32° for En+En and -45° for Ps+Ps) but positive for inter-order cocultures (4.5° for En+Ps) (Figure 3-6D) without significant differences in interaction strength.

The fraction of positive interactions, especially parasitisms, increased with the metabolic distance between the interacting species (Figure 3-6E). Metabolic differences between strains were calculated as the Euclidean distances between strains' carbon source utilization profiles, given by each strain's ability to grow on each car-

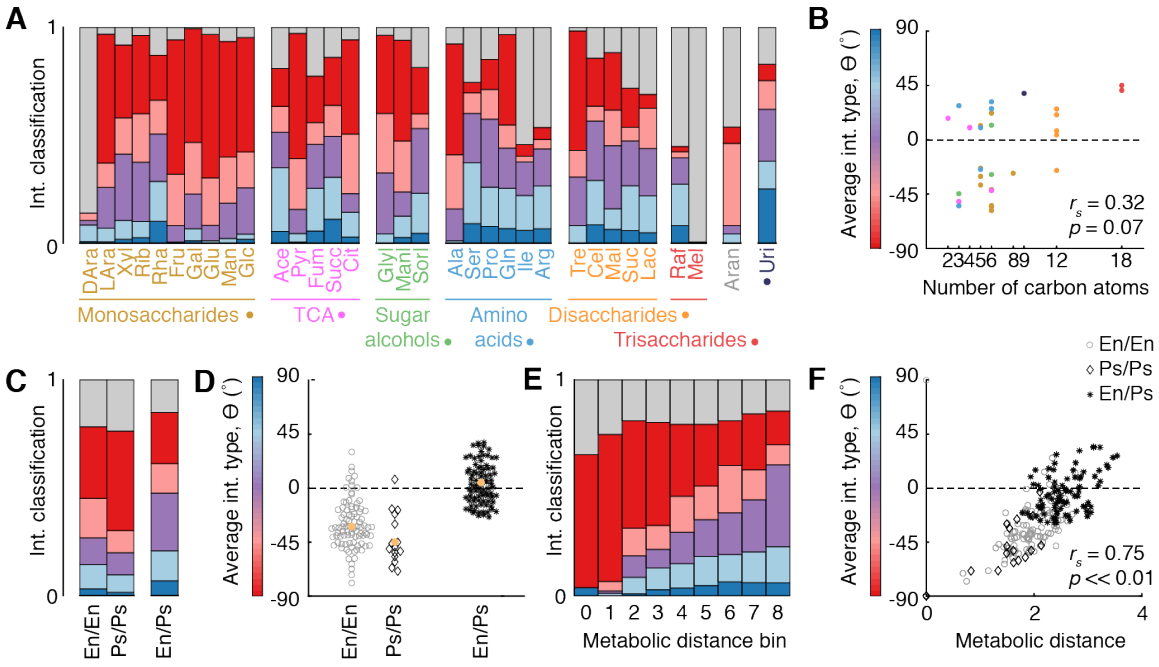

Figure 3-6: **Positive interactions depended strongly on strain properties.** (A) Interaction classification by carbon source organized into biochemical categories. Bar colors indicate interaction classification (legend provided in Figure 3-4A). (B) Average interaction type by number of carbon atoms. Dot color indicates biochemical class. (C) Interaction classification by phylogenetic relatedness of the taxonomic pairs. En = Enterobacterales. Ps = Pseudomonadales. (D) Interaction type by phylogenetic relatedness of the taxonomic pairs, averaged across all carbon sources. (E) Interaction classification by pairwise Euclidean metabolic distance (binned). Bin 0 represents with-self interactions. Bins 1-8 each contain roughly equal numbers of bidirectional interactions. (F) Interaction type by metabolic distance. All data at 72 hr.

bon source in monoculture (Figure 5-10). The distribution of metabolic distances between all strains was a continuous bell-shaped distribution (Figure 5-11), which captured more finely graded functional differences among strains than bimodal phylogenetic distances. Intra-strain mutualisms (i.e. those between labeled strains and their unlabeled counterparts) occurred in a small fraction of strains (3.7%), which we hypothesize was caused by higher starting density (Allee effect [36]). Interactions between metabolically similar strains were also typically competitions (68% for the most metabolically similar but non-identical strains) (Figure 3-6E). As metabolic distance increased, the fraction of pairs that exhibited $\geq$1 positive interaction increased monotonically from 0.2% to 61%, with a growing fraction of these positive interac-

tions occurring within parasitisms. This result indicated a strong and directional dependence of positive interaction frequency on the functional dissimilarity of strain pairs.

As metabolic distance increased, the average interaction type Θ also became increasingly positive (Spearman correlation = 0.75, $p \ll 0.01$) (Figure 3-6F). This trend reflects both shifts in the frequency of interaction classes and quantitative changes in the relative strength of interactions' components. Notably, among parasitisms, the average Θ within each metabolic distance bin increased, indicating an increasingly large facilitative effect relative to the reciprocal inhibitory effect. Since the prevalence of interaction types was correlated with strains' metabolic capabilities, we next tested whether interactions in a given environment were dependent on the interacting strains' abilities to grow in monoculture in that environment.

3.3.4 Monoculture yields shape occurrence of positive interactions in co-culture

Interaction networks for each carbon source suggested that monoculture yield indeed related to interactions (examples in Figure 3-7A and all networks in [400]). Regardless of the carbon source or strain pair, positive interactions (often in the form of parasitisms) commonly occurred between strong and weak growers and sometimes occurred between weak growers; competition commonly occurred between strong growers (Figure 3-7B). These patterns were found to be robust in a systematic analysis of interactions as a function of monoculture yield (Figure 3-7B-C). The frequencies of mutualisms, commensalisms, and parasitisms all increased with the difference between the monoculture yields of interacting strains (individually normalized to the maximum observed growth of each strain across all carbon sources; off-diagonal of Figure 3-7C). By contrast, for interactions of strain pairs able to grow on a carbon source equally well (diagonal of Figure 3-7C), competitions were consistently prevalent and increased in frequency with monoculture yield. Among interactions between the strongest growers, 77% were competitions where both species were in-

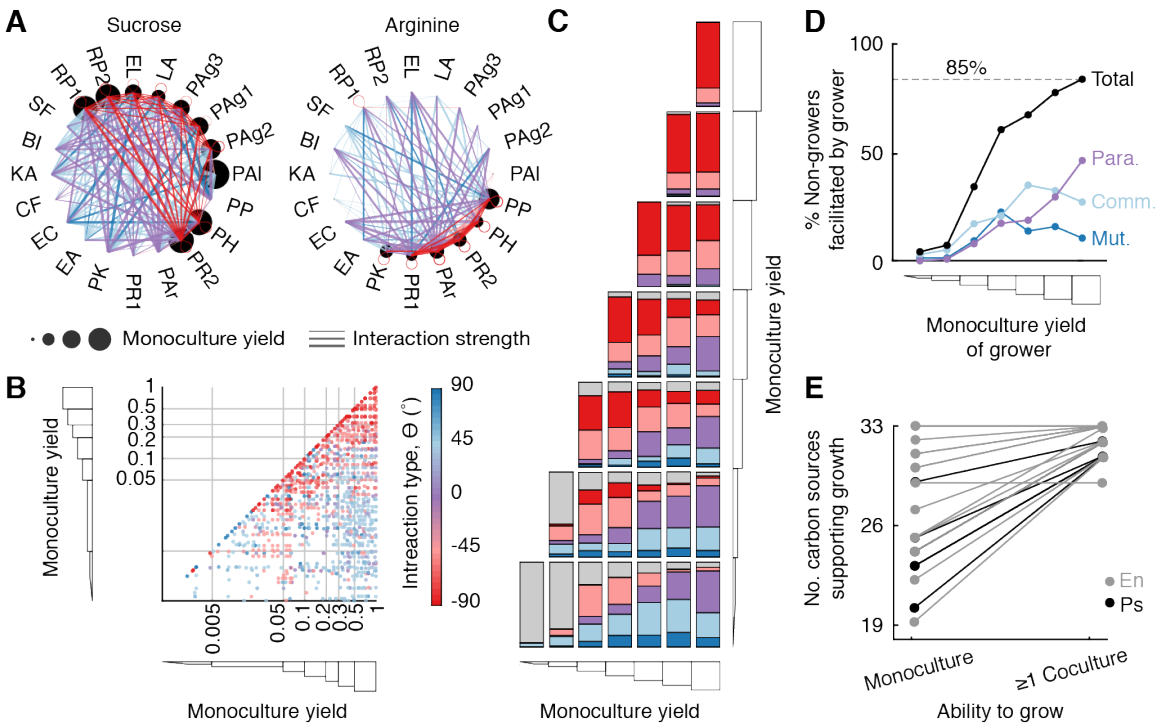

Figure 3-7: **Carbon source utilization capabilities shape interactions.** (A) Example interaction networks for two carbon sources, each with different subsets of strains that could utilize it. The size of the circular nodes represent each strain's growth in monoculture (normalized to each strain's maximum monoculture yield across all carbon sources). Edge color represents interaction classification. Edge thickness represents interaction strength (m). (B) Interaction types (color) by the monocultures yields of both strains in each coculture. Gray lines indicate monoculture growth bins. (C) Interaction classification for each pair of monoculture growth bins. (D) Fractions of non-growers that are obligately facilitated by strains with different monoculture yields. Line colors represent the interaction classifications in which the facilitation occurs (Mut. = mutualism; Comm. = commensalism; Para. = parasitism; Total = total fraction of non-growers facilitated). (E) The total number of carbon sources on which each strain can grow in monoculture and in at least one coculture. Each line represents one strain. All data at 72 hr.

hibited compared to their monoculture growth. Moreover, in 99% of pairs of the strongest growers at least one species was inhibited (77% competitions, 18% amensalisms, and 4% parasitisms). Interaction at 24 hr followed the same trend, with the overall increased prevalence of positive interactions (Figure 5-14) likely reflecting the lack of high monoculture yield values. Overall, interactions appeared to depend heavily on the two interacting strains' individual abilities to grow in a specific environment, and this dependence explained the interaction variability exhibited by certain

strain pairs and certain carbon sources.

These results were also consistent with the fact that there were far fewer positive interactions on low-concentration carbon sources (11% with $\geq$1 positive interaction, as opposed to 35% for their high-concentration counterparts) [400]. The low-concentration carbon sources supported low-to-mid-range monoculture yields that were similar among the strains (Figure 5-10). There were few instances of strong growers paired with weak growers, the regime where positive interactions otherwise emerged [400]. Consequently, the interaction classifications on low-concentration carbon sources were nearly as different to their high-concentration counterparts as the high-concentration carbon sources were to each other [400]. Unlike these low-concentration carbon sources, a mix of 33 carbon sources produced consistently high monoculture yields (most similar to glucose) [400]. As a result, interactions occurred only between strong growers and were consequently highly negative (85% competitions, with 99% containing $\geq$1 inhibition) [400]. Indeed, for each strain pair, the interaction type Θ on the mixed carbon source condition was, with a single exception (PR1+PAr), always lower than its average interaction across all carbon sources [400]. Altogether, differences in monoculture yields provided a common statistical explanation for differences in the interaction distributions across timepoints, carbon source types, and carbon source concentrations.

3.3.5 Non-growing strains are typically facilitated by strongly growing strains

To quantify the prevalence of obligate facilitation, where a non-growing strain required a facilitating partner to grow on a given carbon source, we examined interactions between strains that could grow strongly on a carbon source and those that could not (bottom row of Figure 3-7C). As the yield of the growing strain increased, the frequency of facilitation of the non-grower increased markedly. When paired with the strongest growers, non-growers were facilitated 85% of the time. Moreover, the relative fraction of these positive interactions that occurred as parasitisms also increased

(Figure 3-7D), making up 55% of obligate facilitations between the strongest growers and non-growers. Among these parasitisms, the average interaction type also became increasingly positive [400]: The magnitude of the positive interaction increased monotonically without the inhibited strain experiencing the same degree of yield loss [400].

At least one coculture partner could usually be found to support the growth of otherwise non-growing strains on almost any carbon source. We observed a total of 125 cases where a given carbon source did not support detectable growth of a given strain in monoculture (Figure 3-7E). By contrast, we found only 21 instances of a given carbon source never supporting the growth of a given strain despite its coculture with any partner strain. This common obligate facilitation provides a potential explanation for how biodiversity can be supported when few carbon sources are available or when only a subset of strains can utilize available carbon sources.

3.4 Discussion

In this study, we performed comprehensive pairwise coculturing of 20 culturable soil microbes across 40 different carbon source environments. By measuring interactions across many environments, our study produced many instances in which at least one of the two cocultured strains grew poorly or not at all in monoculture, a regime in which we found positive interactions were far more likely to occur than previously measured [210, 232]. The study also produced instances where both strains grew well and antagonism was common, a result consistent with previous large-scale studies of bacterial interactions [160, 161].

Our study unearthed the wealth of positive interactions that occur in our system. While mutualisms were relatively rare (5%), commensalisms (12%) and parasitisms (18%) were common, and accounted for the majority of cases where total coculture yield was greater than the sum of monocultures (24%)—a criterion previously used to classify cooperative interactions [232]. The prevalence of these positive interactions corroborates predictions from large-scale metabolic models [415, 28, 426]. Our results

are also consistent with the predictions of theories like the Black Queen Hypothesis, which asserts that interspecies cross-feeding of "leaky" public goods are evolutionarily selected for [416, 239]. Finally, our results generalize smaller scale demonstrations that cocultured strains41 and spent media [427] can induce growth of fastidious bacteria. Altogether, positive interactions increasingly appear to play a dominant role in driving community properties, such as resistance to invasion and productivity [401], and in supporting microbial biodiversity [428].

Interactions in our system varied significantly across environments and timepoints (Figure 3-4, 5-14), suggesting that properties of natural communities can display considerable spatial and temporal variability. While interactions did not appear to significantly depend on properties intrinsic to the environment itself, they nonetheless strongly depended on the environment via the ability of each strain to individually grow in it: Negative interactions were frequent between strong growers, while positive interactions occurred commonly between strong and weak growers across all timepoints and environments. Given the widespread differences in growth that occur among bacteria, these results suggest that positive interactions may occur commonly in nature.

A variety of mechanisms could explain the prevalence of positive interactions in our system. First, facilitated strains might have grown on components of accumulating dead cells, though this is unlikely given the timescale of the coculture experiment [429]. Second, the facilitator might have secreted carbon source-degrading enzymes that increased overall carbon availability. This mechanism is consistent with the general prevalence of positive interactions in dimeric and trimeric sugars (Figure 3-6A) but may not explain positive interactions in simple carbon sources like monomeric sugars and TCA cycle intermediates (Figure 3-6A, B). Third, the facilitator may have excreted incompletely oxidized metabolites that were utilized by the facilitated strain [415, 239, 430]. Such "overflow metabolism" would allow strains to indirectly benefit from the biochemical transformation capabilities of their facilitators. Exploitation of newly created niches could explain the positive interactions we observed on simple carbon sources (e.g. the excretion of short-chain fatty acids as a byproduct of

incomplete monosaccharide oxidation). It may also explain the rarity of positive interactions on lower carbon source concentrations, since respiration is known to be favored over fermentation under such conditions and overflow is less likely to occur [430].

Despite the high throughput of our experiment, our system did not capture real-world bacterial diversity or environmental complexity. Our strain library was limited to two taxonomic orders isolated from topsoil. We only chose strains that grew on a minimal medium as part of our culturing protocol, possibly biasing our dataset in a variety of ways, e.g. against obligate facilitations for interactions involving amino acid or vitamin auxotrophies, which are known to be common [431]. Strains were also pre-grown on glucose as the sole carbon source prior to construction coculture/carbon source combinations, imposing glucose consumption as a requisite for inclusion in our strain library and potentially affecting bacterial physiology (e.g. lag phase) and interactions. Moreover, while our carbon source library represented a variety of carbon source types, it was limited to soluble compounds, excluding many polymers on which metabolically-driven positive interactions may be more common. Whether our results extend across additional phylogenetic groups (i.e. those occurring within soil and in other microbial ecosystems) and nutrient environments (i.e. across different and/or multiple carbon sources, concentrations thereof, and non-carbon nutrient requirements) should be investigated in follow-up studies to generalize trends observed in our system.

Our results indicate that knowledge of how strains grow individually in an environment can be strongly predictive of how they interact in that environment. In contrast, knowing how the same strains interact in a different environment or how different strains interact in the same environment do not appear to be very informative. Finally, our results suggest that a potential strategy for inducing the growth of a non-growing or weakly growing strain, independent of growth medium, is to coculture it with a strongly growing strain. Here we uncovered several general, statistical rules governing microbial community structure and function. Such rules deepen our understanding of microbial community ecology and are crucial to enable the efficient

design and control of beneficial microbial communities.

3.5 Materials and methods

Strain isolation from soil samples

Soil samples (two 12cm columns of topsoil, 4cm in diameter) were collected from multiple locations in greater Boston on November 12, 2017 (5.6°C) (specific locations listed in Table S1). Each sample was diluted in PBS within a few hours of collection (5 g soil vortexed in 40 mL PBS). Single strains were first isolated by streaking 70 µL of dilutions of this mixture (10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4}) on 20 different solid (agar) media [Tryptic Soy Broth (TSB) (Bacto), 1% v/v TSB, Lysogeny Broth (LB), 1%v/v LB, Nutrient Broth (NB), 1%v/v NB, M9 salts (Sigma-Aldrich) + 0.5%w/v glucose, M9 salts + 0.005%w/v glucose, M9 salts + 0.005%w/v glucose + 0.2%w/v casamino acids, M9 salts + 0.005%w/v glucose + 0.002%w/v casamino acids, M9 salts + 0.5%w/v glucose + 0.2%w/v casamino acids at pH = 4 and 5, M9 salts + 0.005%w/v glucose + 0.002%w/v casamino acids at pH = 4 and 5, Actinomycete Isolation Agar (Teknova), Brucella Agar, Streptomyces Medium, Campylobacter Medium, Bordetella Medium, and ATCC Medium 1111].

Strains were selected based on the following criteria: growth in LB liquid medium of transferred colony (25°C), frozen glycerol stock revival in LB (OD600 $>$ 0.1) (30°C), and subsequent growth on M9 + 0.5%w/v glucose (OD600 $>$ 0.1) (30°C). We kept 96-well plates of isolates (LB, 25%v/v glycerol) at -80°C. Isolates included in the coculture experiment (Table S1) were selected based on robust revival from glycerol stocks and in subsequent culturing steps and the ability to label the strains via constitutive expression of GFP.

Strain labeling

The soil bacterial isolates were fluorescently labeled with the commercially available plasmid pMRE132 containing GFP2 [432]. pMRE132 is a broad-host range plasmid used to constitutively express fluorescent protein genes in bacteria. First, the soil isolates and the *E. coli* carrying pMRE132 (Ec-pMRE) were grown to saturation

(5 mL of LB media [Bacto] in 50 mL Falcon tubes, loose caps, 30°C, 24 hrs, 15mg/L Chloramphenicol [Sigma-Aldrich] for Ec-pMRE). Second, the saturated culture (SC) of each soil isolate was mixed with the SC of Ec-pMRE (500 µL of each). This and subsequent steps were done in 2 mL 96 deep-well plates (Eppendorf) with a Viaflo 96 liquid handler (Integra) to increase throughput. The mixed SCs were concentrated to 10X by centrifuging (1 min, 7,000 RCF), discarding 900 µL of the supernatant, and re-suspending. Immediately after, 10 µL of each SC was spotted, in three replicates, onto nutrient agar (5 g peptone, 3 g yeast extract, 15 g agar [Bacto] in 1L water). The agar plates were incubated (24 hrs, 30°C) to allow for conjugation and transfer of the plasmid from Ec-pMRE to the soil isolates. The bacterial lawns were picked with sterile wooden applicators (Puritan), suspended in phosphate buffered saline (Corning), and mixed by pipetting 30 times with Viaflo. The suspensions were then serially diluted (100, 10-1 and 10-2) and spotted onto minimal media agar to select for the transconjugants (1X M9 minimal salts [Teknova], 5 g/L Glucose, 15 mg/L Chloramphenicol [Sigma-Aldrich], 15 g agar [Bacto]). Fluorescent colonies were picked and streaked onto minimal media agar to counterselect Ec-pMRE. Finally, SC from single colonies were frozen with 50%glycerol and stored at -80°C.

Multichoose kChip set-up

We generated droplets that each contained the following: (1) Hf-GFP (starting OD600 = 0.02); (2) one isolate (starting OD600 = 0.02, chosen among 14 isolates + 1 no-isolate control + 1 negative control) such that the synthetic communities contained the same initial [Hf-GFP]:[total isolate] ratio if no control droplets were present; and (3) one carbon source. All droplets that received the same carbon source were loaded onto the same kChip such that droplet grouping produced combinations of k = 1, 2, ..., 7, or 19 inputs with the carbon source type and concentration held constant (Figure 5-7). Here we also differentiate between "composition" as a given isolate subset of size s, e.g. the pair [A + B] (s = 2), and a "community" as a larger set of size k that contains the given composition and $\geq$1 additional isolates, e.g. all communities [A + B + X + ... + Y] (k $\geq$ 3). Overall, we produced $\sim$100,000 assay points (which were evenly divided among the carbon sources).

Strain identification and calculation of phylogenetic distance

Each bacterium isolated from soil was classified phylogenetically with its 16S rRNA gene sequence. The full 16S gene sequences (1500bp) were obtained via Sanger sequencing (Quintara Biosciences), quality threshold-trimmed, and classified with Seqmatch 48. Sulfolobus solfataricus, a thermophilic archaeon, was used in the phylogenetic reconstruction as an outgroup species to root the tree. MUSCLE 49 with default parameters was used to align the sequences. PhyML-SMS 50,51 with default parameters was used to select GTR+G+I as the best model and to infer the tree. The inferred branching order and branch lengths in the phylogenetic tree are the maximum likelihood estimate given the sequence alignment. The pairwise phylogenetic distances between strains were calculated directly from the phylogenetic tree with R (Phytools).

Microbial culturing

All labeled and unlabeled monocultures initially underwent two pre-experiment regrowth cycles ("starter phase" in a rich medium and "preculture phase" in minimal medium) and, at the onset of the experiment ("experiment phase"), were normalized to a starting density of OD600 = 0.02 in carbon-less minimal medium.

In the starter phase, glycerol stocks of the unlabeled and labeled strains were inoculated into 525 µL (0.8-mL-deep 96-well plate) of Lysogeny broth (LB) medium (25°C, 220 RPM, 16 hr). Inoculations from glycerol stocks were conducted via pin replicator (sterilized via 70% v/v ethanol bath and flame treatment between inoculations). In the preculture phase, all cultures were washed in carbon-less M9 medium two times and then diluting cultures (1/50) into 1 mL M9 medium with 0.5% w/v glucose (25°C, 220 RPM, 24 hr). Finally, the experimental phase began by washing cells three times in a carbon-less M9 medium to remove residual glucose and normalizing to a starting OD600 of 0.02 (or ∼20 cells/droplet depending on the strain).

Validation experiment in microtiter plates

To validate the interactions measured on the kChip, we cultured in 96-well plates all possible monocultures and cocultures from 7 bacterial strains in 4 different carbon

sources. The resulting 84 conditions represent 2% of the conditions measured on kChip. Each coculture condition in this follow-up experiment was grown in triplicate, and each monoculture in six replicates. The bacterial strains (RP1, KA, PAg2, LA, PP, PR2, PK, and a No-Strain control) were selected to cover a wide range of resource utilization profiles, and the carbon sources (Uridine, Proline, Sucrose, Fumarate, and a No-Carbon control) were selected to cover a wide range of bidirectional interaction types (Figure 3-6A).

The labeled and unlabeled bacterial strains were grown in monoculture in a starter phase (rich medium) and preculture phase (minimal medium) as described above. After washing the precultures with carbon-less M9 medium, and normalizing to a starting OD600 of 0.44 (22X), we mixed the labeled vs unlabeled bacteria in a matrix-like format in a 96-deepwell plate (100 µL each, 11X concentration each). Extra replicates of monocultures and intra-strain cocultures were added to the remaining columns of the same 96-deepwell plate. The four carbon sources were prepared in M9 media as described above, but to a concentration of 0.55% w/v. The experiment was initialized by mixing in triplicates (black-walled 96-well plates, Invitrogen) 136.4 µl of carbon source and 13.6 µl of bacterial mix (bringing the carbon and bacterial concentration to 0.5% w/v and 1X OD600 = 0.02, respectively). The plates were covered with gas-permeable AeraSeal membranes, and incubated for 72 hrs with shaking (220 rpm) and at room temperature (22-25^{o}C). Measurements of OD600 and GFP (ex. 480, em. 515, Optimal Gain, Z-Pos. 20 mm) were taken at 0, 24, 48, and 72hrs with a microplate reader (Infinite 200 Pro, Tecan). Before the 72hr time point, each well in the plates was mixed 20 times with a liquid handler (Viaflo 96, Integra) to break up accumulated biofilms. After the 72hr time point, one plate-replicate of Proline and Sucrose were serially diluted, and the dilutions 10-4, 10-5, and 10-6 were plated onto Nutrient Agar (3g yeast extract, 5g peptone, and 15g agar, Bacto, in one liter of water) with and without 15mg/L Chloramphenicol (plasmid pMRE132 confers resistance). The agar plates were incubated at room temperature for two days, and then the colonies were counted with the aid of a stereo microscope (Leica MZ10 F).

The OD600 values were normalized by subtracting the media-only control. The

GFP values were normalized by dividing by the value of the media-only control from the corresponding time point, since each time point was measured using a different optical gain. Monoculture fold-growth calculated using GFP and OD600 values strongly agreed (Figure 3-5), indicating that bacterial growth can be well-quantified using GFP measurements. To compare the monoculture fold-growth in 96-well plates and on the kChip (Figure 3-5B), the kChip GFP signal of each bacterial strain was normalized by subtracting the 0.2 percentile signal at 0hrs. This subtraction corrects for the background fluorescence of the kChip wells with zero or very few cells (580 arbitrary units of fluorescence), allowing for an estimation of the GFP signal at 0hr due to cells. To calculate the bacterial pairwise effects with GFP data from 96-well plates (Figure 3-5C-D), we utilized the same formula and made the same background normalization as before (Figure 5-12, 5-13), accounting for the inoculum effects on a per-replicate basis and for possible growth on media without carbon. For the Detection Limit (DL) in GFP measurements, we utilized the mean yield of strain A in coculture with strain B with no carbon source. No DL or background subtraction was used for the calculation of effects based on colony counts. Bidirectional interaction (Figure 3-5E) types (Θ) and magnitudes (m) were calculated from the corresponding unidirectional interactions as described for the kChip measurements (Figure 3-3F).

Chapter 4

Conclusion

This thesis shows evidence of bacterial interspecies interactions regulating community assembly in the *C. elegans* intestine and in nanoliter droplets. In both experimental systems, we measured interactions between bacterial species at larger scales than what is usually performed in the field of microbiome sciences. These higher throughput experiments allowed us not only to show that interspecies interactions occur, but also to assess their prevalence in microbial systems. In *C. elegans*, competitive and hierarchical interspecies interactions altered the composition of at least half of the two-species intestinal microbiotas we built and measured. Interspecies interactions were also influential on the assembly of more biodiverse microbiotas, on different worm mutants with different immune activity levels, and on bacteria isolated from worms' intestines or without previous exposure to them. Meanwhile, in the kChip, strongly growing bacteria consistently enabled the growth of strains unable to grow in monoculture, demonstrating a wide occurrence of positive interspecies interactions, mainly in the form of commensalisms and parasitisms. In conclusion, interspecies interactions are a prevalent driver for microbial community assembly.

This thesis contains two microbial ecology projects, each one with a core central idea and also a more innovative, forward-looking proposition. On the *C. elegans* microbiota project, the central theme is the experimental demonstration of the importance of bacterial competition for the assembly of an intestinal microbiota. Such idea is not only thoroughly supported by our data, but it is also widely congruent with the

current view on the microbiome sciences' field. Apart from this, our results on the *C. elegans* microbiota propose that bacteria native to (isolated from) the worm gut are not necessarily more fit to colonize such environment. This forward-looking proposition does not fully align with the developing ideas of phylosymbiosis in the field, which argue that most organisms in natural biological systems are fully co-adapted and co-evolved to live with one another [433]. The central idea of the kChip project, which also aligns well with the field, is that overflow metabolism (cross-feeding) occurs frequently in microbial systems; while the more innovative idea we set forward is that for any poor-growing bacterial species in a given carbon source, there is likely another bacteria capable of benefiting the growth of the former. In conclusion, I expect the central ideas in this thesis to be readily integrated in our current understanding of microbial systems, while the secondary propositions will perhaps incentivize further discussion and better studies in the field.

This thesis has also delved into how the bacterial community context plays a fundamental role in the survival outcome of a bacterial species. Bacteria and other microorganisms are important, among other reasons, due to their capacity to infect and sicken humans. Such dysbiotic states sometimes occur due to uncontrolled bacterial proliferation, which pushes the fractional abundances of a resident microbiota out of equilibrium. The results presented in this thesis hint at the possibility of harnessing the competitive ability of resident bacterial members to prevent the human microflora from being infected. Interspecies interactions are prevalent in microbial communities built in-vitro and in the gut of *C. elegans*, and our all-vs-all pairwise experiments show that the abundance of all bacterial species can be affected by the presence of other community members. Moreover, in this thesis I have illustrated how simple linear forecasts based on feasible experiments have good predictive power of the fractional abundances of bacterial communities of up to 8 species. Future microbial ecology studies of important human ecosystems, as well as the development of infection prevention therapeutics, might benefit from the gain in understanding of bacterial interspecies interactions and microbial community assembly presented in this thesis.

4.1 Follow up experiments and future directions

I provide here some possible follow up experiments that could help to extend the results presented throughout this thesis. Among the many paths to move forward, I believe these specific ones would be appropriate to fill the conceptual gaps in the field and consolidate our understanding of bacterial interspecies interactions and microbial community assembly.

Developing a comprehensive set of fluorescent microbes, beyond Gamaproteobacteria, would help to assess the generality of the positive pairwise interspecies interactions that we measured. This will help to verify that the bacterial bias to not easily receive a fluorescent tag is not also a bias towards different interspecies interactions. The distribution of bidirectional interactions explored in Chapter 3 (Figure 3-4) utilize bacteria from the orders Pseudomonadales and Enterobacterales, both of which belong to the Gammaproteobacteria class (Proteobacteria phylum). Although normal isolation protocols serve well to isolate bacteria from other classes and phyla, such as Firmicutes, Cyanobacteria, Bacteoridetes, and Actinobacteria, it is less straightforward to fluorescently label bacteria from such clades. The protocol we utilize to label bacteria (pMRE plasmids [432], Methods 3.5) is described as a broad host-range insertion protocol that can be implemented as a transposon chromosomal insertion or as a simpler plasmid conjugation. Our labeling efforts with such protocol so far have resulted in a diverse set of fluorescent Pseudomonadales and Enterobacterales. To enlarge this set of bacteria, further labeling attempts could incorporate 1) an electroporation step, 2) an osmotic shock, and/or 3) a temperature shock when co-culturing the tag-donor *E. coli* and the recipient bacteria. With a larger set of fluorescent microbes in hand, a new round of kChip screening experiments[1] could be performed to measure bidirectional interactions and assess if the distribution of positive and negative interactions that we observed before is conserved beyond the Gammaproteobacteria.

Utilizing the conditioned (leftover) media of one bacteria to grow other species

[1]Concerto Biosciences, a novel enterprise implementing the kChip technology, could serve as an alternative to run such experiments.

in the kChip could reveal bacterial inhibition patterns at larger scales than before. Although these conditioned media experiments have been explored multiple times before, it would be reasonable to utilize the kChip, together with a comprehensive set of fluorescent microbes, to scale-up such results. Specifically, having multiple species from the same genus, together with multiple phyla represented in the same screen could allow to reconcile previous studies reporting increasing [159, 160] and decreasing [161] probabilities of bacterial inhibition as a function of phylogenetic distances. These conditioned media experiments would test for bacterial inhibition by means of excreted compounds into the environment, which is a common form of interspecies interactions [210].

A long-term co-culturing experiment, where a small set of pairs of bacterial species are grown across successive passages, could serve as a test for evolutionary patterns in simple microbial communities. Two specific hypothesis to test with such long-term co-culturing experiment are: 1) if there is a decrease in the frequency and magnitude of bacterial pairwise inhibition, and 2) if there is an increase in the productivity and yield of the co-cultured strains to grow in the metabolic byproducts of the partners. A recent study by Meroz et al. (2021) [434] explores a similar experimental framework of evolution in microbial communities, where they co-culture bacterial species known to coexist, and they keep track of the changes in community composition. Instead of propagating coexisting species, I propose to introduce *de novo* one of the partners at each dilution passage–this would lead to one bacterial partner staying virtually unchanged, while the other partner would perhaps accumulate mutations and evolve. On the order of 5-10 species, conditioned with each of the other ones, would lead to 25-100 different evolved pairs, which is a feasible number of conditions to grow over approximately 400 generations. The original interaction between species A and B would serve as the comparison control for the novel interaction between A_B and B.

Other ideas for further experiments are: 1) Building more complex microbiotas by using very low nutrients, longer timescales and/or porous solid environments. 2) Select for a bacterium that does not leak byproducts by co-culturing a handful of high-growing bacteria with a non-grower in a spatially defined medium. Uti-

lizing a non-grower species that is highly competitive by means of producing toxins or other compounds capable of killing at local scales could cause the growing bacterial species to be inhibited if they leak byproducts and enable the growth of the non-grower. 3) Extend the comparisons in the *C. elegans* microbiota between MYb strains and phylogenetically similar non-native strains. If there is still little difference in competitive ability between native and non-native strains, figure out why the MYb haven't evolved higher fitness for the worm gut environment.

Chapter 5

Appendix

5.1 Bacterial competition in the *C. elegans* gut

5.1.1 Extended materials and methods

A synchronous adult population of *C. elegans* was obtained by: 1) Collecting gravid worms from 8 to 10 plates and extracting eggs with the standard egg prep protocol [396]. 2) Letting the eggs hatch on M9 Worm Buffer (WB) (3 g KH2PO4, 6 g Na2HPO4, 5 g NaCl, H2O to 1 liter. Sterilize by autoclaving and then add 1 mL 1 M MgSO4. Sigma-Aldrich for all chemicals, unless otherwise specified) overnight to arrest development of L1 larvae. 3) Transfer larvae to NGM plates with lawns of *E. coli* OP50 and incubate at room temperature for 2 days. The synchronized adult worms were then transferred and kept for 24 hours in 50 mL Falcon tubes (Corning) with 5 mL S medium, 100 µg/mL gentamicin, and 5X heat-killed *E. coli* OP50 to kill any bacteria inhabiting the intestine, resulting in germ-free worms. S Medium: 1 liter S Basal, 10 mL 1 M potassium citrate pH 6, 10 mL trace metals solution (Teknova), 3 mL 1 M CaCl2, 3 mL 1 M MgSO4. Add components using sterile technique; do not autoclave. S Basal: 5.85 g NaCl, 1 g K2HPO4, 6 g KH2PO4, H2O to 1 liter. Sterilize by autoclaving and then add 1 mL cholesterol (5 mg/mL in ethanol). Heat-killed OP50 was prepared by concentrating 50x a saturated culture of *E. coli* OP50, and it was used to trigger feeding behavior in the worms. The adult worms were washed

via sucrose flotation to remove debris before bacterial colonization.

The natural microbiota strains of *C. elegans* were isolated by growing *C. elegans* on different types of rotten organic material, followed by washing and sterilizing the worms on the outside, grinding the worms, and plating the resulting bacterial suspension on agar plates. Different types of compost and rotten fruits and vegetables were fed to the worms. Rotten apples were directly collected from the outside. Other fruits like apples, celery, almonds and parsnip were placed on local soil from Boston, MA in a household plastic box (Sterilite) with semi-open lid and incubated at room temperature until the fruits were strongly decayed (3 weeks). The compost samples were taken from two local compost piles in Boston, MA, that mostly contained kitchen and garden waste. Some amount of PBS and glass beads were added to the samples. The samples were homogenized by vortexing at high speeds. The resulting solution was filtered (Millex-SV 5 µm, MerckMillipore) to remove bigger particles. The resulting emulsion was spread on S media agar plates without citrate. *C. elegans* N2 were first grown on OP50 lawn on NGM plates, sterilized with antibiotics (5 mL S medium, 100 µg/mL gentamicin, 5X heat-killed *E. coli* OP50) and added to the plates with rotten organic material for approximately one week. After that time, the worms were washed off the plates with M9 + 0.1% Tx. The worms were washed twice with 1 mL WB + 0.1% Tx (centrifugation at 2K RCF, 10s). Afterwards the worms were re-suspended in 1 mL ice-cold WB + 0.1% Tx and incubated on ice for 10 mins. 2 µL bleach (Clorox) were added and the worms were incubated for 6 mins on ice. Afterwards the worms were washed 3x with ice-cold WB + 0.1% Tx. Single worms were transferred into 0.6 mL reaction tubes (Eppendorf) and ground with a motorized pestle (Kimble Kontes Pellet Pestle, Fisher Scientific) for at least 1 min. The resulting solution was plated onto a tryptic soy broth (Teknova) agar plate (2% agar, 150mm petri dish). From the resulting colonies, physiological unique colonies were picked, and streaked-out again on tryptic soy broth agar to isolate the strain. Finally, the bacteria were grown in tryptic soy broth at 30°C and stored as glycerol stocks. The species identity was analyzed by 16S Sanger sequencing (Genewiz).

The hierarchy score of a matrix with fractional abundances is calculated by: 1)

Ordering its rows and columns ascendingly based on the mean fractional abundance; and 2) Taking the mean value of the half-matrix under the diagonal. In a perfectly hierarchical matrix, each competitor will drive to extinction every other species with a lower rank, reaching a hierarchy score of 1. Random matrices were generated to calculate the significance of the observed high hierarchy score. We conserved the distribution of fractional abundances by: 1) Sampling with replacement 55 values of the original matrix; 2) Assigning these random fractional abundances to the lower triangle of a new matrix; and 3) Assigning to the upper triangle of the matrix the values of 1-transpose. For each random matrix generated, a new hierarchy score is calculated as previously described.

5.1.2 Supplementary figures

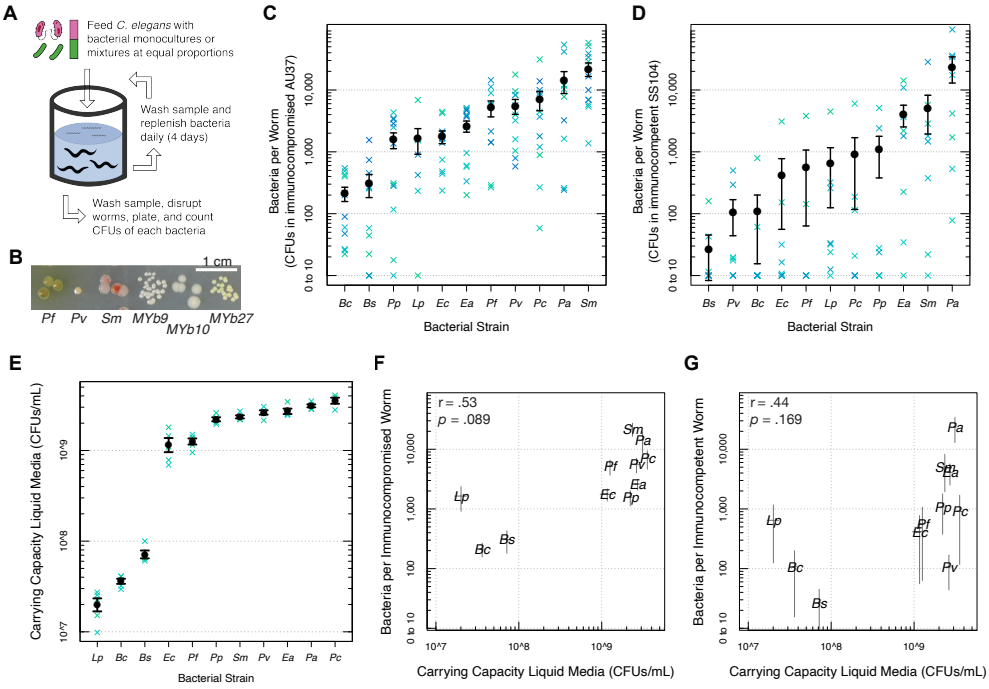

Figure 5-1: Monoculture colonization of *C. elegans* intestine is variable. (A) To construct and measure simple microbiotas in *C. elegans*, a defined number of bacterial species are fed in liquid culture to a same-age adult population of worms raised on *E. coli* OP50 and sterilized with antibiotics. The liquid feeding substrate is restored every day to maintain equal bacterial concentrations during the four days of colonization. After four days of feeding, worms are mechanically disrupted in batches of 20, plated onto nutrient agar in triplicates, and the added counts of colony forming units (CFUs) are used to determine bacterial population sizes in the worm gut. (B) Photograph of bacterial CFUs of six selected species used in this article, plated onto Nutrient Agar and grown at 25ºC for 2 days (*P. fluorescens, P. veronii, S. marcescens, Achromobacter MYb9, Acinetobacter MYb10, Arthrobacter MYb27*). (C-D) Different bacterial species reach widely different average population sizes (black points) during monoculture colonization of the *C. elegans* intestine. Each species also reaches widely different population sizes across biological replicates. Each blue cross is a different mechanical disruption of a batch of 20 worms, and error bars are the standard error of the mean (s.e.m.). (E) Bacterial carrying capacities in the liquid media used as feeding substrate (1%AXN, Methods) were measured by plating the saturated batch cultures (24hrs at 25ºC, OD600 in stationary phase) onto Nutrient Agar and counting CFUs afterwards. The blue crosses show the five biological replicates per species, and the error bars are the s.e.m. (F-G) Comparison of the carrying capacities in the liquid media and the population sizes in the intestine of *C. elegans* AU37 (F) and SS104 (G). The three Firmicutes (*Lp, Bs, Bc*) reach low population sizes in the worm gut and low carrying capacities in the liquid media, but the carrying capacities in the liquid medium don't explain the variation in monoculture colonization (Pearson correlation coefficients and p-values in the upper left corners). Bars are the s.e.m.

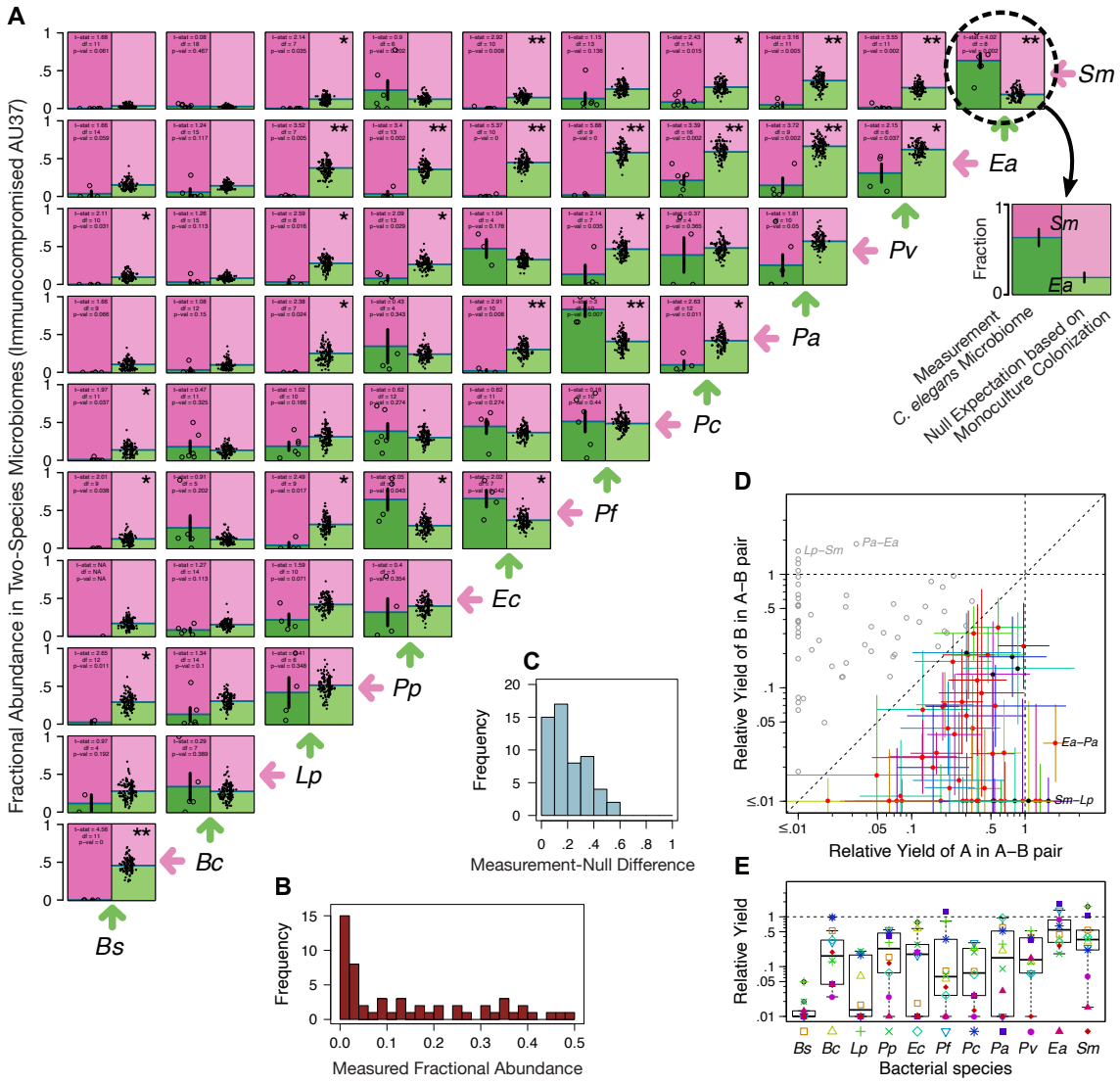

Figure 5-2: Two-species microbiotas in *C. elegans* intestine contain mostly competitive interactions.

Figure 5-2 *(previous page)*: (A) LEFT: Fractional abundances of 55 different two-species microbiotas constructed from 11 non-native bacterial species in an immunocompromised worm. Points are the replicates and error bars are the s.e.m. Most conditions had four or more biological replicates. RIGHT: Null expectation based on monoculture colonization. The underlying points, from where the mean null expectation was calculated, are the fractions obtained from combining the monoculture replicates of the two bacterial species, and the error bars are the s.e.m. * and ** represent a statistically significant difference between left and right panels at p-values of .05 and .01, respectively (Welch's T-test). Bacterial species are ordered from left to right by their mean fraction across the 10 different two-species microbiotas. (B) Histogram of the fractional abundances in the two-species microbiotas show that only a minority of the pairs displayed competitive exclusion (14/55 25%). From a pair of bacteria reaching fractions 57%-43%, only the lower quantity, 43%, was plotted. (C) Distribution of the differences between the left and right panels in (A). Across all pairs, the measured fraction deviates from the null expectation by a mean distance of 20%, with a peak at low distance and a long tail corresponding to cases where the interspecies interactions are particularly important (Figure S2C). (D) Relative yields (RY) less than one in two-species microbiotas are indicative of competition. The RY of a bacterial strain is calculated by dividing its population size in a co-culture experiment by its monoculture population size. Each point is the mean RY from 1000 bootstrap replicates sampling simultaneously over the pairwise data and the monoculture data. Error bars are the standard deviation of this bootstrap replicates (or the s.e.m. RY). The points were colored red, or black if the total abundance in co-culture (A+B) is lower or higher than the higher of the monoculture population sizes plus its s.e.m., respectively. Some points were labeled, error bars were colored, and dotted lines were added to improve clarity. The points above the diagonal are equivalent to the points below the diagonal. (E) Box plots of the RY of each bacterial species. Partner species is denoted by the colored symbols.

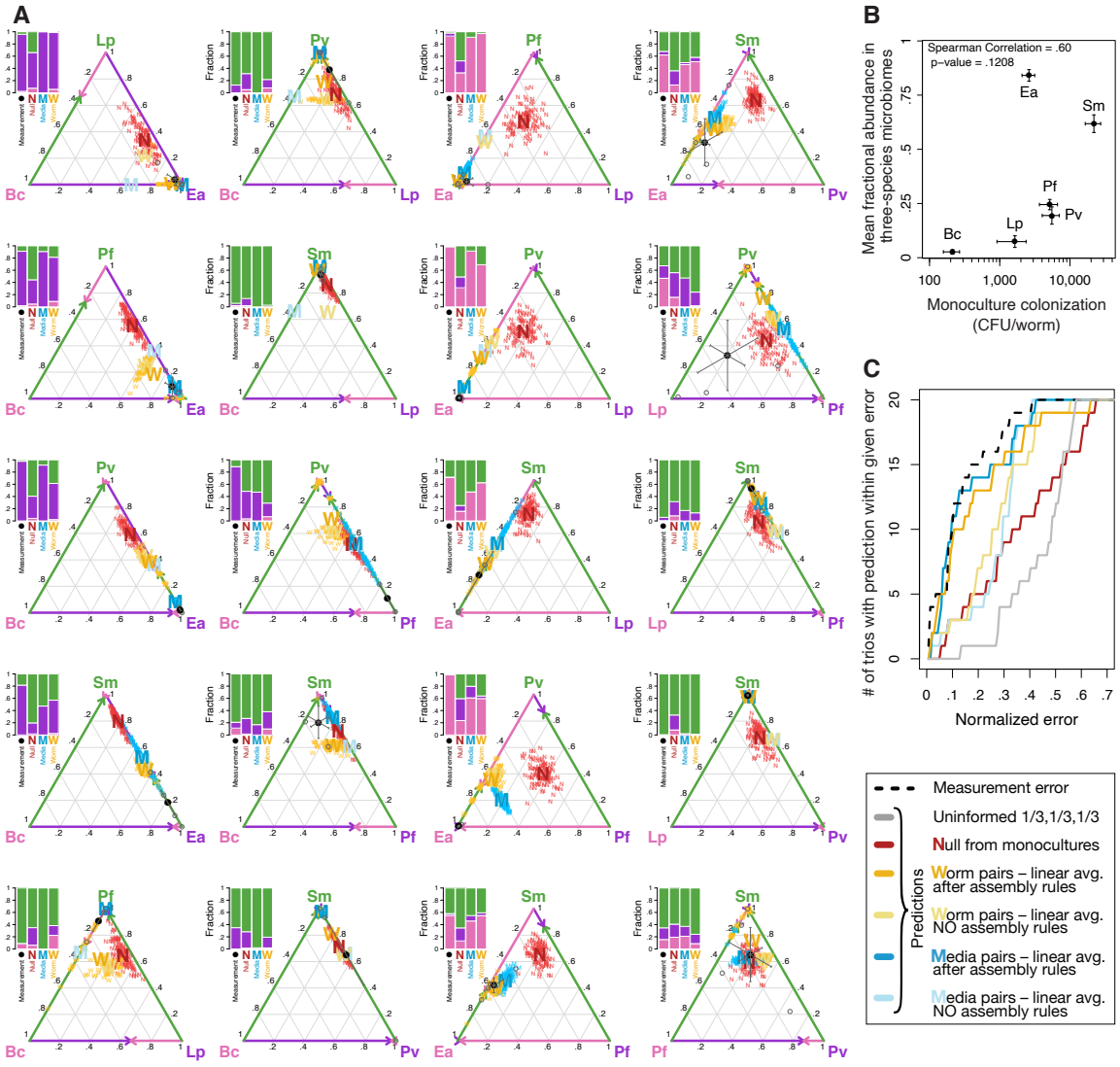

Figure 5-3: Three-species microbiotas in the *C. elegans* intestine are well predicted based on pairwise outcomes.

Figure 5-3 *(previous page)*: (A) Each set of bar graphs and the triangle (simplex) to the right displays the measured fractional abundances and predictions of one three-species microbiota. These 20 different three-species microbiotas are all the possible combinations with the species Bc, Lp, Pf, Pv, Ea, and Sm. The edges in the simplexes depict the fractional abundances in the two-species microbiotas. 'N': Null expectation based on monocultures, where each bacterial species reaches its population size in monoculture colonization. 'W': predictions based on two species microbiotas in worm gut (normalized arithmetic mean with or without assembly rules). 'M': predictions based on pairwise outcomes in vitro liquid media. The error bars on the measurement are the s.e.m. of 4 biological replicates, and the clouds of points around predictions are 100 bootstrap replicates ('N's sampling the monoculture data, and 'W's and 'M's sampling the pairwise data). (B) The mean fractional abundance in three-species microbiotas correlates mildly with monoculture population size. Error bars in Y-axis are the propagated error from the s.e.m. of the trios that include the given species. (C) Cumulative distribution of the error of the predictions of three-species microbiotas with and without assembly rules. The error of the predictions based on pairs can be reduced by removing a bacterial species from the trio prediction when it cannot survive all co-culture experiments—assembly rules. Errors calculated as the linear distance (L2 norm) between the fractional abundances in the prediction and measurement (normalized by the maximal distance, $\sqrt{2}$). The dashed line represents the mean distance between the measured mean and the biological replicates of each trio, and it is an upper bound for the error. The gray line is the error of an uninformed prediction of '1/3, 1/3, 1/3', and it is a lower bound for the error.

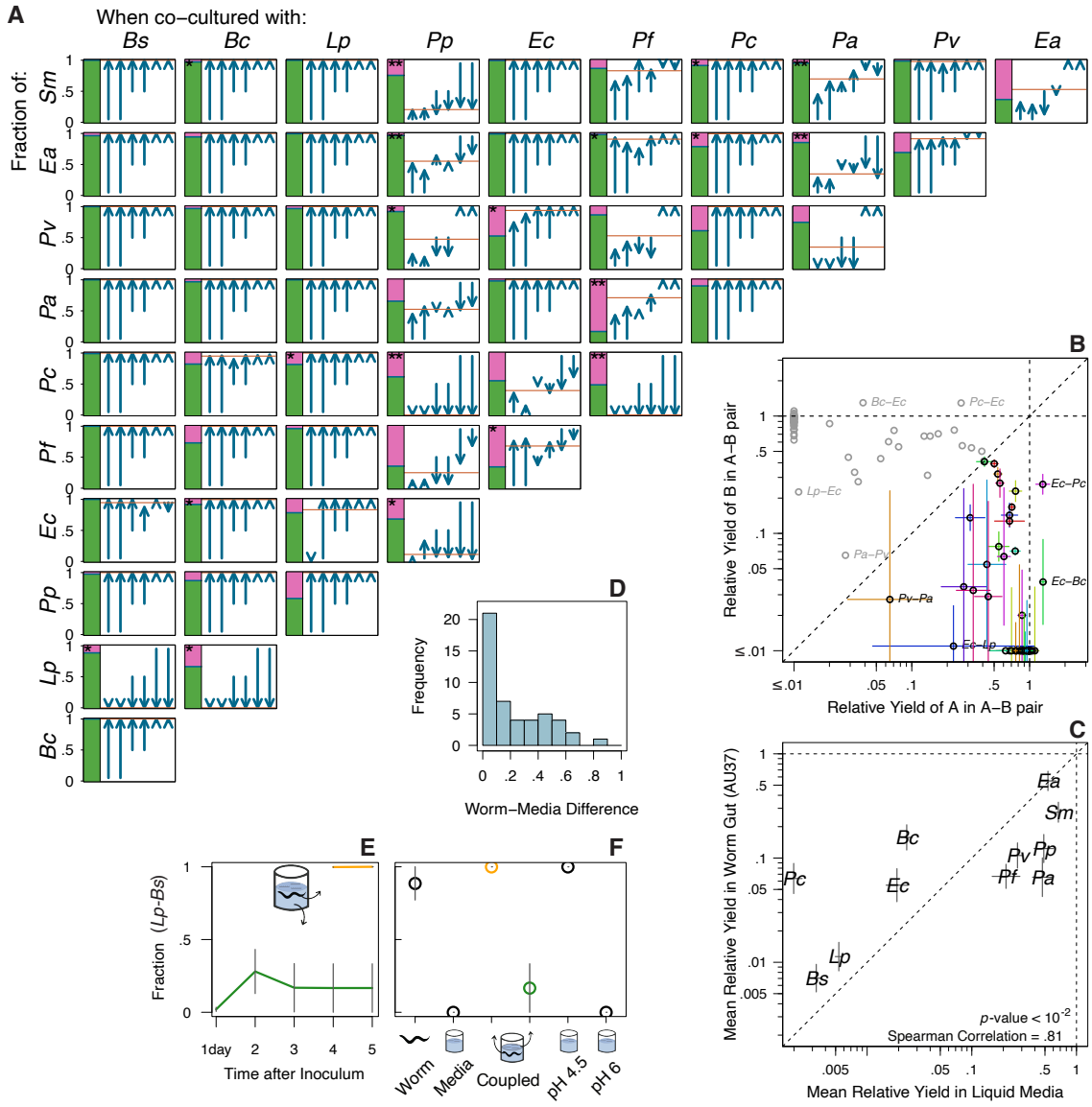

Figure 5-4: Co-culture experiments in vitro liquid media reveal the cases where the worm gut environment determines microbiota composition.

Figure 5-4 *(previous page)*: (A) Fractional abundances in co-culture experiments in the liquid medium used as feeding substrate (rich medium with peptone, yeast extract, etc., Methods), with each subpanel representing one bacterial pairwise combination, and each arrow representing a replicate. The beginning and the end of each arrow are the starting fraction of each replicate (95-5%, 50-50%, and 5-95%), and the measured fraction after 7 cycles of 100x daily dilution, respectively. Orange lines are the mean across replicates. Green-pink bar-plot on the left of each subpanel is the fractional abundance of the two-species microbiota in AU37 (Figure S2A). * and ** represent a statistically significant difference between worm and liquid media at p-values of .05 (19 cases) and .01, respectively (Welch's T-test). (B) Low relative yields in liquid media experiments indicate that competition is also the norm in this in vitro environment. Each point is the mean RY from 1000 bootstrap replicates sampling simultaneously over the pairwise data and the monoculture data; error bars are the standard deviation of this bootstrap replicates (or the s.e.m. RY). (C) Mean relative yield of each bacteria in vitro liquid media and in vivo worm gut are correlated. (D) Distribution of differences between co-culture experiments in vitro liquid media and in vivo worm intestine. (E) *L. plantarum* and *B. subtilis* reach different fractional abundances in vivo worm gut and in vitro liquid media on a coupled experiment. (F) An acidic version of the media resembling the average pH of the worm intestine (4.5) shifts back the pairwise outcome to a worm-like state; error bars are the s.e.m. of at least 4 replicates.

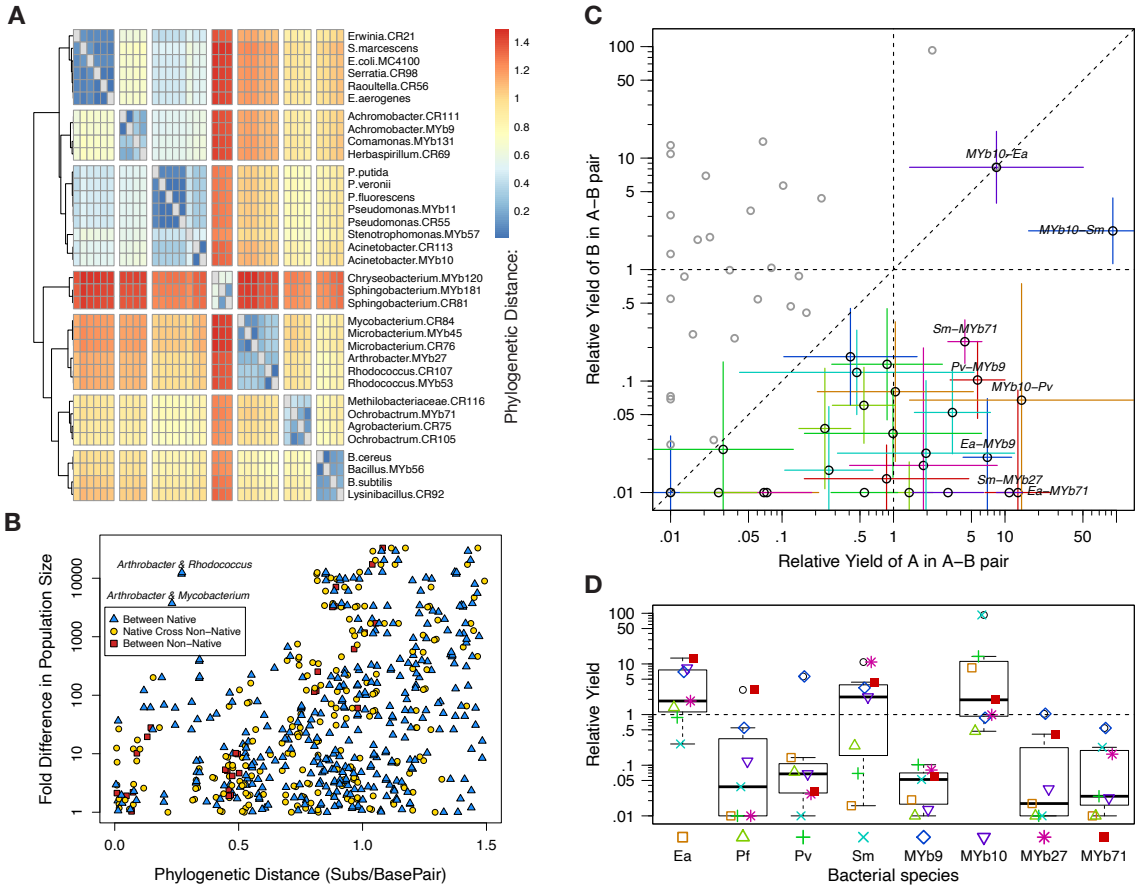

Figure 5-5: Bacteria isolated from *C. elegans* intestine show more facilitative interactions in two-species microbiotas.

Figure 5-5 *(previous page)*: (A) Heat-map of phylogenetic distances between native and non-native bacteria in substitutions per base-pair. The distances were calculated directly from the phylogenetic tree (Figure 4A) that was inferred with maximum likelihood from the multiple sequence alignment of the of full-length 16S gene of the bacterial species. Both axis are clustered into 7 bins for ease of interpretation. The phylogenetic tree matches the nomenclature of all bacteria, except for Stenotrophomonas MYb57, which fell outside the split between Gamma- and Betaproteobacteria instead of being inside the Gamma class. More genomic information might be needed to correctly classify MYb57 [435]. (B) Similar bacteria reach similar population sizes, and as phylogenetic distance increases, the difference in colonization ability tends to increase as well (rs=.39, p<10-15). Fold differences in population sizes calculated as the ratio of better colonizer over worse colonizer (1 added to each population size to avoid dividing by zero). The phylum Actinobacteria showed the highest variation in colonization (e.g. Arthrobacter and Rhodococcus), which suggests that genomic differences between members of the Actinobacteria might set them apart when colonizing the worm gut. Pseudomonas CR55 is also especially different from its close relatives. (C) RY>1 in two-species microbiotas are indicative of facilitation. Each point is the mean RY from 1000 bootstrap replicates sampling simultaneously over the pairwise data and the monoculture data; error bars are the standard deviation of this bootstrap replicates (or the s.e.m. RY). (D) Box plots of the RY of each bacterial species. Partner species is denoted by the colored symbols. *E. aerogenes, S. marcescens,* and *Acinetobacter MYb10* reach a RY>1 in repeated occasions, especially with native isolates as the co-culture partners.

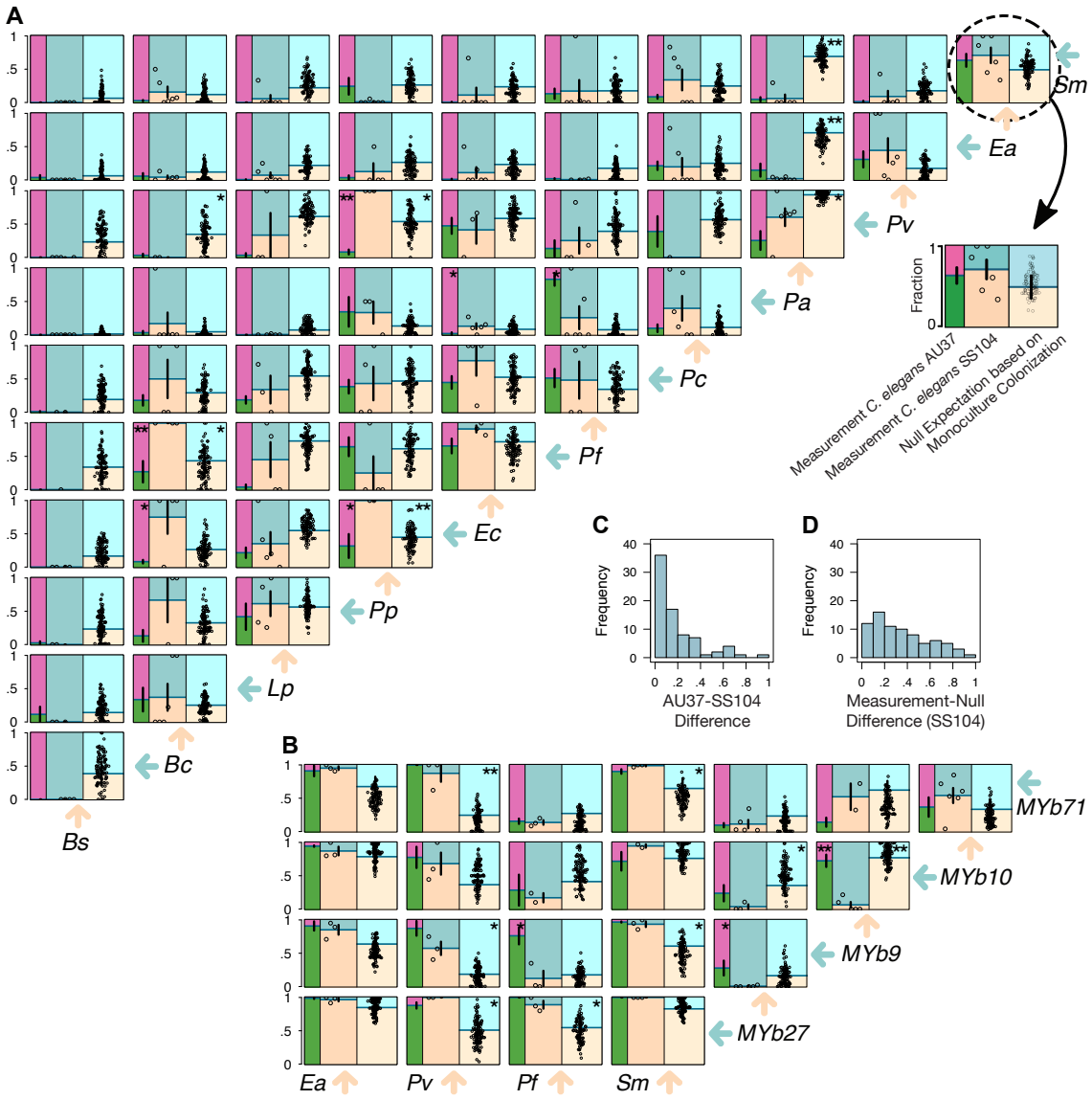

Figure 5-6: Two-species microbiotas in p38-immunocompetent *C. elegans* reveal the cases where the immune system determines microbiota composition. (A-B) The pink-green bar-plots display the fractional abundances of two-species microbiotas in immunocompromised *C. elegans* AU37 (previously shown in Figures 2A and 4C). The aqua-beige bar-plots display the equivalent two-species microbiotas and the null expectation based on monocultures for immunocompetent worm SS104. * and ** represent a statistically significant difference at p-values of .05 and .01, respectively (Welch's T-test). (C) 9 out of 77 two-species microbiotas are significantly different between immunocompetent and immunocompromised worms. (D) 15 out of 77 two-species microbiotas in immunocompetent worm are significantly different from the null expectation based on monocultures.

5.2 Massively parallel screening of synthetic microbial communities

5.2.1 Growth of labeled and unlabeled strains can be profiled across environmental conditions

kChip screening allows for rapid functional profiling of both fluorescently labeled and unlabeled strains across libraries of environmental conditions, e.g. antibiotics, natural products, and carbon sources, with desired temporal resolution (limited only by kChip scan time, <15 minutes at 2X magnification). Carbon utilization profiles, i.e. growth curves for each strain across different single carbon sources in a minimal medium (SI Appendix, Section S2), were attained for a panel of droplet monocultures as well as conventional 96-well plate monocultures for comparison (SpectraMax plate reader). We pooled a library of microbe-containing droplets with a library of carbon source-containing droplets and loaded the droplets onto a k = 2 Chip (i.e. all microwells on the kChip grouped 2 droplets). From microwells that received one droplet from each library ($\sim$ 1/2 the total microwells on the k = 2 Chip, $\sim$17,000), we profiled the growth of each strain on each carbon source.

To track growth on each carbon source, we used one of two assays: (1) measurement of a constitutively expressed fluorescent protein (GFP or YFP) (Figure 5-7A-C); or (2) reduction of resazurin dye to its fluorescent product resorufin by cellular metabolism (proportional to cell density), a label-less assay that can be used with unlabeled or genetically-intractable strains (Figure 5-7E-G). We selected 10 fluorescently labeled strains, and first confirmed that glucose utilization was recapitulated on a k = 1 Chip. We then crossed this panel with 13 carbon sources. Carbon utilization profiles produced from GFP or YFP signal on a k = 2 Chip and 96-well plates agreed strongly (Pearson r = 0.868) (Figure 5-7C) with on-chip consistency between technical replicates (R^2 = 0.968). To assay growth of unlabeled bacteria, the resazurin assay, which has previously been used to quantify cell viability in droplets [436], was used by incorporating resazurin in all carbon source-containing droplets. For a panel

of three strains profiled across four carbon sources, growth curves produced via resorufin fluorescence measurements on a k = 2 Chip and OD600 measurements on a 96-well plate agreed strongly (Pearson r = 0.969) (Figure 5-7).

5.2.2 Outlook and perspective for the kChip screening platform

Droplet microfluidics have recently been applied across a diverse range of assay types in single-cell transcriptomics, drug discovery, and microbiology [279, 437]. The kChip platform expands upon these technologies to enable the rapid construction and high-throughput screening of beyond-pairwise species combinations. Here [259] we have demonstrated that the kChip screening paradigm is compatible with fluorescently labeled species, diverse environmental isolates (14 tested, Table S3), and environmental conditions (16 tested, Table S2). We showed that the platform enables phenotypic screening of fluorescently labeled species across combinations of biotic (isolate) and abiotic (carbon source) settings, as well as growth profiling of unlabeled isolate libraries via the resazurin assay. Demonstrating the utility of kChip screening, we discovered and validated compositions that facilitate the model plant symbiont Herbaspirillum frisingense in a manner robust to carbon source environment and community background. We further extrapolated ecological trends in the data, generated hypotheses about consortia design principles, and demonstrated consistency with these principles for top-scoring compositions identified in the screen.

Data generated through kChip screening is a valuable resource to explore the underlying ecology of cellular interactions among microorganisms (bacterial, algal, and/or fungal) and their environmental dependencies. Taking the carbon sources glucose and galactose, for which conserved glycolytic pathways are used ubiquitously by the bacterial kingdom, we observed pervasive competition that we expect is attributable to nutrient competition. Taking the more complex oligomers (sucrose, lactose, and raffinose) on which Hf-GFP monocultures grew poorly in monoculture, a possible explanation for the common facilitation we observed (Figure 3-2A) is the

secretion of enzymes that increase carbon availability to Hf-GFP [438]. In the case of sucrose, for example, we can speculate that extracellular enzymes produced by facilitating strains cleave sucrose into glucose and fructose, monomers that are then utilized by Hf-GFP. We found that these facilitative effects were typically robust to community context, suggesting that this facilitation is driven by key interactions and agnostic to the presence of additional strains. Improvement to the median yield of Hf-GFP with community richness (Figure 3-2A) could be explained by the probability of sampling primary facilitator strains that individually facilitate Hf-GFP. In the case of sucrose, we observed that the effect of a primary facilitator could be bolstered by additional strains, suggesting additive and/or higher-order effects. Finally, the rarity of robustness to both carbon source and community context (Figure 5-8E) suggests that facilitative mechanisms depend highly on the environment and indicates a need for testing under many conditions to identify mechanisms and interactions robust to carbon source variation.

The kChip has numerous applications in elucidating microbial community ecology, namely the phenotypic characterization of a given species or cell type across an array of biotic and/or abiotic settings. Datasets can be used to parameterize or assess mathematical models of growth or interactions as well as to determine of how biotic metrics (e.g. the genetic and metabolic diversity of co-cultured species) and abiotic factors (e.g. the complexity, concentration, or ratio of carbon substrates provided) drive metabolic decision making and interactions. Screens can also be used to detect when higher-order interactions emerge that are unpredictable from measured pairwise interactions [263, 266, 267, 245, 168] and to produce hypotheses about community design principles [439, 440] and the environments that induce desirable interactions [441, 413, 442].

Beyond basic ecology, kChip screens could be used to identify promising compositions for development into probiotics. Inspired by the success of microbiota transplantation to counter ecological dysbiosis in animal and plant hosts [443, 444], standardized interventions remain difficult to develop for a variety of reasons including a lack of mechanistic understanding and the explosion of possible strain combina-

tions. Analogous to in vitro compound screening to generate therapeutic candidates, kChip screens can generate short lists of "hit" microbial cocktails that are also robust to relevant biotic and abiotic perturbations and constitute attractive candidates for validation and follow-up studies. For example, combinations of soil species can be identified that robustly facilitate plant growth-promoting rhizobacteria (PGPR), which have been shown to improve crop yields substantially [410, 445] by providing the plant with nutrients and resisting pathogen colonization [446]. On the other hand, screens to identify combinations that robustly suppress the growth of pathogens may be particularly useful in the context of dysbiotic human microbiomes. Indeed, defined probiotics are under development to address infections like Clostridium difficile [447], Salmonella [448], vancomycin-resistant enterococci [449], and Staphylococcus aureus [450]. Screening combinations of species from healthy, pathogen-resistant microbiomes may expedite probiotic discovery or identify higher-performing formulations. More broadly, any optically detectable community-wide phenotype can be screened, e.g. the degradation of a fluorescently labeled recalcitrant organic compound (or the growth of a fluorescently labeled species that consumes one of its byproducts); community-induced changes in gene expression (via promoter-GFP reporter fusions); and the production of cryptic, interaction-mediated metabolites that impact growth, such as antimicrobials [451].

5.2.3 Extended materials and methods

Microbial Culture Input Preparation.

All bacterial cultures underwent an initial "starter phase," whereby glycerol stocks of environmental isolates and fluorescently labeled strains were inoculated into 525 µL of lysogeny broth (LB) medium (2 mL deep, 96-well plate via pin replicator) and 4 mL of LB (15 mL culture Falcon tube), respectively (30 °C, 220 rpm, 16 h). A subsequent "preculture phase" (30 °C, 220 rpm, 24 h) consisted of washing cells and resuspending them in M9 minimal medium (MM) with 0.5% (wt/vol) glucose at an initial OD600 of 0.01. The "experimental phase" consisted of washing and resuspending cells typically to a starting OD600 of 0.02 in MM (or ∼20 cells per droplet).

kChip Input Preprocessing.

For the screen described in Figure 3-1, each droplet contained Hf-GFP, a single cocultured isolate, and a single carbon source. Every input received a "color code," a unique ratio of three fluorescent dyes (1 µM or 10 µM) before generating droplets. Every input received 0.05% (wt/vol) bovine serum albumin to aid in retention of hydrophobic small molecules during droplet pooling and loading (32, 45).

Droplet Making and kChip Loading.

Droplets were produced on a Bio-Rad QX200 Droplet Generator in a fluorocarbon oil (3M Novec 7500). Droplets were pooled to prepare a total of 200 μL of droplet suspension (∼5 min) and injected into a custom-built kChip loading apparatus [259]. Each microwell randomly sampled k droplets (∼5–10 min). The kChip was scanned at 2× magnification to identify the droplets in each microwell from their color codes (∼12–15 min). Droplets were merged within their microwells via ∼10 s of exposure to an alternating-current electric field (4.5 MHz, 10,000–45,000 volts; Electro-Technic Products corona treater).

Fluorescence Microscopy.

All fluorescence microscopy was performed using a Nikon Ti-E inverted fluorescence microscope with fluorescence excitation by a Lumencor Sola light-emitting diode illuminator. Images were collected by means of a Hamamatsu ORCA-Flash 4.0 CMOS camera.

5.2.4 Supplementary figures

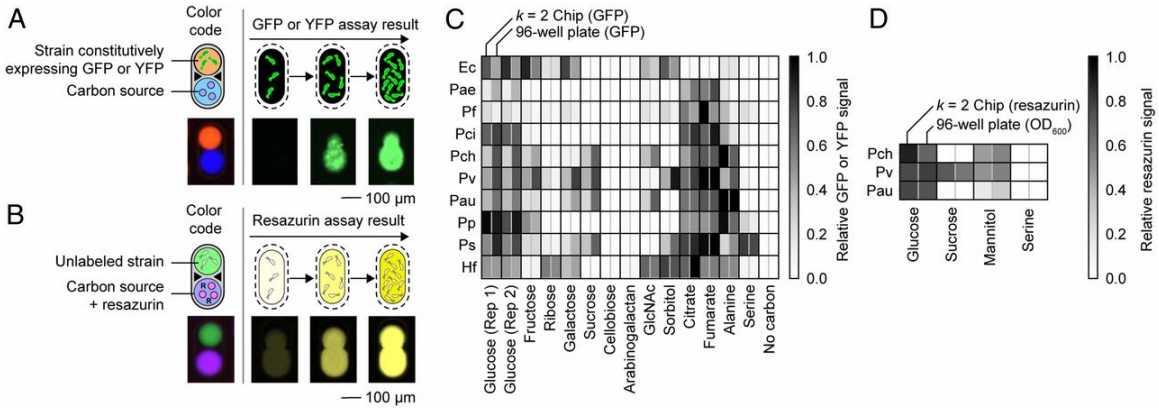

Figure 5-7: Carbon utilization profiles of labeled and unlabeled strains were measured on k = 2 Chips. (A) Droplet libraries can be made from a library of fluorescently labeled strains and a set of carbon sources. The ability of each strain to grow on each carbon source can be measured by monitoring microwells that receive one microbe-containing droplet and one carbon source-containing droplet. (B) To measure growth of unlabeled strains, the dye resazurin is added to carbon source inputs before droplet production (postmerge concentration of 40 µM). Resazurin is reduced to the fluorescent product resorufin in the presence of metabolically active cells. (C) We measured fluorescence for a panel of 10 fluorescent strains (starting OD600 = 0.02) across 15 conditions [13 carbon sources at 0.5% (wt/vol), one additional glucose replicate control, and one negative control (no carbon)] in k = 2 Chip microwells (21 °C, no shaking) as well as 200-µL cultures in 96-well plates (21 °C, 220 rpm). Heatmaps show the relative signal at 50 h, with interleaved columns corresponding to the kChip and 96-well plates (Pearson r = 0.868) (full time course is shown in [259] SI Appendix, Figure S6). (D) We measured the resazurin signal (fluorescence due to resorufin accumulation) for three strains (starting OD600 = 0.005) across four carbon source conditions in k = 2 Chip microwells (21 °C, no shaking) and compared those signals with OD600 measurements from 200-µL cultures in 96-well plates (21 °C, 220 rpm). Heatmaps show signal at 22 h (Pearson r = 0.969). In C and D, the relative signal for each row is obtained by normalizing to the maximum across all carbon sources and time points after background subtraction. Ec, *Escherichia coli*; GlcNAc, N-acetylglucosamine; Hf, *Herbaspirillum frisingense*; Pae, *Pseudomonas aeruginosa*; Pau, *Pseudomonas aurantiaca*; Pch, *Pseudomonas chlororaphis*; Pci, *Pseudomonas citronellolis*; Pf, *Pseudomonas fluorescens*; Pp, *Pseudomonas putida*; Ps, *Pseudomonas syringae*; Pv, *Pseudomonas veronii*; Rep, replicate.

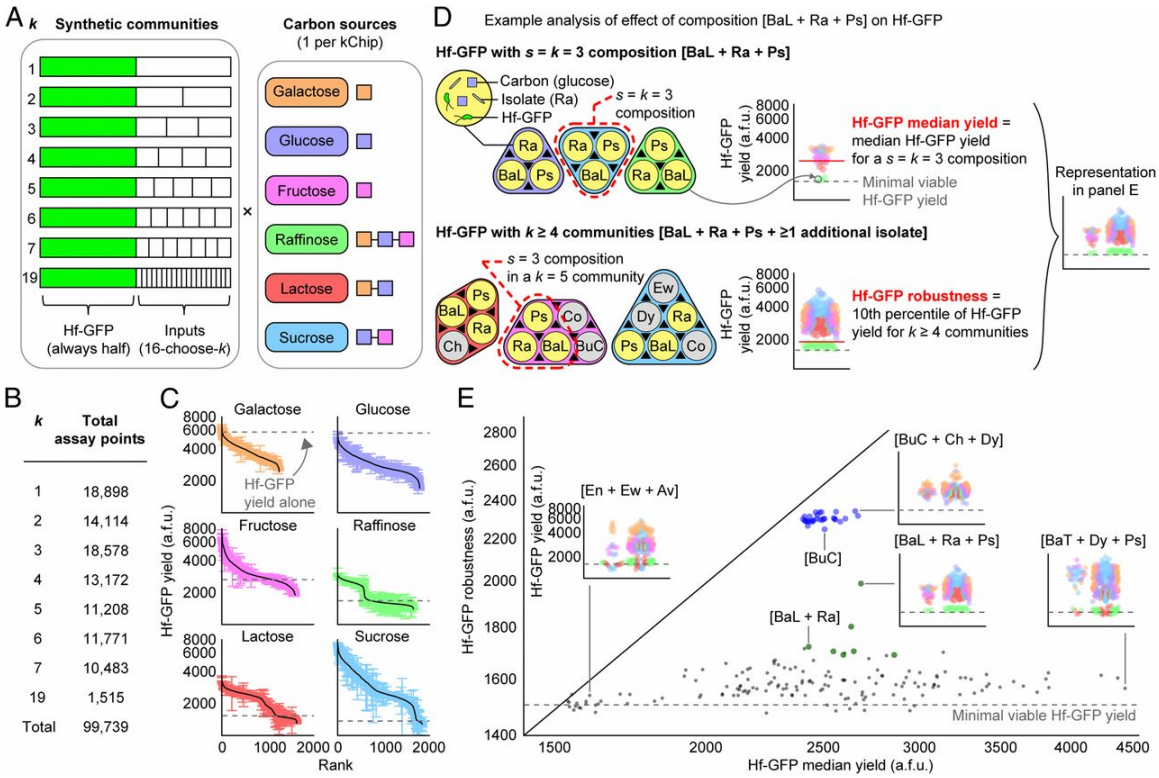

Figure 5-8: High-throughput kChip screening identifies *H. frisingense*-promoting compositions that are robust to carbon source and community context.

Figure 5-8 *(previous page)*: (A) Screen schematic to identify Hf-GFP–promoting compositions. Assays are constructed whereby Hf-GFP represents half of the starting biomass (starting Hf-GFP OD600 = 0.02) and the other half is divided evenly among one to seven or 19 soil isolate inputs (starting total isolate OD600 = 0.02 if no control droplets are present). Each of these communities is constructed in one of six media that each contain a single carbon source. Each carbon source enables a different Hf-GFP monoculture yield. Droplets containing the same carbon source are pooled and loaded onto the same kChip (six kChips in total, 21 °C, no shaking). After droplet merging, Hf-GFP yield is measured (24 h, 48 h, and 72 h) in each community/carbon source environment. (B) Total number of assay points collected for different values of k (about evenly divided among the six kChips; [259] Dataset S2). (C) Ranked Hf-GFP yield at 72 h for all constructed compositions. A median is represented when a composition is replicated more than one time (with a mean calculated in instances of two replicates), error bar = s.e.m., and dotted line = Hf-GFP yield in monoculture. (D) Effect of each s-sized composition on Hf-GFP was analyzed in two ways to identify the most facilitative and robust compositions. Here, the composition [BaL + Ra + Ps] is used as an example. (Top) First, all instances of [BaL + Ra + Ps] appearing in k = 3 microwells were identified across all carbon sources, and the median Hf-GFP yield for these was calculated ("Hf-GFP median yield"). (Bottom) Second, all instances of [BaL + Ra + Ps + one or more additional isolate] in k $\geq$ 4 microwells were identified across all carbon sources, and the 10th percentile of Hf-GFP yield for these was calculated ("Hf-GFP robustness"). The color of each data point indicates the carbon source. Gray dotted line = minimal viable Hf-GFP yield (1,500 GFP counts, or 1 SD above Hf-GFP monoculture yield in sucrose medium). (E) For compositions represented 30 or more times across all carbon sources (only k = 1, k = 2, and k = 3 compositions met this criterion), Hf-GFP median yield and Hf-GFP robustness were calculated as described in D. Dark blue points indicate a composition contains at least BuC. Dark green points indicate a composition contains at least [BaL + Ra]. The diagonal line is the x = y line. a.f.u., arbitrary fluorescence units; Av, *Averyella dalhousiensis*; Ch, *Chryseobacterium* sp.; Co, *Collimonas* sp.; Ew, *Ewingella americana*; Ps, *Pseudomonas fluorescens*.

5.3 Positive interactions are common among culturable bacteria

5.3.1 Extended materials and methods

Medium construction

The medium used in the coculture experiment was an M9 minimal medium consisting of 1X M9 salts (Teknova), 1X trace metals (Teknova), 0.1 mM calcium chloride, and 2 mM magnesium sulfate. We additionally added 0.05% w/v bovine serum albumin (BSA) to the medium to improve the retention of fluorescent dyes used in the droplet color codes 5 (see section "Input color coding") and, presumably, other small molecules 52. We previously showed that addition of BSA does not affect bacterial growth 5.

A bank of kChip-deployable carbon source growth substrates was developed from which libraries could be readily created. Carbon compounds in this bank met the following criteria: the compounds were soluble at 2% w/v in water; the solutions were emusifiable using Bio-rad QX200 cartridges; the integrity of the color code dye signals were maintained despite the presence of the carbon compound (see section "Input color coding"). A total of 33 carbon compounds were chosen that passed these criteria, representing a diversity of growth substrates including monosaccharides, oligosaccharides, polysaccharides, carboxylate ions, amino acids, sugar alcohols, and a nucleic acid (Figure 5-10).

In the coculture experiment, a total of 40 environmental conditions were used. These included the 33 chosen compounds at 0.5% w/v; five of these compounds (glucose, glycerol, pyruvate, proline, and sucrose) at 0.05% w/v; an even mix of all 33 compounds (totaling 0.5% w/v); and a no-carbon control.

Carbon source plates consisting of the 40 carbon source conditions in water (4X experiment concentration) were frozen at -20°C and thawed at the onset of each round of the experiment (see section "kChip coculture construction"). We previously demonstrated preservation of the frozen carbon substrate plates by showing tight cor-

respondence in the growth of Escherichia coli K-12 MG1655 on the freshly prepared carbon substrate plate and plates stored in -20C for 3 days and 15 days 5.

Input color coding Every unique input to the coculture screen (e.g. a bacterial culture or environmental condition) received a "color code", or unique ratio of three fluorescent dyes (standardized to a total final dye concentration of 2.5 μM), prior to generating droplets [400]. Each set of three dyes collectively labeled each specific input. The dyes included Alexa Fluor 555, Alexa Fluor 594, and Alexa Fluor 647, all of which have distinct excitation and emission spectra and dyes did not interfere with GFP. We previously showed that the color coding dyes do not affect bacterial growth 5.

Droplet preparation and pooling Each 1-nL droplet contained either a labeled/unlabeled pair or a single carbon source (Figure 3-3A). The labeled strain was projected to each unlabeled strain (1:1 ratio) just before making droplets. Droplets were produced on a Bio-Rad QX200 Droplet Generator (which generated roughly 20,000 1-nL emulsifications prepared per 20 μL input for eight inputs at a time; three minutes per 8-input cartridge). The continuous phase was a fluorocarbon oil (3M Novec 7500). For droplet making, 2% w/w fluorosurfactant (RAN Biotech 008 FluoroSurfactant) was added to stabilize droplets.

For each kChip loading (see section "kChip coculture construction"), about 5,000 droplets per each of the labeled/unlabeled inputs (22 + 2 empty controls, or a total of 24 x 5,000 = 120,000 droplets) and about 3,000 droplets per carbon source (39 + 1 empty control, or a total of 40 x 3,000 = 120,000 droplets) were pooled, generating a total of about 240,000 droplets where half contained cultures and the other half contained carbon sources. As a result, about a half of randomly generated pairwise droplet combinations on the kChip contained one labeled/unlabeled pair (premerge OD600 = 0.04 for each strain; post-merge OD600 = 0.02 for each strain, or 20 cells/droplet) and one carbon source droplet (final concentration = 0.5% w/v or 0.05% w/v).

kChip coculture construction

The following steps were performed for each kChip, with 6 kChips run in sequence per day (for a total of 24 kChip run over 4 days, with 2 kChips repeated due to poor droplet loading). As previously described, a single labeled strain was projected to all unlabeled strains, each labeled/unlabeled input and each carbon source input received a unique color code, droplets were made from all inputs, and all droplets were pooled together (Figure 3-3A, Steps 1-2). With the assistance of a loading apparatus, the droplets were loaded onto a kChip in one pipetting step (Figure 3-3A, Step 3).

A kChip loading apparatus, which was required to load droplets into microwells, consisted of a "loader top" and "loader base" [400]. The loader base held in place a piece of custom-cut glass (Brain Research Laboratories; 1.2 mm thickness) made hydrophobic via pretreatment with Aquapel. The top side of the kChip, which was not coated with parylene, spontaneously formed a seal with the loader top. Four neodymium magnet pairs were oriented such that the two acrylic pieces repelled each other. Working against this repulsive force, the loader top was lowered toward the loader base via tightening nuts until the desired standoff between the glass and kChip was attained (500-700 µm) to create a space for flow under the microwells. Via an open slot passing through the loader top and kChip, the flow space was pre-wetted with an injection of oil (3 mL to fill the entirety of the flow space) followed by the pooled droplets. Each kChip consisted of an array of microwells, each designed to randomly group a specific number of droplets. Here we used only k = 2 microwells that each group two droplets. Buoyant in the surrounding oil, the droplets were distributed around the flow space via tilting the loading apparatus. After the droplets had passed through the flow space and entered the k = 2 microwells, additional oil (no fluorosurfactant) was flushed through the device to wash away excess droplets and fluorosurfactant. The kChip, coupled to the top piece, was removed, and sealed with a transparent PCR film 35, limiting inter-microwell crosstalk 34. The edges of the film were trimmed off. A second set of four neodymium magnets were used to couple to the top piece/kChip to a microscope stage adapter.

Fluorescence imaging

All fluorescence microscopy was performed using a Nikon Ti-E inverted fluorescence microscope with fluorescence excitation by a Lumencor Sola light emitting diode illuminator (100% power setting). Images were taken across four fluorescence channels—three for the color codes and one additional channel for fluorescence-based assays. Each dye and the assay signal was detected with a different excitation wavelength generated by a collection of excitation filters: GFP by Semrock GFP-1828A (blue excitation); Alexa Fluor 555 dye by Semrock SpGold-B (green excitation); Alexa Fluor 594 dye by Semrock FF03-575/25-25 [excitation filter] + FF01-615/24-25 [emission filter] (yellow excitation); and Alexa Fluor 647 dye by Semrock LF635-B (red excitation). The emission signals corresponding to each dye channel were used to identify the contents of a given droplet within each droplet grouping prior to droplet merging (Figure 3-6B). The final channel was used post-merge and at subsequent time points to quantify the assay signal (Figure 3-6B). The camera possessed 16-bit depth (total pixel intensity range of 0-65,535 counts). Exposure times were chosen such that all assay signals and dye signals did not saturate (<20,000 counts).

Bootstrap resampling and interaction classification

To measure EB→A, the effect of unlabeled strain B on labeled strain A for a given carbon source environment, the log2 of the ratio of its yield of A in coculture (median of A+B replicates) to monoculture (median of A_{mono} replicates) was calculated (Figure 3-3D, Figure 5-12A). In instances where either of these values fell below the detection limit (DL), they were replaced with DL, which was calculated as the 90% percentile of the distribution of A+B when no carbon source was present. Several examples of this calculation are provided in Figure 5-12B. Positive values indicated facilitation, negative values indicated inhibition, and 0 indicated no detected effect.

The coordinate (EB→A, EA→B) represented both sides of an interaction. To qualitatively classify this interaction, this point was plotted on a Cartesian plane (and, as necessary, reflected to the left of the identity line y = x). Uncertainty was calculated via bootstrapping: 1000 calculations were performed for EB→A and EA→B via resampling A+B, Amono, B+A, and Bmono. The 25th and 75th per-

centiles of the resulting distributions were plotted (Figure 3-3E, Figure 5-13A). An interaction was classified as mutualism (+,+) if both sets of uncertainty bars fell within the first quadrant; as a parasitism if they both fell within the second quadrant (+,–); and as a competition if they both fell within the third quadrant (–,–). Error bars passing over quadrants indicated that a classification for at least one of the two effects did not adequately separate from no effect. If an uncertainty bar passed over the y-axis, an interaction was classified as a commensalism (+,0); if an uncertainty bar passed over the x-axis, the interaction was classified as an amensalism (–,0); if both uncertainty bars passed over the x- and y-axes, the interaction was classified as a neutralism (0,0). Examples of these pairwise classifications are shown in Figure 5-13A.

The point (EB→A, EA→B) occupied a radial continuum of possible pairwise interactions, with the magnitude m and the angle Θ of this point in polar coordinates providing a quantitative description of the interaction strength and type, respectively (Figure 3-3F). The value Θ specifically represented the relative size of the effects of two strains on each other. To determine Θ, the line y = –x was assigned to 0° (representing a balanced parasitism of equal and opposite effects). Values $-90° < \Theta < -45°$ quantified competition, with –90° indicating that two strains inhibited each other equally; –45° indicated amensalism; $-45° < \Theta < 0°$ quantified parasitism where the inhibitory effect outweighed the facilitative effect; $0° < \Theta < 45°$ quantified parasitism where the facilitative effect outweighed the inhibitory effect; 45° indicated commensalism; and $45° < \Theta < 90°$ quantified mutualism, with 90° indicating that two strains facilitated each other equally. The distribution of m and Θ are given in [400].

For a given set of interactions (e.g. all pairs among a phylogenetic group) the mean interaction [μ(EB→A), μ(EA→B)] was sometimes calculated (Figure 5-12B). The average interaction magnitude m and average interaction type Θ of this average interaction were also calculated. The variance in a set of interactions can also be calculated as an interaction diversity metric. This analytical framework was applied to the entire dataset (Figure 3-4, Figure 5-14) and to subsets of the data organized by

properties of the environments, such as biochemical classification (Figure 3-6B), or properties of strains, such or phylogenetic distances (Figure 3-6D). The distribution of the average m and average Θ for interactions grouped by coculture pair and by carbon source are given in [400].

Binning interactions by metabolic distance or monoculture growth

Using measured resource utilization profiles (monoculture growth values across all carbon sources normalized to the maximum growth per strain, per time point) (Figure 5-10), the Euclidean distance between each resource utilization profile was computed for each strain pair (Figure 5-11C) as a measure of metabolic similarity. Whereas the phylogenetic distances among pairs produced a bimodal distribution that reflected the two taxonomic orders, the metabolic distance values produced a continuous bell-shaped distribution [400].

We binned the 190 possible cocultures into 8 discrete metabolic distance bins of increasing dissimilarity and counted interaction types in each bin across all carbon sources (Figure 3-6E). (Cocultures of a labeled strains cocultured with their unlabeled counterparts, for which the metabolic distance was 0, were grouped as Bin #0.) In the center of the metabolic distance distribution (Bins #3, #4, #5, and #6), each bin spanned 0.25 Euclidean distance units. Because there were fewer data points nearer the tails of the distribution, Bins #2 and #7 each spanned 0.5 Euclidean distance units, and Bins #1 and #8 spanned 1 Euclidean distance unit. The resulting bins each had roughly equal numbers of data points.

Unlike binning by metabolic distance, binning by monoculture growth disregarded any larger interaction patterns within a given strain pair, i.e. each interaction generated among a pair across each carbon source was independently binned only by the degree to which each strain grew on the given carbon source. Normalized monoculture yields (Figure 5-10) were placed into one of seven bins (cutoffs: 0, 0.005, 0.05, 0.1, 0.2, 0.3, 0.5, 1). The first bin, [0 0.005), represented undetectable growth (within background noise) as qualified as "no growth" in analyses of obligate facilitation (Figure 3-6D). With the exception of the second bin, [0.005 0.05), which spanned a decade,

bins cutoffs were roughly based on exponential doublings.

Interaction networks and binning

Interaction networks were constructed for all pairwise interactions occurring per carbon source (examples in Figure 3-6A and all networks in [400]). The nodes, each representing a strain, were arranged concentrically by carbon utilization similarity (same order as in Figure 5-10). The size of the node corresponded linearly to the normalized monoculture growth (Figure 5-10). Edges between nodes, each representing a pairwise interaction, were colored by interaction classification (with neutralisms not shown). The thickness of the edge corresponded to interaction strength m.

5.3.2 Supplementary figures

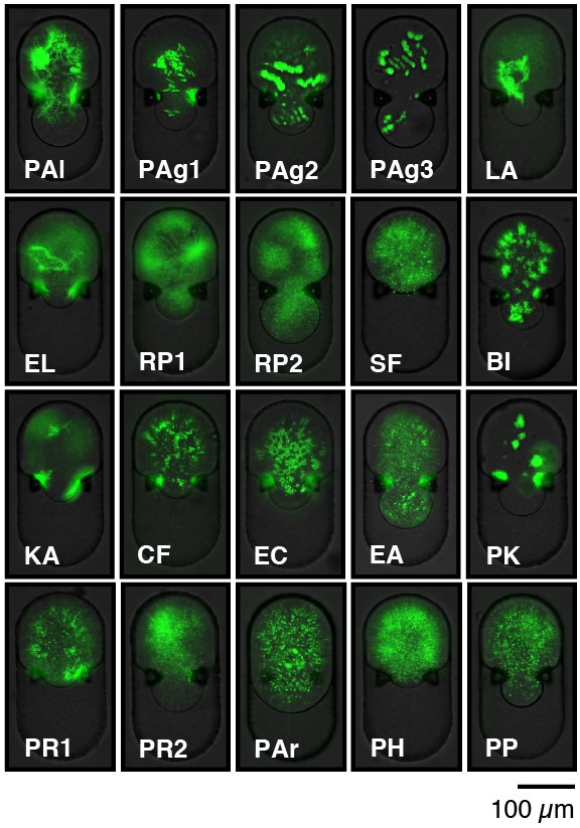

Figure 5-9: Images of labeled strains in example cocultures (with arbitrary unlabeled strain and carbon source) (10X magnification). Full strain names are provided in Table S1

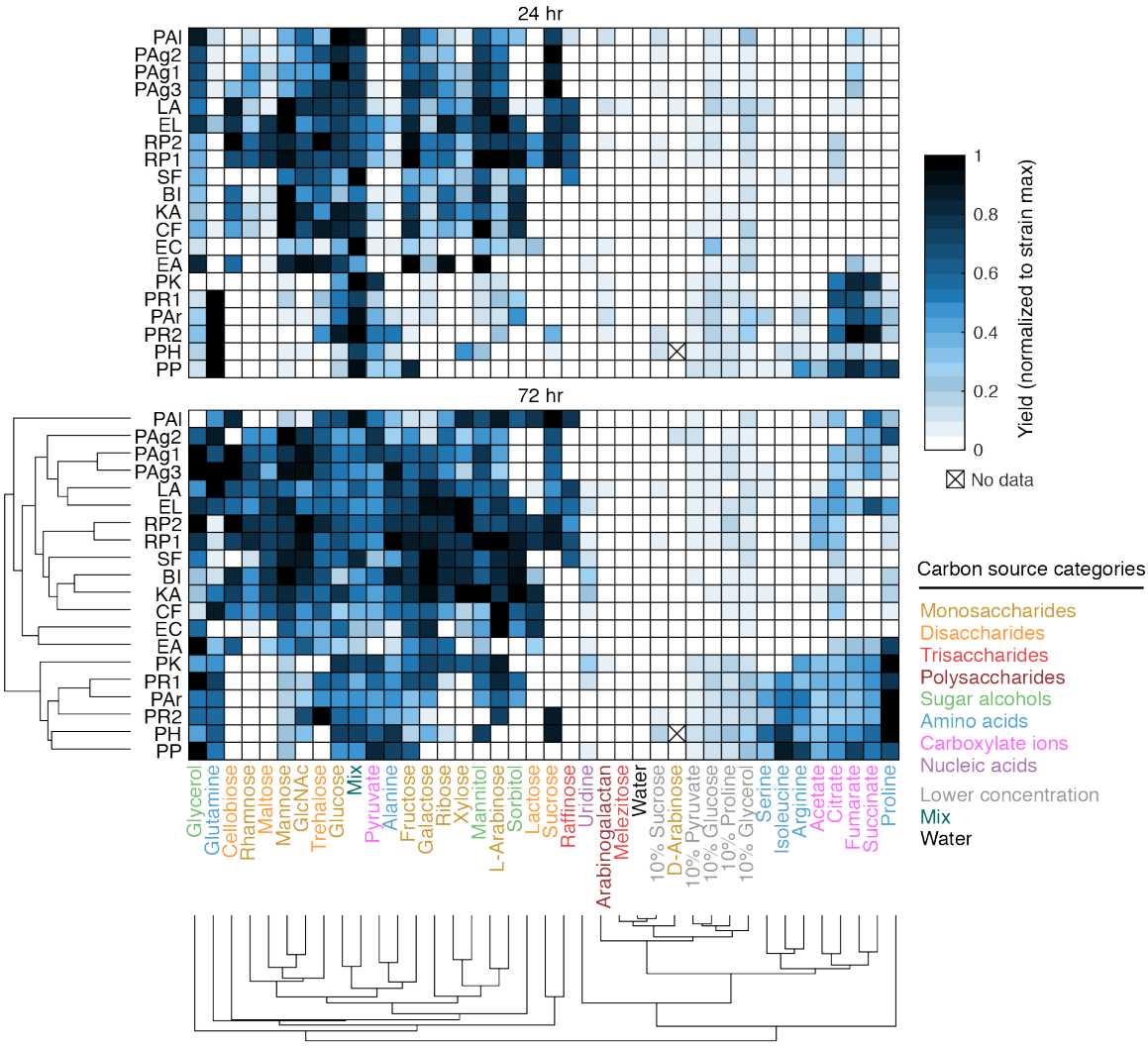

Figure 5-10: Carbon source utilization profiles of soil bacterial strains. Each strain's ability to grow on each carbon source was determined from kChip microwells containing a labeled strain in monoculture and a carbon source. Yield values (median GFP measurement across replicates) were background-subtracted (background = median yield of labeled strain with no carbon) and normalized to the maximum yield value observed (per strain per time point). Carbon sources and strains were hierarchically clustered at the 72 hr time point.

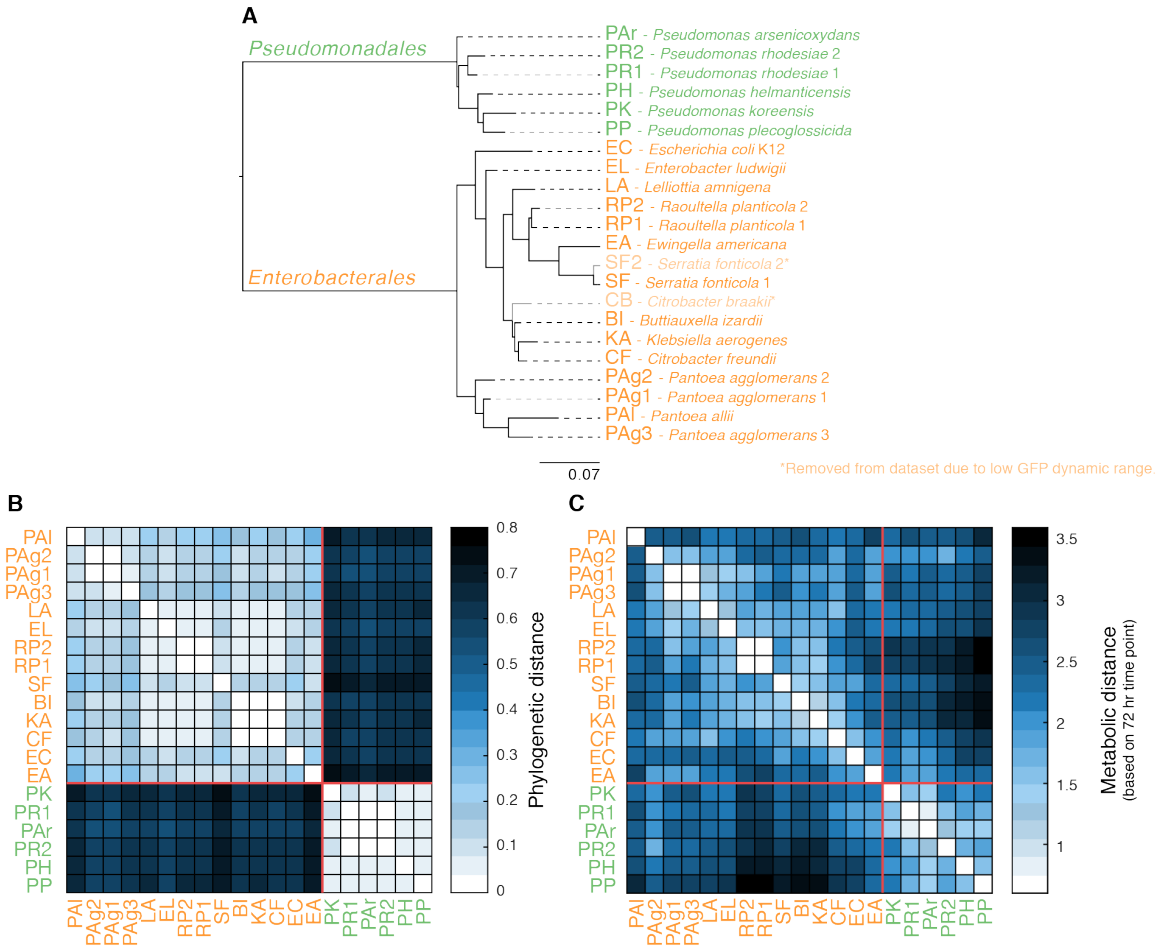

Figure 5-11: Strain pairwise comparison metrics. (A) Phylogenetic tree built with maximum likelihood estimate, utilizing alignment of full 16S sequences. Sequences of the full 16S rRNA gene (1500bp) were obtained via sanger sequencing. Sequences were trimmed to quality threshold ≥20 before merging the forward and reverse reads of each bacterial 16S. The sequence for *E. coli* K12 was obtained directly from NCBI. Sulfolobus solfataricus, a thermophilic archaeon, was used in the phylogenetic reconstruction as an outgroup species to root the tree. MUSCLE with default parameters was used to align the sequences. PhyML-SMS with default parameters was used to select GTR+G+I as the best model and to infer the tree. Units are 'Substitutions per base pair of the 16S gene. (B) Pairwise phylogenetic distances between strains were calculated directly from the phylogenetic tree with R (Library Phytools, command 'drop.tip'). (C) Metabolic distances ordered by carbon source utilization similarity. These distances were calculated as the Euclidean distance between two carbon source utilization profiles (Figure 5-10). Red line = separation of two taxonomic families. Orange font = strains of the order Enterobacterales. Green font = strains of the order Pseudomonadales.

A

$$E_{B \to A} = \text{Effect of unlabeled B on focal A} = \log_2 \frac{\max(A_{+B}, DL)}{\max(A_{mono}, DL)}$$

A_{+B} = background-subtracted median yield of focal A in coculture with B with given carbon source

A_{mono} = background-subtracted median yield of focal A in monoculture with given carbon source

DL = background-subtracted 90th percentile of A in coculture with B with no carbon source

B

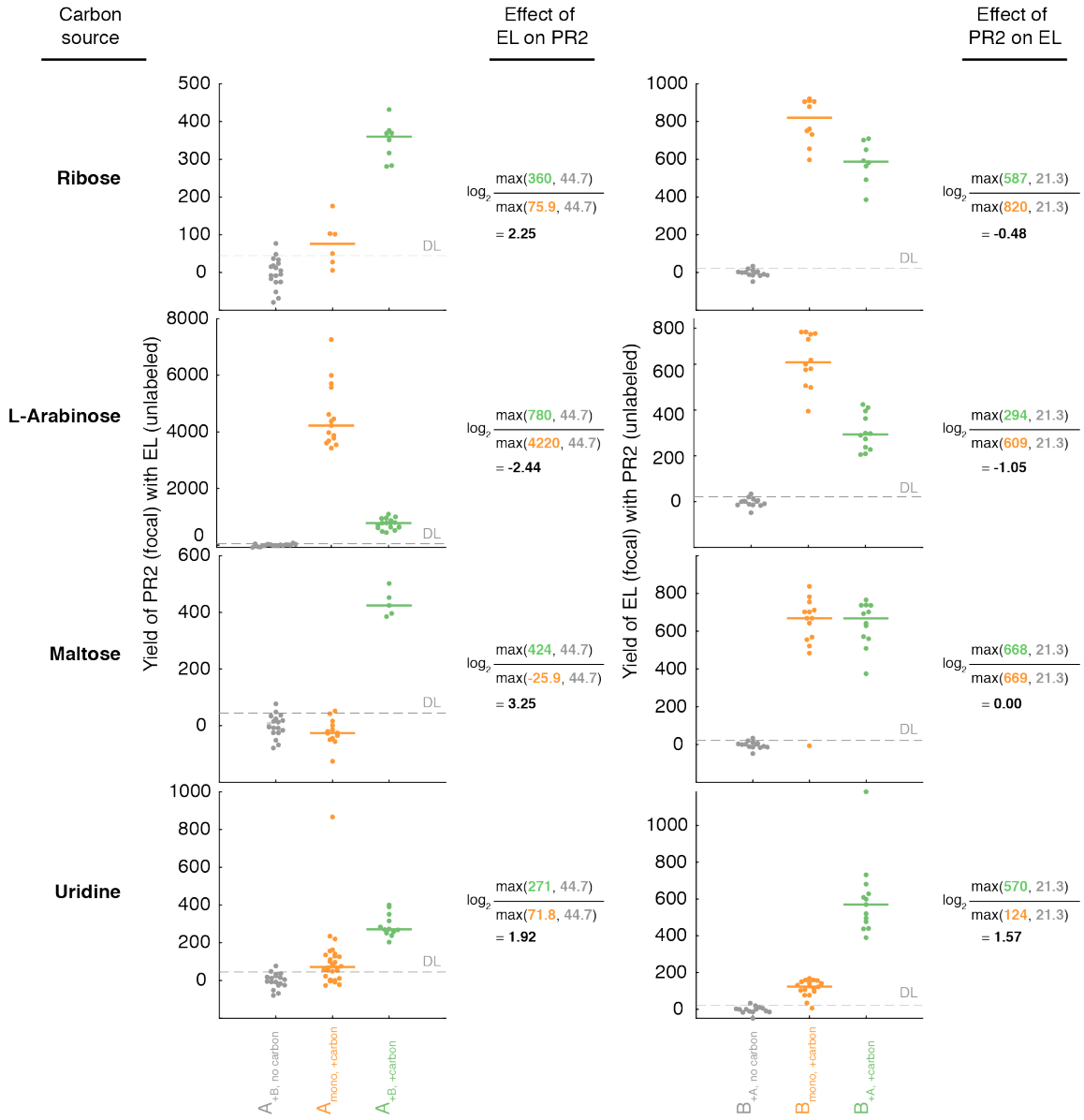

Figure 5-12: Interaction effect calculation. (A) Formula for calculating effect of unlabeled strain B on yield of labeled strain A. (B) Effect calculations for example coculture [Enterobacter ludwigii (EL) and Pseudomonas rhodesiae #2 (PR2)] on four example carbon sources.

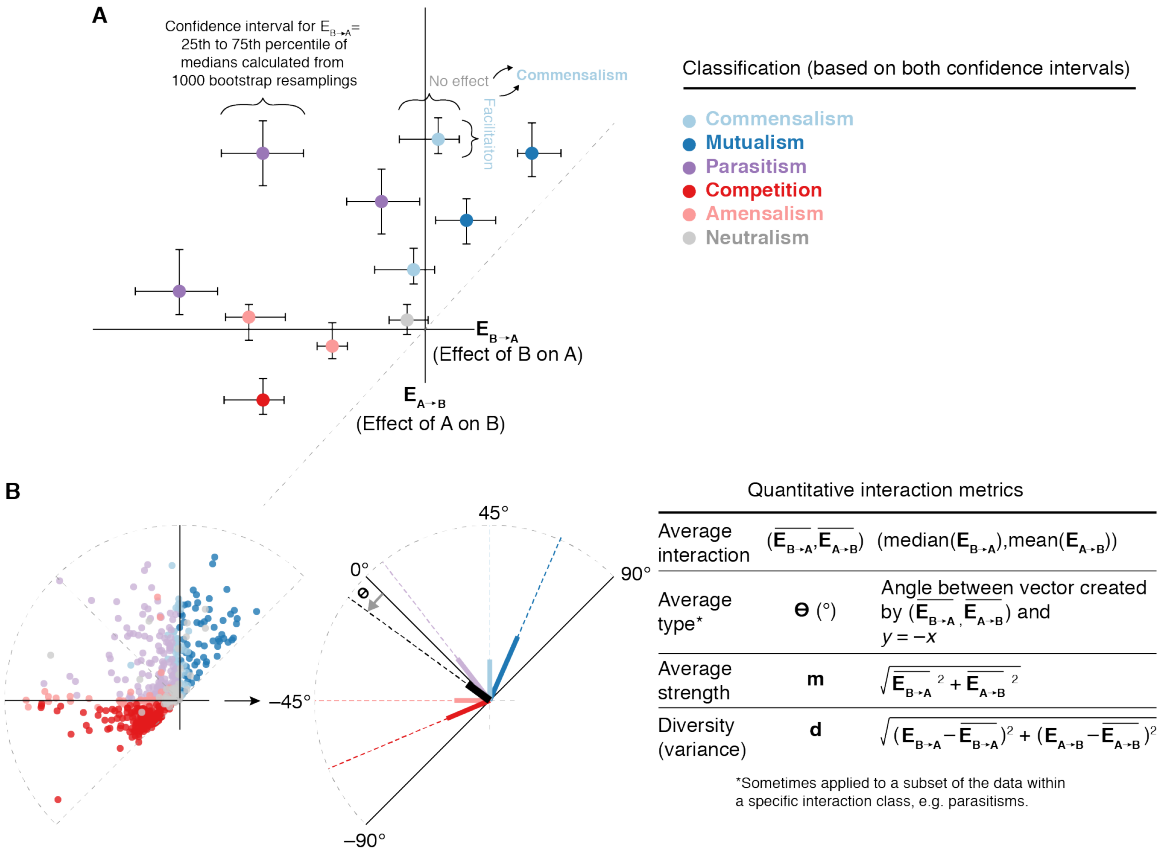

Figure 5-13: Pairwise interaction classification and quantitative metrics. (A) Method for assigning pairwise interaction classification (commensalism, mutualism, parasitism, competition, amensalism, or neutralism). Confidence intervals for all one-way effects were generated by bootstrap resampling the effect calculation (Figure 5-12). One-way interactions (facilitation, inhibition or no effect) were classified by the effect size and the fraction of the confidence interval that fell within the effect size's quadrant. Two-way interactions were classified by combining two one-way classifications. (B) Quantitative metrics used to describe pairwise interactions. These include metrics for type, strength, and diversity (with diversity only applying to a set of interactions). The type metric was also used to subset of interactions belonging to specific interaction classifications.

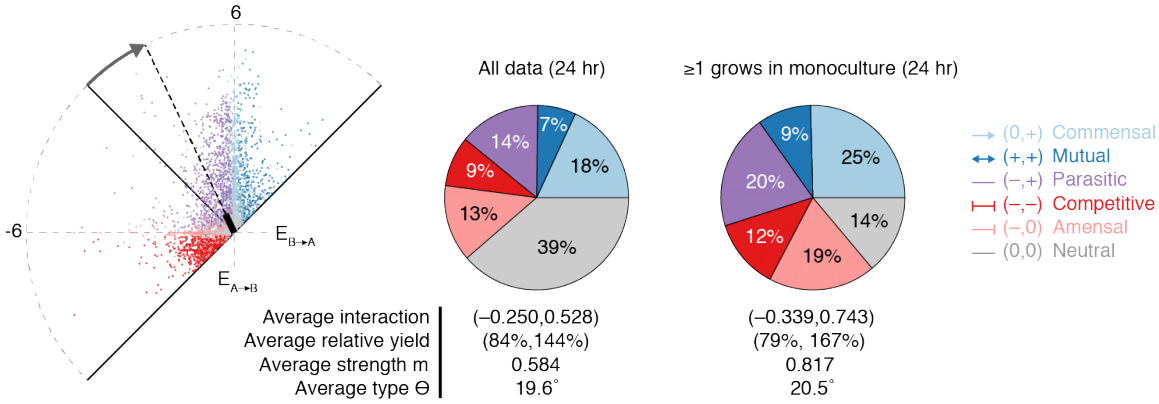

Figure 5-14: Interaction distributions for full dataset at 24 hr. (Left) All pairwise interactions at 24 hr on 33 distinct carbon sources. (Middle) Interaction classification of all data. (Right) Interaction classification excluding cases in which both strains comprising a coculture showed no detectable growth as monocultures on a given carbon source.

5.4 Role of phylogenetic classification after accounting for metabolic capabilities

Does phylogeny contain extra information to predict bacterial pairwise effects after accounting for metabolic capabilities?[1] I have shown previously that the measurements of bacterial effects ($\frac{GFP_{X|Y}}{GFP_X}$) depend on how much the affected bacteria 'X' and the effector 'Y' can grow on the medium. If X can grow well in monoculture, the presence of a partner Y will lower its yield in most cases, but if X can't grow on its own, the presence of a good growing partner will often times facilitate it. Taking into account these growth capabilities of the bacteria, we want to know if the bacterial 16S phylogenetic classification provides further understanding into when do positive interactions occur.

The twenty bacterial isolates utilized in this study fall into two orders: Pseudomonadales and Enterobacterales. The phylogenetic distances between pairs of isolates within an order (intra-order) are between 0 and .3 substitutions per site, while the distances across orders are between .55 and .75 subs/site. This bimodal distribution of phylogenetic distances allows us to simplify the metric into high (inter-order) and low (intra-order) distances. The question then becomes: does monoculture fold growth relate to bacterial effects in the same way for high and low phylogenetic distances?

In the three graphs of Figure 5-15, we are plotting the bacterial effects as a function of the effector's monoculture growth. We are only focusing on the effects that poor growing bacteria receive because these are the only effects that are quantitatively predictable based on the effector's monoculture growth. Each graph considers a different subset of the data based on phylogeny, and each graph displays a linear regression of the effects as a function of the effector's growth. The similarity in the parameters of the linear regressions show that the estimated relationship between monoculture growth and bacterial effects is not statistically different across phyloge-

[1] This question came up during the second meeting I had with my Thesis Committee, Otto Cordero and Tami Lieberman.

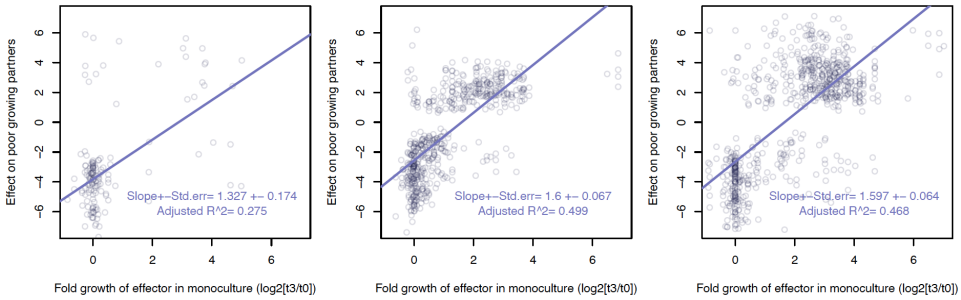

Figure 5-15: **Linear regressions show that bacterial effects vary with monoculture growth similarly across different phylogenetic distances.** Monoculture growth was calculates as the fold change in GFP fluorescence intensity after 72hrs of growth. Calculation of bacterial interaction effects is explained in Figure 5-12.

netic distances. Also, as shown in Figure 5-16, the percentage of positive bacterial effects changes significantly with the effector's fold growth for all phylogenetic bins (error bars are ± one std. dev.). The expected percentage of positive interactions appears to be lower for pairs of *Pseudomonas*, but the few 20 conditions where a high growing *Pseudomonas* was paired with a poor growing *Pseudomonas* give us low confidence on such trend.

To further test if phylogenetics adds value to the prediction of bacterial effects, we compared different multiple linear regressions with and without phylogenetic distance in the model. Here we are considering all bacterial effects instead of focusing only on

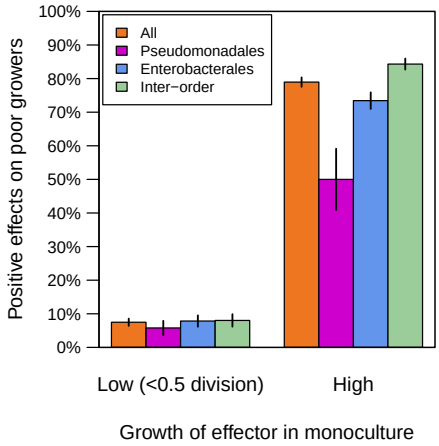

Figure 5-16: **Percentage of positive interactions on poor growers as a function of taxonomy and the effector's monoculture growth.**

Coefficients	MFGs	+PhylDist
R^2 of Model	0.30415	0.31703
Intercept	-2.3711	-2.83869
Std. error	0.05197	0.06826
MFG effector	1.1437	1.26945
Std. error	0.02257	0.0314
MFG affected	0.24394	0.21055
Std. error	0.02159	0.02893
Phyl. Distance		1.65113
Std. error		0.153481

Table 5.1: Coefficients for multiple linear regressions.

poor growing bacteria. Table 5.1 shows the R^2 and the estimated coefficients for the best two models: 'MFGs' uses the monoculture fold growth of both strains and their two-term product to predict all the significant bacterial effects, while '+PhylDist' also uses phylogeny and the two other two-term products. Although the inclusion of every variable was statistically significant ($p<.05$), the quality of the predictions (R^2) increased merely 1.2% with the addition of phylogenetic distances. In other words, if we don't break the dataset into low/medium/high monoculture growers and instead we try to predict all of the bacterial effects based on monoculture fold growths, this prediction is almost as good as a prediction that also includes phylogenetic distances. In conclusion, these results show that phylogeny adds little value after accounting for MFGs.

5.5 Pairing off: a bottom-up approach to the human gut microbiome

[2]The human gut microbiome has been implicated in a variety of health outcomes, and extensive research has aimed to understand its composition and function, primarily via metagenomic analyses. An examination of how the microbiome develops and interacts through interspecies competition and cooperation has been lacking so far. In their recent work, Venturelli et al (2018) [273][3] build a synthetic gut community and accurately predict its dynamics with a simple network of pairwise interactions.

Imagine if a single organ of the human body was proposed to be implicated in a wide variety of diseases and disorders, including Parkinson's disease, autism, depression, obesity, ulcerative colitis, cardiovascular disease, diabetes, and multiple cancer types. The gut microbiome, a microbial ecosystem inhabiting the digestive tract of humans and other animals, is frequently referred to as an additional organ since its weight (∼1 kg) is comparable to that of a small organ, and its role in physiology and disease has become evident in the recent years. In fact, the gut microbiome has been linked to all the conditions listed above as well as further pathologies [296]. While several of the proposed links between the microbiome and human diseases remain to be further investigated and might after all prove to be overblown, they nevertheless have drawn tremendous attention to the gut microbiome, and research continues at a fever pitch to link the microbiome to health outcomes. Metagenomic analyses have facilitated a "top-down" look at gut community composition [7], but "bottom-up" approaches can be instrumental for understanding how these complex bacterial communities form, interact, and affect the health of their host. In their recent study, Venturelli et al (2018) use a bottom-up approach to community assembly by performing experiments with a synthetic gut microbial community and analyze the data using a predictive dynamic computational model. They find that pairwise interactions are

[2]Abreu C, Ortiz Lopez A, Gore J. Pairing off: a bottom-up approach to the human gut microbiome. Molecular systems biology. 2018 Jun;14(6):e8425.

[3]Venturelli OS, Carr AV, Fisher G, Hsu RH, Lau R, Bowen BP, Hromada S, Northen T, Arkin AP. Deciphering microbial interactions in synthetic human gut microbiome communities. Molecular systems biology. 2018 Jun;14(6):e8157.

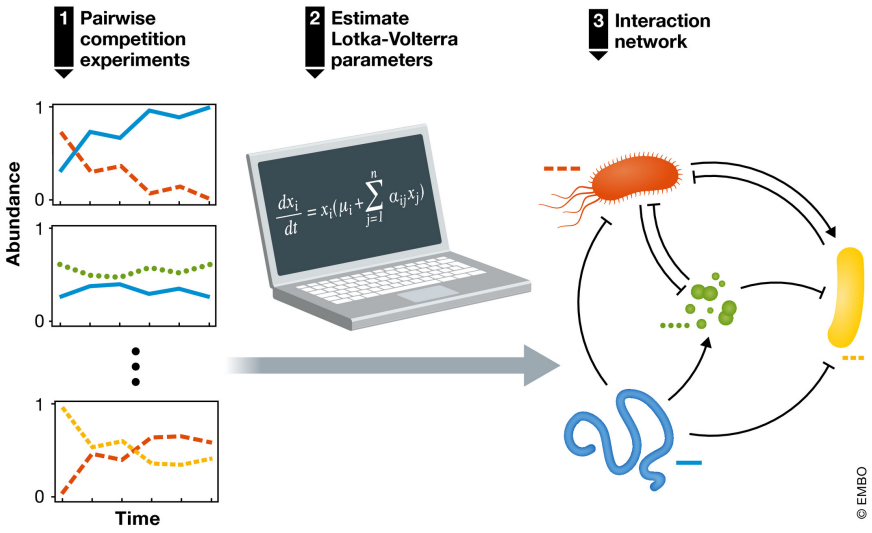

Figure 5-17: Interaction network of microbiome constructed from pairwise dynamics

sufficient to model multispecies community assembly and that certain pairwise motifs may be key to making a healthy microbiome resilient over time.

Although the authors' model community is less complex and diverse than real gut microbiome communities, its 12 species span the major phyla of human-associated intestinal bacteria. Venturelli et al (2018) performed all possible two-species combination experiments and tracked the relative abundances across time with multiplexed 16S sequencing and the total biomass with OD600 measurements. By fitting the pairwise dynamics to a generalized Lotka–Volterra model, they estimated the network of interactions 5-17 and were able to accurately predict the majority of temporal changes in 11-species (single-species dropout communities) and 12-species (full communities) communities. The predictions from pairwise experiments were significantly better compared to predictions based on single-species growth. This finding agrees with work from our group [27] and others [245] showing that pairwise outcomes are adequate for predicting multispecies community assembly and community-level metabolic rates. In contrast to some theoretical work and empirical evidence [207, 452, 266], these results suggest that higher-order interactions may play a relatively minor role in driving community assembly and that a network of pairwise interactions is an adequate description of the system.

The temporal dynamics of two-species bacterial communities were used by Ven-

turelli et al. [273] to estimate the parameters of a Lotka–Volterra model. The parameters of this model represent growth rates, intra- and interspecies interactions, and the interspecies interaction coefficients can be used to visualize the microbial ecological network.

The types of links that connect pairs of bacteria in the network influence the ecological dynamics of the communities in which they are embedded. The authors found that approximately half of the interactions are negative (56 vs. 21% positive) and note that, generally, negative interactions introduce stabilizing feedbacks to networks. Negative interactions are generally thought to arise from competition for resources or direct conflict involving toxin production, while positive interactions can occur when one species secretes a metabolite utilized by another species, a process known as "cross-feeding." However, even net-negative interactions could contain some cross-feeding, a concept similar to a casino offering free drinks to gamblers, making a negative interaction less negative. The prevalence of negative interactions in the network is in line with previous results from studies of culturable microbes [232].

Studying pairwise dynamics may be useful for predicting the survival of an invading pathogen or probiotic, or on the other hand, the effects of removing a species from a community. The authors found that several of the analyzed species, which happened to be the strongly interacting and abundant ones, had a dramatic effect on the structure of the 11- and 12-species communities. These results suggest that pairwise interactions are important in understanding how a gut community would respond to a species invasion or removal. Unhealthy gut microbiota states have been shown to be less stable to perturbations [453]. The authors noted that the prevalence of negative interactions in the network could be a stabilizing force and, more specifically, they highlighted a particular pairwise motif that appeared frequently. The positive/negative interaction motif was associated with stable coexistence in experiments and simulations and is one potential mechanism that may lead to emergent stability of a community.

The advantage of analyzing synthetic communities is that they are experimentally tractable. However, the tradeoff of this controllability is the uncertainty of how well

the findings apply to real, complex gut communities. Experiments performed in liquid cultures have less spatial structure than natural intestinal systems, and such spatial separation has been shown to promote coexistence. Additionally, serial transfer experiments use a dilution rate to model colonic transit time, but this approximation might err in its simplicity. Notwithstanding these differences, the relative abundance of species appearing in both experiments and real human guts was found to be correlated.

The work of Venturelli et al. contributes to the ongoing effort to understand the ecology of the human microbiome. The insightful results from a synthetic model community highlight the merits of applying a "bottom-up" approach to the analysis of the gut microbiome. Nontrivial patterns emerge from the 12-species network, built by simple two-species competition experiments. The use of a simple phenomenological model was also validated by its accurate predictions of community assembly and dynamics. While such models are considered to be context-dependent in comparison with mechanistic models, which incorporate more information such as metabolomics data, in this case the authors reported that fewer predictive insights were gained from metabolite measurements. Future experiments may reveal why some communities are more easily perturbed than others and whether the "healthy state" of a microbiome community can be explained by its pairwise network. Such experiments could test the predictions of Venturelli and colleagues, bridging synthetic and natural communities [454].

Bibliography

[1] Milton Wainwright. An alternative view of the early history of microbiology. *Advances in applied microbiology*, 52:333–356, 2003.

[2] Rino Rappuoli, Mariagrazia Pizza, Giuseppe Del Giudice, and Ennio De Gregorio. Vaccines, new opportunities for a new society. *Proceedings of the National Academy of Sciences*, 111(34):12288–12293, 2014.

[3] George A Zavarzin. Winogradsky and modern microbiology. *Microbiology*, 75(5):501–511, 2006.

[4] Martin Dworkin and David Gutnick. Sergei winogradsky: a founder of modern microbiology and the first microbial ecologist. *FEMS microbiology reviews*, 36(2):364–379, 2012.

[5] Marjon GJ de Vos, Marcin Zagorski, Alan McNally, and Tobias Bollenbach. Interaction networks, ecological stability, and collective antibiotic tolerance in polymicrobial infections. *Proceedings of the National Academy of Sciences*, 114(40):10666–10671, 2017.

[6] Waldan K Kwong and Nancy A Moran. Evolution of host specialization in gut microbes: the bee gut as a model. *Gut Microbes*, 6(3):214–220, 2015.

[7] Curtis Huttenhower, Dirk Gevers, Rob Knight, Sahar Abubucker, Jonathan H Badger, Asif T Chinwalla, Heather H Creasy, Ashlee M Earl, Michael G FitzGerald, Robert S Fulton, et al. Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402):207, 2012.

[8] Shinichi Sunagawa, Luis Pedro Coelho, Samuel Chaffron, Jens Roat Kultima, Karine Labadie, Guillem Salazar, Bardya Djahanschiri, Georg Zeller, Daniel R Mende, Adriana Alberti, et al. Structure and function of the global ocean microbiome. *Science*, 348(6237), 2015.

[9] Mohammad Bahram, Falk Hildebrand, Sofia K Forslund, Jennifer L Anderson, Nadejda A Soudzilovskaia, Peter M Bodegom, Johan Bengtsson-Palme, Sten Anslan, Luis Pedro Coelho, Helery Harend, Jaime Huerta-Cepas, Marnix H Medema, Mia R Maltz, Sunil Mundra, Pål Axel Olsson, Mari Pent, Sergei Põlme, Shinichi Sunagawa, Martin Ryberg, Leho Tedersoo, and Peer Bork. Structure and function of the global topsoil microbiome. *Nature*, 560(7717):233–237, August 2018.

[10] Frederic Edward Clements. *Plant succession: an analysis of the development of vegetation*. Carnegie Institution of Washington, 1916.

[11] Jonathan M Chase. Community assembly: when should history matter? *Oecologia*, 136(4):489–498, 2003.

[12] Eugene P. Odum. The strategy of ecosystem development. *Science*, 164(3877):262–270, 1969.

[13] Abdennaceur Hassen, Kaouala Belguith, Naceur Jedidi, Ameur Cherif, Mohamed Cherif, and Abdellatif Boudabous. Microbial characterization during composting of municipal solid waste. *Bioresource technology*, 80(3):217–225, 2001.

[14] Lucas J Stal, Hans van Gemerden, and Wolfgang E Krumbein. Structure and development of a benthic marine microbial mat. *FEMS Microbiology Ecology*, 1(2):111–125, 1985.

[15] Manoshi S Datta, Elzbieta Sliwerska, Jeff Gore, Martin F Polz, and Otto X Cordero. Microbial interactions lead to rapid micro-scale successions on model marine particles. *Nature communications*, 7(1):1–7, 2016.

[16] Tim N Enke, Manoshi S Datta, Julia Schwartzman, Nathan Cermak, Désirée Schmitz, Julien Barrere, Alberto Pascual-García, and Otto X Cordero. Modular assembly of polysaccharide-degrading marine microbial communities. *Current Biology*, 29(9):1528–1535, 2019.

[17] Julian Huxley et al. Evolution. the modern synthesis. *Evolution. The Modern Synthesis.*, 1942.

[18] Beth Gibson, Daniel J Wilson, Edward Feil, and Adam Eyre-Walker. The distribution of bacterial doubling times in the wild. *Proceedings of the Royal Society B*, 285(1880):20180789, 2018.

[19] Paul D Sniegowski, Philip J Gerrish, and Richard E Lenski. Evolution of high mutation rates in experimental populations of e. coli. *Nature*, 387(6634):703–705, 1997.

[20] Seyfullah Enes Kotil and Kalin Vetsigian. Emergence of evolutionarily stable communities through eco-evolutionary tunnelling. *Nature ecology & evolution*, 2(10):1644–1653, 2018.

[21] Alvaro Sanchez and Jeff Gore. Feedback between population and evolutionary dynamics determines the fate of social microbial populations. *PLoS Biology*, 11(4):e1001547, 2013.

[22] Simon A Levin. The problem of pattern and scale in ecology: the robert h. macarthur award lecture. *Ecology*, 73(6):1943–1967, 1992.

[23] Robert H. MacArthur and Edward O. Wilson. *The Theory of Island Biogeography*. Princeton University Press, 1967.

[24] Stephen P. Hubbell. *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton University Press, 2001.

[25] Robert E Ricklefs. Community diversity: relative roles of local and regional processes. *Science*, 235(4785):167–171, 1987.

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

[27] Jonathan Friedman, Logan M Higgins, and Jeff Gore. Community structure follows simple assembly rules in microbial microcosms. *Nature Ecology & Evolution*, 1(5):109, March 2017.

[28] Joshua E Goldford, Nanxi Lu, Djordje Bajić, Sylvie Estrela, Mikhail Tikhonov, Alicia Sanchez-Gorostiaga, Daniel Segrè, Pankaj Mehta, and Alvaro Sanchez. Emergent simplicity in microbial community assembly. *Science*, 361(6401):469–474, 2018.

[29] Mark Vellend. Conceptual synthesis in community ecology. *The Quarterly review of biology*, 85(2):183–206, 2010.

[30] Diana R Nemergut, Steven K Schmidt, Tadashi Fukami, Sean P O'Neill, Teresa M Bilinski, Lee F Stanish, Joseph E Knelman, John L Darcy, Ryan C Lynch, Phillip Wickey, et al. Patterns and processes of microbial community assembly. *Microbiology and molecular biology reviews: MMBR*, 77(3):342, 2013.

[31] Viviane Cordovez, Francisco Dini-Andreote, Víctor J Carrión, and Jos M Raaijmakers. Ecology and evolution of plant microbiomes. *Annual review of microbiology*, 73:69–88, 2019.

[32] Nicole M Vega and Jeff Gore. Stochastic assembly produces heterogeneous communities in the caenorhabditis elegans intestine. *PLoS Biology*, 15(3):e2000633, March 2017.

[33] Charles Darwin, Alfred Russel Wallace, Sir Charles Lyell, and Joseph Dalton Hooker. On the tendency of species to form varieties: and on the perpetuation of varieties and species by natural means of selection. In *Nineteenth-Century Evolutionism*. Linnean Society of London, 1858.

[34] Franck Courchamp, Tim Clutton-Brock, and Bryan Grenfell. Inverse density dependence and the allee effect. *Trends in ecology & evolution*, 14(10):405–410, 1999.

[35] Timothy H Keitt, Mark A Lewis, and Robert D Holt. Allee effects, invasion pinning, and species' borders. *The American Naturalist*, 157(2):203–216, 2001.

[36] RajReni B Kaul, Andrew M Kramer, Fred C Dobbs, and John M Drake. Experimental demonstration of an allee effect in microbial populations. *Biology letters*, 12(4):20160070, 2016.

[37] Georgii Frantsevich Gause. Experimental studies on the struggle for existence: I. mixed population of two species of yeast. *Journal of experimental biology*, 9(4):389–402, 1932.

[38] Sarah P Hammarlund, Jeremy M Chacón, and William R Harcombe. A shared limiting resource leads to competitive exclusion in a cross-feeding system. *Environmental microbiology*, 21(2):759–771, 2019.

[39] Rutger De Wit and Thierry Bouvier. 'everything is everywhere, but, the environment selects'; what did baas becking and beijerinck really say? *Environmental microbiology*, 8(4):755–758, 2006.

[40] Nathan JB Kraft, Peter B Adler, Oscar Godoy, Emily C James, Steve Fuller, and Jonathan M Levine. Community assembly, coexistence and the environmental filtering metaphor. *Functional ecology*, 29(5):592–599, 2015.

[41] Maureen Berg, Ben Stenuit, Joshua Ho, Andrew Wang, Caitlin Parke, Matthew Knight, Lisa Alvarez-Cohen, and Michael Shapira. Assembly of the caenorhabditis elegans gut microbiota from diverse soil microbial environments. *The ISME Journal*, 10(8):1998–2009, 2016.

[42] Silke Langenheder and Anna J Székely. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *The ISME Journal*, 5(7):1086–1094, 2011.

[43] Bin Ni, Remy Colin, Hannes Link, Robert G Endres, and Victor Sourjik. Growth-rate dependent resource investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis. *Proceedings of the National Academy of Sciences*, 117(1):595–601, 2020.

[44] Steven Smriga, Vicente I Fernandez, James G Mitchell, and Roman Stocker. Chemotaxis toward phytoplankton drives organic matter partitioning among marine bacteria. *Proceedings of the National Academy of Sciences*, 113(6):1576–1581, 2016.

[45] Weirong Liu, Jonas Cremer, Dengjin Li, Terence Hwa, and Chenli Liu. An evolutionarily stable strategy to colonize spatially extended habitats. *Nature*, 575(7784):664–668, 2019.

[46] Ali Ebrahimi, Julia Schwartzman, and Otto X Cordero. Multicellular behaviour enables cooperation in microbial cell aggregates. *Philosophical Transactions of the Royal Society B*, 374(1786):20190077, 2019.

[47] Yutaka Yawata, Otto X Cordero, Filippo Menolascina, Jan-Hendrik Hehemann, Martin F Polz, and Roman Stocker. Competition–dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations. *Proceedings of the National Academy of Sciences*, 111(15):5622–5627, 2014.

[48] Evgeni M Frenkel, Michael J McDonald, J David Van Dyken, Katya Kosheleva, Gregory I Lang, and Michael M Desai. Crowded growth leads to the spontaneous evolution of semistable coexistence in laboratory yeast populations. *Proceedings of the National Academy of Sciences*, 112(36):11306–11311, 2015.

[49] Sebastian Gude, Erçağ Pinçe, Katja M Taute, Anne-Bart Seinen, Thomas S Shimizu, and Sander J Tans. Bacterial coexistence driven by motility and spatial competition. *Nature*, 578(7796):588–592, 2020.

[50] Cole D Larsen and Anna L Hargreaves. Miniaturizing landscapes to understand species distributions. *Ecography*, 43(11):1625–1638, 2020.

[51] Jared M Diamond. The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. *Biological conservation*, 7(2):129–146, 1975.

[52] Etienne Low-Décarie, Marcus Kolber, Paige Homme, Andrea Lofano, Alex Dumbrell, Andrew Gonzalez, and Graham Bell. Community rescue in experimental metacommunities. *Proceedings of the National Academy of Sciences*, 112(46):14307–14312, 2015.

[53] Benjamin Kerr, Margaret A Riley, Marcus W Feldman, and Brendan JM Bohannan. Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature*, 418(6894):171–174, 2002.

[54] Shreyas Gokhale, Arolyn Conwill, Tanvi Ranjan, and Jeff Gore. Migration alters oscillatory dynamics and promotes survival in connected bacterial populations. *Nature communications*, 9(1):1–10, 2018.

[55] James C Stegen, Xueju Lin, Jim K Fredrickson, Xingyuan Chen, David W Kennedy, Christopher J Murray, Mark L Rockhold, and Allan Konopka. Quantifying community assembly processes and identifying features that impose them. *The ISME Journal*, 7(11):2069–2079, 2013.

[56] Linwei Wu, Daliang Ning, Bing Zhang, Yong Li, Ping Zhang, Xiaoyu Shan, Qiuting Zhang, Mathew Robert Brown, Zhenxin Li, Joy D Van Nostrand, et al. Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nature microbiology*, 4(7):1183–1195, 2019.

[57] Thea Whitman, Rachel Neurath, Adele Perera, Ilexis Chu-Jacoby, Daliang Ning, Jizhong Zhou, Peter Nico, Jennifer Pett-Ridge, and Mary Firestone. Microbial community assembly differs across minerals in a rhizosphere microcosm. *Environmental microbiology*, 20(12):4444–4460, 2018.

[58] Stilianos Louca, Saulo MS Jacques, Aliny PF Pires, Juliana S Leal, Diane S Srivastava, Laura Wegener Parfrey, Vinicius F Farjalla, and Michael Doebeli. High taxonomic variability despite stable functional structure across microbial communities. *Nature ecology & evolution*, 1(1):1–12, 2016.

[59] Marta Goberna, Alicia Montesinos-Navarro, Alfonso Valiente-Banuet, Yannick Colin, Alicia Gómez-Fernández, Santiago Donat, Jose A Navarro-Cano, and Miguel Verdú. Incorporating phylogenetic metrics to microbial co-occurrence networks based on amplicon sequences to discern community assembly processes. *Molecular ecology resources*, 19(6):1552–1564, 2019.

[60] Shao-peng Li, Pandeng Wang, Yongjian Chen, Maxwell C Wilson, Xian Yang, Chao Ma, Jianbo Lu, Xiao-yong Chen, Jianguo Wu, Wen-sheng Shu, et al. Island biogeography of soil bacteria and fungi: similar patterns, but different mechanisms. *The ISME Journal*, 14(7):1886–1896, 2020.

[61] David Tilman. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences*, 101(30):10854–10861, 2004.

[62] Ryan M Keane and Michael J Crawley. Exotic plant invasions and the enemy release hypothesis. *Trends in ecology & evolution*, 17(4):164–170, 2002.

[63] Wim H Van der Putten, John N Klironomos, and David A Wardle. Microbial ecology of biological invasions. *The ISME Journal*, 1(1):28–37, 2007.

[64] Erik S Wright and Kalin H Vetsigian. Inhibitory interactions promote frequent bistability among competing bacteria. *Nature communications*, 7(1):1–7, 2016.

[65] Harry Mickalide and Seppe Kuehn. Higher-Order interaction between species inhibits bacterial invasion of a Phototroph-Predator microbial community. *Cell Systems*, 9(6):521–533.e10, December 2019.

[66] Cynthia Portal-Celhay and Martin J Blaser. Competition and resilience between founder and introduced bacteria in the caenorhabditis elegans gut. *Infection and Immunity*, 80(3):1288–1299, March 2012.

[67] Saurabh R Gandhi, Eugene Anatoly Yurtsev, Kirill S Korolev, and Jeff Gore. Range expansions transition from pulled to pushed waves as growth becomes more cooperative in an experimental microbial population. *Proceedings of the National Academy of Sciences*, 113(25):6922–6927, 2016.

[68] Cyrus Alexander Mallon, X Le Roux, GS Van Doorn, F Dini-Andreote, F Poly, and JF Salles. The impact of failure: unsuccessful bacterial invasions steer the soil microbial community away from the invader's niche. *The ISME Journal*, 12(3):728–741, 2018.

[69] Daniel R Amor, Christoph Ratzke, and Jeff Gore. Transient invaders can induce shifts between alternative stable states of microbial communities. *Science advances*, 6(8):eaay8676, 2020.

[70] Guido Bonthond, Till Bayer, Stacy A Krueger-Hadfield, Nadja Stärck, Gaoge Wang, Masahiro Nakaoka, Sven Künzel, and Florian Weinberger. The role of host promiscuity in the invasion process of a seaweed holobiont. *The ISME Journal*, pages 1–12, 2021.

[71] Brent Nowinski and Mary Ann Moran. Niche dimensions of a marine bacterium are identified using invasion studies in coastal seawater. *Nature Microbiology*, 6(4):524–532, 2021.

[72] Jean CC Vila, Matt L Jones, Matishalin Patel, Tom Bell, and James Rosindell. Uncovering the rules of microbial community invasions. *Nature ecology & evolution*, 3(8):1162–1171, 2019.

[73] Jimmy J Qian and Erol Akçay. The balance of interaction types determines the assembly and stability of ecological communities. *Nature ecology & evolution*, 4(3):356–365, 2020.

[74] Michael Sieber, Arne Traulsen, Hinrich Schulenburg, and Angela E Douglas. On the evolutionary origins of host–microbe associations. *Proceedings of the National Academy of Sciences*, 118(9), 2021.

[75] Travis J Wiles, Brandon H Schlomann, Elena S Wall, Reina Betancourt, Raghuveer Parthasarathy, and Karen Guillemin. Swimming motility of a gut bacterial symbiont promotes resistance to intestinal expulsion and enhances inflammation. *PLoS Biology*, 18(3):e3000661, 2020.

[76] Catherine D Robinson, Helena S Klein, Kyleah D Murphy, Raghuveer Parthasarathy, Karen Guillemin, and Brendan JM Bohannan. Experimental bacterial adaptation to the zebrafish gut reveals a primary role for immigration. *PLoS Biology*, 16(12):e2006893, 2018.

[77] Edo Kussell. Evolution in microbes. *Annual review of biophysics*, 42:493–514, 2013.

[78] Olivier Tenaillon, Jeffrey E Barrick, Noah Ribeck, Daniel E Deatherage, Jeffrey L Blanchard, Aurko Dasgupta, Gabriel C Wu, Sébastien Wielgoss, Stéphane Cruveiller, Claudine Médigue, et al. Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature*, 536(7615):165–170, 2016.

[79] Sarit Avrani, Evgeni Bolotin, Sophia Katz, and Ruth Hershberg. Rapid genetic adaptation during the first four months of survival under resource exhaustion. *Molecular biology and evolution*, 34(7):1758–1769, 2017.

[80] Avihu H Yona, Eric J Alm, and Jeff Gore. Random sequences rapidly evolve into de novo promoters. *Nature communications*, 9(1):1–10, 2018.

[81] Otmane Lamrabet, Jacqueline Plumbridge, Mikaël Martin, Richard E Lenski, Dominique Schneider, and Thomas Hindré. Plasticity of promoter-core sequences allows bacteria to compensate for the loss of a key global regulatory gene. *Molecular biology and evolution*, 36(6):1121–1133, 2019.

[82] Abigail Trejo-Hernández, Andrés Andrade-Domínguez, Magdalena Hernández, and Sergio Encarnacion. Interspecies competition triggers virulence and mutability in candida albicans–pseudomonas aeruginosa mixed biofilms. *The ISME Journal*, 8(10):1974–1988, 2014.

[83] Antoine Frenoy and Sebastian Bonhoeffer. Death and population dynamics affect mutation rate estimates and evolvability under stress in bacteria. *PLoS Biology*, 16(5):e2005056, 2018.

[84] Gregory I Lang, Daniel P Rice, Mark J Hickman, Erica Sodergren, George M Weinstock, David Botstein, and Michael M Desai. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*, 500(7464):571–574, 2013.

[85] Joseph J Vitti, Sharon R Grossman, and Pardis C Sabeti. Detecting natural selection in genomic data. *Annual review of genetics*, 47:97–120, 2013.

[86] Jason M Peters, Alexandre Colavin, Handuo Shi, Tomasz L Czarny, Matthew H Larson, Spencer Wong, John S Hawkins, Candy HS Lu, Byoung-Mo Koo, Elizabeth Marta, et al. A comprehensive, crispr-based functional analysis of essential genes in bacteria. *Cell*, 165(6):1493–1506, 2016.

[87] Morgan N Price, Kelly M Wetmore, R Jordan Waters, Mark Callaghan, Jayashree Ray, Hualan Liu, Jennifer V Kuehl, Ryan A Melnyk, Jacob S Lamson, Yumi Suh, et al. Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature*, 557(7706):503–509, 2018.

[88] Philip Arevalo, David VanInsberghe, Joseph Elsherbini, Jeff Gore, and Martin F Polz. A reverse ecology approach based on a biological definition of microbial populations. *Cell*, 178(4):820–834, 2019.

[89] Eugene V Koonin, Kira S Makarova, and L Aravind. Horizontal gene transfer in prokaryotes: quantification and classification. *Annual Reviews in Microbiology*, 55(1):709–742, 2001.

[90] Shannon M Soucy, Jinling Huang, and Johann Peter Gogarten. Horizontal gene transfer: building the web of life. *Nature Reviews Genetics*, 16(8):472–482, 2015.

[91] Daniel JP Engelmoer, Ian Donaldson, and Daniel E Rozen. Conservative sex and the benefits of transformation in streptococcus pneumoniae. *PLoS Pathogens*, 9(11):e1003758, 2013.

[92] Nelson Frazão, Ana Sousa, Michael Lässig, and Isabel Gordo. Horizontal gene transfer overrides mutation in escherichia coli colonizing the mammalian gut. *Proceedings of the National Academy of Sciences*, 116(36):17906–17915, 2019.

[93] John T Sullivan and Clive W Ronson. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-trna gene. *Proceedings of the National Academy of Sciences*, 95(9):5145–5149, 1998.

[94] María J Lorite, María J Estrella, Francisco J Escaray, Analía Sannazzaro, Isabel M Videira e Castro, Jorge Monza, Juan Sanjuán, and Milagros León-Barrios. The rhizobia-lotus symbioses: deeply specific and widely diverse. *Frontiers in microbiology*, 9:2055, 2018.

[95] Dana E Hunt, Dirk Gevers, Nisha M Vahora, and Martin F Polz. Conservation of the chitin utilization pathway in the vibrionaceae. *Applied and environmental microbiology*, 74(1):44, 2008.

[96] Chris S Smillie, Mark B Smith, Jonathan Friedman, Otto X Cordero, Lawrence A David, and Eric J Alm. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature*, 480(7376):241–244, 2011.

[97] Rafael Peña-Miller, Rogelio Rodríguez-González, R Craig MacLean, and Alvaro San Millan. Evaluating the effect of horizontal transmission on the stability of plasmids under different selection regimes. *Mobile genetic elements*, 5(3):29–33, 2015.

[98] Abigail A Salyers, Anamika Gupta, and Yanping Wang. Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends in microbiology*, 12(9):412–416, 2004.

[99] Christian S Riesenfeld, Robert M Goodman, and Jo Handelsman. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environmental microbiology*, 6(9):981–989, 2004.

[100] Mathieu Groussin, Mathilde Poyet, Ainara Sistiaga, Sean M Kearney, Katya Moniz, Mary Noel, Jeff Hooker, Sean M Gibbons, Laure Segurel, Alain Froment, et al. Elevated rates of horizontal gene transfer in the industrialized human microbiome. *Cell*, 184(8):2053–2067, 2021.

[101] Rodney A Welch, V Burland, GIII Plunkett, P Redford, P Roesch, D Rasko, EL Buckles, S-R Liou, A Boutin, Jeremiah Hackett, et al. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic escherichia coli. *Proceedings of the National Academy of Sciences*, 99(26):17020–17024, 2002.

[102] Carl R Woese and George E Fox. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences*, 74(11):5088–5090, 1977.

[103] Christophe Fraser, Eric J Alm, Martin F Polz, Brian G Spratt, and William P Hanage. The bacterial species challenge: making sense of genetic and ecological diversity. *Science*, 323(5915):741–746, 2009.

[104] B Jesse Shapiro, Jonathan Friedman, Otto X Cordero, Sarah P Preheim, Sonia C Timberlake, Gitta Szabó, Martin F Polz, and Eric J Alm. Population genomics of early events in the ecological differentiation of bacteria. *Science*, 336(6077):48–51, 2012.

[105] B Jesse Shapiro and Martin F Polz. Ordering microbial diversity into ecologically and genetically cohesive units. *Trends in microbiology*, 22(5):235–247, 2014.

[106] Sydney Brenner. The genetics of caenorhabditis elegans. *Genetics*, 77(1):71–94, 1974.

[107] Catharine H Rankin. From gene to identified neuron to behaviour in caenorhabditis elegans. *Nature Reviews Genetics*, 3(8):622–630, 2002.

[108] Dennis H Kim, Rhonda Feinbaum, Geneviève Alloing, Fred E Emerson, Danielle A Garsin, Hideki Inoue, Miho Tanaka-Hino, Naoki Hisamoto, Kunihiro Matsumoto, Man-Wah Tan, et al. A conserved p38 map kinase pathway in caenorhabditis elegans innate immunity. *Science*, 297(5581):623–626, 2002.

[109] Danielle A Garsin, Jacinto M Villanueva, Jakob Begun, Dennis H Kim, Costi D Sifri, Stephen B Calderwood, Gary Ruvkun, and Frederick M Ausubel. Long-lived c. elegans daf-2 mutants are resistant to bacterial pathogens. *Science*, 300(5627):1921–1921, 2003.

[110] Gustavo V Mallo, C Léopold Kurz, Carole Couillault, Nathalie Pujol, Samuel Granjeaud, Yuji Kohara, and Jonathan J Ewbank. Inducible antibacterial defense system in c. elegans. *Current Biology*, 12(14):1209–1214, 2002.

[111] Maureen Berg, David Monnin, Juhyun Cho, Lydia Nelson, Alex Crits-Christoph, and Michael Shapira. TGFβ/BMP immune signaling affects abundance and function of c. elegans gut commensals. *Nature Communications*, 10(1), 2019.

[112] Cynthia Portal-Celhay, Ellen R Bradley, and Martin J Blaser. Control of intestinal bacterial proliferation in regulation of lifespan in caenorhabditis elegans. *BMC microbiology*, 12(1):1–17, 2012.

[113] Megan Taylor and NM Vega. Host immunity alters community ecology and stability of the microbiome in a caenorhabditis elegans model. *Msystems*, 6(2), 2021.

[114] Mark G Sterken, L Basten Snoek, Jan E Kammenga, and Erik C Andersen. The laboratory domestication of caenorhabditis elegans. *Trends in Genetics*, 31(5):224–231, 2015.

[115] Arnaud Labrousse, Sophie Chauvet, Carole Couillault, C Léopold Kurz, and Jonathan J Ewbank. Caenorhabditis elegans is a model host for salmonella typhimurium. *Current Biology*, 10(23):1543–1545, 2000.

[116] Bing Han, Priya Sivaramakrishnan, Chih-Chun J Lin, Isaiah AA Neve, Jingquan He, Li Wei Rachel Tay, Jessica N Sowa, Antons Sizovs, Guangwei Du, Jin Wang, et al. Microbial genetic composition tunes host longevity. *Cell*, 169(7):1249–1262, 2017.

[117] Stephen Jay Gould and Richard C Lewontin. The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the royal society of London. Series B. Biological Sciences*, 205(1161):581–598, 1979.

[118] Sergi Valverde, Jordi Piñero, Bernat Corominas-Murtra, Jose Montoya, Lucas Joppa, and Ricard Solé. The architecture of mutualistic networks as an evolutionary spandrel. *Nature ecology & evolution*, 2(1):94–99, 2018.

[119] Kevin R Foster, Jonas Schluter, Katharine Z Coyte, and Seth Rakoff-Nahoum. The evolution of the host microbiome as an ecosystem on a leash. *Nature*, 548(7665):43–51, 2017.

[120] Motoo Kimura et al. Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626, 1968.

[121] Rees Kassen and Paul B Rainey. The ecology and genetics of microbial diversity. *Annual Reviews in Microbiology*, 58:207–231, 2004.

[122] Les Dethlefsen, Paul B Eckburg, Elisabeth M Bik, and David A Relman. Assembly of the human intestinal microbiota. *Trends in ecology & evolution*, 21(9):517–523, 2006.

[123] Jamie M Kneitel and Jonathan M Chase. Trade-offs in community ecology: linking spatial scales and species coexistence. *Ecology letters*, 7(1):69–80, 2004.

[124] Stefanie Widder, Rosalind J Allen, Thomas Pfeiffer, Thomas P Curtis, Carsten Wiuf, William T Sloan, Otto X Cordero, Sam P Brown, Babak Momeni, Wenying Shou, Helen Kettle, Harry J Flint, Andreas F Haas, Béatrice Laroche, Jan-Ulrich Kreft, Paul B Rainey, Shiri Freilich, Stefan Schuster, Kim Milferstedt, Jan R van der Meer, Tobias Groβkopf, Jef Huisman, Andrew Free, Cristian Picioreanu, Christopher Quince, Isaac Klapper, Simon Labarthe, Barth F Smets, Harris Wang, Isaac Newton Institute Fellows, and Orkun S Soyer. Challenges in microbial ecology: building predictive understanding of community function and dynamics. *The ISME Journal*, 10(11):2557–2568, November 2016.

[125] Jennifer B Hughes Martiny, Brendan JM Bohannan, James H Brown, Robert K Colwell, Jed A Fuhrman, Jessica L Green, M Claire Horner-Devine, Matthew Kane, Jennifer Adams Krumins, Cheryl R Kuske, et al. Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, 4(2):102–112, 2006.

[126] Alban Ramette and James M Tiedje. Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. *Microbial ecology*, 53(2):197–207, 2007.

[127] China A Hanson, Jed A Fuhrman, M Claire Horner-Devine, and Jennifer BH Martiny. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology*, 10(7):497–506, 2012.

[128] Nicholas J Gotelli and Gary R Graves. *Null models in ecology*. Smithsonian Institution Press, 1996.

[129] Nicholas J Gotelli. Null model analysis of species co-occurrence patterns. *Ecology*, 81(9):2606–2621, 2000.

[130] Nicholas J Gotelli and Robert K Colwell. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology letters*, 4(4):379–391, 2001.

[131] Nicholas J Gotelli and Declan J McCabe. Species co-occurrence: a meta-analysis of jm diamond's assembly rules model. *Ecology*, 83(8):2091–2096, 2002.

[132] M Claire Horner-Devine, Jessica M Silver, Mathew A Leibold, Brendan JM Bohannan, Robert K Colwell, Jed A Fuhrman, Jessica L Green, Cheryl R Kuske, Jennifer BH Martiny, Gerard Muyzer, et al. A comparison of taxon co-occurrence patterns for macro-and microorganisms. *Ecology*, 88(6):1345–1353, 2007.

[133] Lianwei Li and Zhanshan Sam Ma. Testing the neutral theory of biodiversity with human microbiome datasets. *Scientific reports*, 6(1):1–10, 2016.

[134] Afrah Shafquat, Regina Joice, Sheri L Simmons, and Curtis Huttenhower. Functional and phylogenetic assembly of microbial communities in the human microbiome. *Trends in microbiology*, 22(5):261–266, 2014.

[135] Ariel Amir and Nathalie Q Balaban. Learning from noise: how observing stochasticity may aid microbiology. *Trends in microbiology*, 26(4):376–385, 2018.

[136] Monika Scholz, Aaron R Dinner, Erel Levine, and David Biron. Stochastic feeding dynamics arise from the need for information and energy. *Proceedings of the National Academy of Sciences*, 114(35):9261–9266, 2017.

[137] Luisa De Sordi, Varun Khanna, and Laurent Debarbieux. The gut microbiota facilitates drifts in the genetic diversity and infectivity of bacterial viruses. *Cell host & microbe*, 22(6):801–808, 2017.

[138] Stilianos Louca, Martin F Polz, Florent Mazel, Michaeline BN Albright, Julie A Huber, Mary I O'Connor, Martin Ackermann, Aria S Hahn, Diane S Srivastava, Sean A Crowe, et al. Function and functional redundancy in microbial systems. *Nature ecology & evolution*, 2(6):936–943, 2018.

[139] Laura A Herndon, Peter J Schmeissner, Justyna M Dudaronek, Paula A Brown, Kristin M Listner, Yuko Sakano, Marie C Paupard, David H Hall, and Monica Driscoll. Stochastic and genetic factors influence tissue-specific decline in ageing c. elegans. *Nature*, 419(6909):808–814, 2002.

[140] Deqing Wu, Shane L Rea, Anatoli I Yashin, and Thomas E Johnson. Visualizing hidden heterogeneity in isogenic populations of c. elegans. *Experimental gerontology*, 41(3):261–270, 2006.

[141] Mikihiro Hashimoto, Takashi Nozoe, Hidenori Nakaoka, Reiko Okura, Sayo Akiyoshi, Kunihiko Kaneko, Edo Kussell, and Yuichi Wakamoto. Noise-driven growth rate gain in clonal cellular populations. *Proceedings of the National Academy of Sciences*, 113(12):3251–3256, 2016.

[142] Jan-Willem Veening, Wiep Klaas Smits, and Oscar P Kuipers. Bistability, epigenetics, and bet-hedging in bacteria. *Annual Reviews in Microbiology*, 62:193–210, 2008.

[143] Martin Ackermann. A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Reviews Microbiology*, 13(8):497–508, 2015.

[144] Akshit Goyal and Sergei Maslov. Diversity, stability, and reproducibility in stochastically assembled microbial ecosystems. *Physical Review Letters*, 120(15):158102, 2018.

[145] Gabriel Birzu, Sakib Matin, Oskar Hallatschek, and Kirill S Korolev. Genetic drift in range expansions is very sensitive to density dependence in dispersal and growth. *Ecology Letters*, 22(11):1817–1827, 2019.

[146] Michael T Pearce, Atish Agarwala, and Daniel S Fisher. Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos. *Proceedings of the National Academy of Sciences*, 117(25):14572–14583, 2020.

[147] Lana Descheemaeker and Sophie De Buyl. Stochastic logistic models reproduce experimental time series of microbial communities. *Elife*, 9:e55650, 2020.

[148] John H Vandermeer. The competitive structure of communities: an experimental approach with protozoa. *Ecology*, 50(3):362–371, 1969.

[149] Daniel Rodríguez Amor and Martina Dal Bello. Bottom-Up approaches to synthetic cooperation in microbial communities. *Life*, 9(1), February 2019.

[150] Wenzheng Liu, Samuel Jacquiod, Asker Brejnrod, Jakob Russel, Mette Burmølle, and Søren J Sørensen. Deciphering links between bacterial interactions and spatial organization in multispecies biofilms. *The ISME Journal*, 13(12):3054–3066, 2019.

[151] Georgij Francevic Gauze. *The Struggle for Existence*. Hafner, 1934.

[152] Jeannine Cavender-Bares, Kenneth H Kozak, Paul VA Fine, and Steven W Kembel. The merging of community ecology and phylogenetic biology. *Ecology letters*, 12(7):693–715, 2009.

[153] Jerry A Coyne. Genetics and speciation. *Nature*, 355(6360):511–515, 1992.

[154] John A Wiens. Spatial scaling in ecology. *Functional ecology*, 3(4):385–397, 1989.

[155] G Evelyn Hutchinson. The paradox of the plankton. *The American Naturalist*, 95(882):137–145, 1961.

[156] Carl R Woese, Otto Kandler, and Mark L Wheelis. Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya. *Proceedings of the National Academy of Sciences*, 87(12):4576–4579, 1990.

[157] Donovan H Parks, Maria Chuvochina, David W Waite, Christian Rinke, Adam Skarshewski, Pierre-Alain Chaumeil, and Philip Hugenholtz. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature biotechnology*, 36(10):996–1004, 2018.

[158] Akshit Goyal, Leonora S Bittleston, Gabriel E Leventhal, Lu Lu, and Otto X Cordero. Interactions between strains govern the eco-evolutionary dynamics of microbial communities. *BioRxiv*, 2021.

[159] Kalin Vetsigian, Rishi Jajoo, and Roy Kishony. Structure and evolution of streptomyces interaction networks in soil and in silico. *PLoS Biology*, 9(10):e1001184, 2011.

[160] Otto X Cordero, Hans Wildschutte, Benjamin Kirkup, Sarah Proehl, Lynn Ngo, Fatima Hussain, Frederique Le Roux, Tracy Mincer, and Martin F Polz. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science*, 337(6099):1228–1231, 2012.

[161] Jakob Russel, Henriette L Røder, Jonas S Madsen, Mette Burmølle, and Søren J Sørensen. Antagonism correlates with metabolic similarity in diverse bacteria. *Proceedings of the National Academy of Sciences*, 114(40):10684–10688, 2017.

[162] Matti Gralka, Rachel Szabo, Roman Stocker, and Otto X Cordero. Trophic interactions and the drivers of microbial community assembly. *Current Biology*, 30(19):R1176–R1188, 2020.

[163] L Charles Birch. The meanings of competition. *The American Naturalist*, 91(856):5–18, 1957.

[164] Robert M Pringle, Tyler R Kartzinel, Todd M Palmer, Timothy J Thurman, Kena Fox-Dobbs, Charles CY Xu, Matthew C Hutchinson, Tyler C Coverdale, Joshua H Daskin, Dominic A Evangelista, et al. Predator-induced collapse of niche structure and species coexistence. *Nature*, 570(7759):58–64, 2019.

[165] Canan Karakoç, Adam Thomas Clark, and Antonis Chatzinotas. Diversity and coexistence are influenced by time-dependent species interactions in a predator–prey system. *Ecology letters*, 23(6):983–993, 2020.

[166] Aleksandar D Kostic, Michael R Howitt, and Wendy S Garrett. Exploring host-microbiota interactions in animal models and humans. *Genes & development.*, 27(7):701–718, April 2013.

[167] Peter D Newell and Angela E Douglas. Interspecies interactions determine the impact of the gut microbiota on nutrient allocation in drosophila melanogaster. *Appl. Environ. Microbiol.*, 80(2):788–796, January 2014.

[168] Alison L Gould, Vivian Zhang, Lisa Lamberti, Eric W Jones, Benjamin Obadia, Nikolaos Korasidis, Alex Gavryushkin, Jean M Carlson, Niko Beerenwinkel, and William B Ludington. Microbiome interactions shape host fitness. *Proceedings of the National Academy of Sciences*, 115(51):E11951–E11960, December 2018.

[169] Samuel Slowinski, Isabella Ramirez, Vivek Narayan, Medha Somayaji, Maya Para, Sarah Pi, Niharika Jadeja, Siavash Karimzadegan, Barbara Pees, and Michael Shapira. Interactions with a complex microbiota mediate a trade-off between the host development rate and heat stress resistance. *Microorganisms*, 8(11):1781, 2020.

[170] Kyaw Aung, Yanjuan Jiang, and Sheng Yang He. The role of water in plant–microbe interactions. *The Plant Journal*, 93(4):771–780, 2018.

[171] Leo Breiman et al. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, 16(3):199–231, 2001.

[172] Robert M May. Will a large complex system be stable? *Nature*, 238(5364):413–414, 1972.

[173] Peter Chesson. Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, 31(1):343–366, 2000.

[174] Robert MacArthur. Species packing and competitive equilibrium for many species. *Theoretical population biology*, 1(1):1–11, 1970.

[175] Hirotsugu Matsuda, Naofumi Ogita, Akira Sasaki, and Kazunori Satō. Statistical mechanics of population: the lattice lotka-volterra model. *Progress of theoretical Physics*, 88(6):1035–1049, 1992.

[176] Andrew D Letten and Daniel B Stouffer. The mechanistic basis for higher-order interactions and non-additivity in competitive communities. *Ecology Letters*, 22(3):423–436, 2019.

[177] Pragya Singh and Gaurav Baruah. Higher order interactions and species coexistence. *Theoretical Ecology*, 14(1):71–83, 2021.

[178] Clare I Abreu, Jonathan Friedman, Vilhelm L Andersen Woltz, and Jeff Gore. Mortality causes universal changes in microbial community composition. *Nat. Commun.*, 10(1):2120, May 2019.

[179] Katharine Z Coyte, Chitong Rao, Seth Rakoff-Nahoum, and Kevin R Foster. Ecological rules for the assembly of microbiome communities. *PLoS Biology*, 19(2):e3001116, 2021.

[180] Simon P Hart, Robert P Freckleton, and Jonathan M Levine. How to quantify competitive ability. *Journal of Ecology*, 106(5):1902–1909, 2018.

[181] Adam Thomas Clark, Lindsay Ann Turnbull, Andrew Tredennick, Eric Allan, W Stanley Harpole, Margaret M Mayfield, Santiago Soliveres, Kathryn Barry, Nico Eisenhauer, Hans de Kroon, et al. Predicting species abundances in a grassland biodiversity experiment: Trade-offs between model complexity and generality. *Journal of ecology*, 108(2):774–787, 2020.

[182] Robert MacArthur. Fluctuations of animal populations and a measure of community stability. *ecology*, 36(3):533–536, 1955.

[183] Jordi Bascompte, Pedro Jordano, and Jens M Olesen. Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, 312(5772):431–433, 2006.

[184] Ugo Bastolla, Miguel A Fortuna, Alberto Pascual-García, Antonio Ferrera, Bartolo Luque, and Jordi Bascompte. The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, 458(7241):1018–1020, 2009.

[185] Alberto Pascual-García and Ugo Bastolla. Mutualism supports biodiversity when the direct competition is weak. *Nature Communications*, 8(1):1–13, 2017.

[186] Stefano Allesina and Si Tang. Stability criteria for complex ecosystems. *Nature*, 483(7388):205–208, 2012.

[187] Pengjuan Zu, Karina Boege, Ek Del-Val, Meredith C Schuman, Philip C Stevenson, Alejandro Zaldivar-Riverón, and Serguei Saavedra. Information arms race explains plant-herbivore chemical communication in ecological communities. *Science*, 368(6497):1377–1381, 2020.

[188] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[189] Albert-Laszlo Barabasi and Zoltan N Oltvai. Network biology: understanding the cell's functional organization. *Nature reviews genetics*, 5(2):101–113, 2004.

[190] Daniela Ledezma-Tejeida, Cecilia Ishida, and Julio Collado-Vides. Genome-wide mapping of transcriptional regulation and metabolism describes information-processing units in escherichia coli. *Frontiers in microbiology*, 8:1466, 2017.

[191] Jonathan R Karr, Jayodita C Sanghvi, Derek N Macklin, Miriam V Gutschow, Jared M Jacobs, Benjamin Bolival Jr, Nacyra Assad-Garcia, John I Glass, and Markus W Covert. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389–401, 2012.

[192] Ruchir A Khandelwal, Brett G Olivier, Wilfred F M Röling, Bas Teusink, and Frank J Bruggeman. Community flux balance analysis for microbial consortia at balanced growth. *PLoS One*, 8(5):e64567, May 2013.

[193] Jason H Yang, Sarah N Wright, Meagan Hamblin, Douglas McCloskey, Miguel A Alcantar, Lars Schrübbers, Allison J Lopatkin, Sangeeta Satish, Amir Nili, Bernhard O Palsson, et al. A white-box machine learning approach for revealing antibiotic mechanisms of action. *Cell*, 177(6):1649–1661, 2019.

[194] Tobias Reichenbach, Mauro Mobilia, and Erwin Frey. Mobility promotes and jeopardizes biodiversity in rock–paper–scissors games. *Nature*, 448(7157):1046–1049, 2007.

[195] Mario E Muscarella and James P O'Dwyer. Species dynamics and interactions via metabolically informed consumer-resource models. *Theoretical Ecology*, 13(4):503–518, 2020.

[196] Crawford S Holling. The components of predation as revealed by a study of small-mammal predation of the european pine sawfly1. *The Canadian Entomologist*, 91(5):293–320, 1959.

[197] Matthieu Barbier, Jean-François Arnoldi, Guy Bunin, and Michel Loreau. Generic assembly patterns in complex ecological communities. *Proceedings of the National Academy of Sciences*, 115(9):2156–2161, 2018.

[198] Martin A Nowak. Evolutionary dynamics, 2006.

[199] Tamás L Czárán, Rolf F Hoekstra, and Ludo Pagie. Chemical warfare between microbes promotes biodiversity. *Proceedings of the National Academy of Sciences*, 99(2):786–790, 2002.

[200] Monica I Abrudan, Fokko Smakman, Ard Jan Grimbergen, Sanne Westhoff, Eric L Miller, Gilles P Van Wezel, and Daniel E Rozen. Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proceedings of the National Academy of Sciences*, 112(35):11054–11059, 2015.

[201] Kevin E Bassler, Erwin Frey, and RKP Zia. Coevolution of nodes and links: Diversity-driven coexistence in cyclic competition of three species. *Physical Review E*, 99(2):022309, 2019.

[202] Santiago Soliveres, Fernando T Maestre, Werner Ulrich, Peter Manning, Steffen Boch, Matthew A Bowker, Daniel Prati, Manuel Delgado-Baquerizo, José L Quero, Ingo Schöning, et al. Intransitive competition is widespread in plant communities and maintains their species richness. *Ecology letters*, 18(8):790–798, 2015.

[203] Logan M Higgins, Jonathan Friedman, Hao Shen, and Jeff Gore. Co-occurring soil bacteria exhibit a robust competitive hierarchy and lack of non-transitive interactions. *BioRxiv*, page 175737, 2017.

[204] J Arjan Gm De Visser and Joachim Krug. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics*, 15(7):480–490, 2014.

[205] Alvaro Sanchez. Defining higher-order interactions in synthetic ecology: lessons from physics and quantitative genetics. *Cell Systems*, 9(6):519–520, 2019.

[206] Anja Worrich, Niculina Musat, and Hauke Harms. Associational effects in the microbial neighborhood. *The ISME Journal*, 13(9):2143–2149, 2019.

[207] Eyal Bairey, Eric D Kelsic, and Roy Kishony. High-order species interactions shape ecosystem diversity. *Nature communications*, 7(1):1–7, 2016.

[208] Deepika Sundarraman, Edouard A Hay, Dylan M Martins, Drew S Shields, Noah L Pettinari, and Raghuveer Parthasarathy. Higher-order interactions dampen pairwise competition in the zebrafish gut microbiome. *Mbio*, 11(5), 2020.

[209] Jonathan M Levine, Jordi Bascompte, Peter B Adler, and Stefano Allesina. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, 546(7656):56–64, 2017.

[210] Melanie Ghoul and Sara Mitri. The ecology and evolution of microbial competition. *Trends Microbiol.*, 24(10):833–845, October 2016.

[211] Karoline Faust and Jeroen Raes. Microbial interactions: from networks to models. *Nature Reviews Microbiology*, 10(8):538–550, 2012.

[212] Ana Gomez Fons and Michel Tuomo Karjalainen. Mechanisms of colonisation and colonisation resistance of the digestive tract part 2: bacteria/bacteria interactions. *Microbial Ecology in Health and Disease*, 12(2):240–246, 2000.

[213] Bernardo Cervantes. *Tool development for the rapid identification of microbiome manipulating agents*. PhD thesis, Massachusetts Institute of Technology, 2020.

[214] Michael E Hibbing, Clay Fuqua, Matthew R Parsek, and S Brook Peterson. Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.*, 8(1):15–25, January 2010.

[215] Elisa T Granato, Thomas A Meiller-Legrand, and Kevin R Foster. The evolution and ecology of bacterial warfare. *Current biology*, 29(11):R521–R537, 2019.

[216] Cyrille Violle, Zhichao Pu, and Lin Jiang. Experimental demonstration of the importance of competition under disturbance. *Proceedings of the National Academy of Sciences*, 107(29):12925–12929, 2010.

[217] Nuno M Oliveira, Esteban Martinez-Garcia, Joao Xavier, William M Durham, Roberto Kolter, Wook Kim, and Kevin R Foster. Biofilm formation as a response to ecological competition. *PLoS Biology*, 13(7):e1002191, 2015.

[218] Christoph Ratzke and Jeff Gore. Modifying and reacting to the environmental ph can drive bacterial interactions. *PLoS Biology*, 16(3):e2004248, March 2018.

[219] Lawrence H Uricchio, S Caroline Daws, Erin R Spear, and Erin A Mordecai. Priority effects and nonhierarchical competition shape species composition in a complex grassland community. *The American Naturalist*, 193(2):213–226, 2019.

[220] Yandong Xiao, Marco Tulio Angulo, Jonathan Friedman, Matthew K Waldor, Scott T Weiss, and Yang-Yu Liu. Mapping the ecological networks of microbial communities. *Nature communications*, 8(1):1–12, 2017.

[221] Phuongan Dam, Luis M Rodriguez-R, Chengwei Luo, Janet Hatt, Despina Tsementzi, Konstantinos T Konstantinidis, and Eberhard O Voit. Model-based comparisons of the abundance dynamics of bacterial communities in two lakes. *Scientific reports*, 10(1):1–12, 2020.

[222] Daniel Aguirre de Cárcer. Experimental and computational approaches to unravel microbial community assembly. *Computational and Structural Biotechnology Journal*, 18:4071, 2020.

[223] Chitong Rao, Katharine Z Coyte, Wayne Bainter, Raif S Geha, Camilia R Martin, and Seth Rakoff-Nahoum. Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature*, 591(7851):633–638, 2021.

[224] Roie Levy and Elhanan Borenstein. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proceedings of the National Academy of Sciences*, 110(31):12804–12809, 2013.

[225] Benjamin G Morgan, Paul Warren, Ryan E Mewis, and Damian W Rivett. Bacterial dominance is due to effective utilisation of secondary metabolites produced by competitors. *Scientific reports*, 10(1):1–8, 2020.

[226] Jonathan Friedman and Eric J Alm. Inferring correlation networks from genomic survey data. *PLoS Computational Biology*, 8(9):e1002687, 2012.

[227] Zachary D Kurtz, Christian L Müller, Emily R Miraldi, Dan R Littman, Martin J Blaser, and Richard A Bonneau. Sparse and compositionally robust inference of microbial ecological networks. *PLoS Computational Biology*, 11(5):e1004226, 2015.

[228] Hokuto Hirano and Kazuhiro Takemoto. Difficulty in inferring microbial community structure based on co-occurrence network approaches. *BMC bioinformatics*, 20(1):1–14, 2019.

[229] Daniel S Maynard, J Timothy Wootton, Carlos A Serván, and Stefano Allesina. Reconciling empirical interactions and species coexistence. *Ecology letters*, 22(6):1028–1037, 2019.

[230] Julia Johnke, Philipp Dirksen, and Hinrich Schulenburg. Community assembly of the native C . elegans microbiome is influenced by time, substrate and individual bacterial taxa. *Environmental Microbiology*, 22(4):1265–1279, 2020.

[231] Johannes Zimmermann, Nancy Obeng, Wentao Yang, Barbara Pees, Carola Petersen, Silvio Waschina, Kohar A Kissoyan, Jack Aidley, Marc P Hoeppner, Boyke Bunk, Cathrin Spröer, Matthias Leippe, Katja Dierking, Christoph Kaleta, and Hinrich Schulenburg. The functional repertoire contained within the native microbiota of the model nematode caenorhabditis elegans. *The ISME Journal*, 14(1):26–38, January 2020.

[232] Kevin R Foster and Thomas Bell. Competition, not cooperation, dominates interactions among culturable microbial species. *Current Biology*, 22(19):1845–1850, 2012.

[233] Xiaoqian Yu, Martin F Polz, and Eric J Alm. Interactions in self-assembled microbial communities saturate with diversity. *The ISME Journal*, 13(6):1602–1617, 2019.

[234] Bernard J Crespi. The evolution of social behavior in microorganisms. *Trends in ecology & evolution*, 16(4):178–183, 2001.

[235] Nicole M Vega, Kyle R Allison, Amanda N Samuels, Mark S Klempner, and James J Collins. Salmonella typhimurium intercepts escherichia coli signaling to enhance antibiotic tolerance. *Proceedings of the National Academy of Sciences*, 110(35):14420–14425, 2013.

[236] Nicole M Vega and Jeff Gore. Collective antibiotic resistance: mechanisms and implications. *Current opinion in microbiology*, 21:28–34, 2014.

[237] Hongyue Dang and Charles R Lovell. Microbial surface colonization and biofilm development in marine environments. *Microbiology and molecular biology reviews: MMBR*, 80(1):91, 2016.

[238] Thomas Pfeiffer and Sebastian Bonhoeffer. Evolution of cross-feeding in microbial populations. *The American naturalist*, 163(6):E126–E135, 2004.

[239] Glen D'Souza, Shraddha Shitut, Daniel Preussger, Ghada Yousif, Silvio Waschina, and Christian Kost. Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports*, 35(5):455–488, 2018.

[240] OM Neijssel and DW Tempest. The regulation of carbohydrate metabolism in klebsiella aerogenes nctc 418 organisms, growing in chemostat culture. *Archives of Microbiology*, 106(3):251–258, 1975.

[241] James B Russell and Gregory M Cook. Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiological reviews*, 59(1):48, 1995.

[242] Muriel Cocaign-Bousquet and Nicholas D Lindley. Pyruvate overflow and carbon flux within the central metabolic pathways of corynebacterium glutamicum during growth on lactate. *Enzyme and microbial technology*, 17(3):260–267, 1995.

[243] Kelly P Williams, Joseph J Gillespie, Bruno WS Sobral, Eric K Nordberg, Eric E Snyder, Joshua M Shallom, and Allan W Dickerman. Phylogeny of gammaproteobacteria. *Journal of bacteriology*, 192(9):2305–2314, 2010.

[244] Na-Ri Shin, Tae Woong Whon, and Jin-Woo Bae. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends in biotechnology*, 33(9):496–503, 2015.

[245] Xiaokan Guo and James Q Boedicker. The contribution of high-order metabolic interactions to the global activity of a four-species microbial community. *PLoS Computational Biology*, 12(9):e1005079, 2016.

[246] Mark Loftus, Sayf Al-Deen Hassouneh, and Shibu Yooseph. Bacterial associations in the healthy human gut microbiome across populations. *Scientific reports*, 11(1):1–14, 2021.

[247] Stuart A West, Ashleigh S Griffin, Andy Gardner, and Stephen P Diggle. Social evolution theory for microorganisms. *Nature reviews microbiology*, 4(8):597–607, 2006.

[248] Jeff Gore, Hyun Youk, and Alexander Van Oudenaarden. Snowdrift game dynamics and facultative cheating in yeast. *Nature*, 459(7244):253–256, 2009.

[249] Giles ED Oldroyd. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology*, 11(4):252–263, 2013.

[250] Joelle Sasse, Enrico Martinoia, and Trent Northen. Feed your friends: do plant exudates shape the root microbiome? *Trends in plant science*, 23(1):25–41, 2018.

[251] Saritha Mohanram and Praveen Kumar. Rhizosphere microbiome: revisiting the synergy of plant-microbe interactions. *Annals of Microbiology*, 69(4):307–320, 2019.

[252] Abhishek Shrivastava, Visha K Patel, Yisha Tang, Susan Connolly Yost, Floyd E Dewhirst, and Howard C Berg. Cargo transport shapes the spatial organization of a microbial community. *Proceedings of the National Academy of Sciences*, 115(34):8633–8638, 2018.

[253] Jessica L Mark Welch, Blair J Rossetti, Christopher W Rieken, Floyd E Dewhirst, and Gary G Borisy. Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences*, 113(6):E791–800, February 2016.

[254] Cristal Zuñiga, Chien-Ting Li, Geng Yu, Mahmoud M Al-Bassam, Tingting Li, Liqun Jiang, Livia S Zaramela, Michael Guarnieri, Michael J Betenbaugh, and Karsten Zengler. Environmental stimuli drive a transition from cooperation to competition in synthetic phototrophic communities. *Nature microbiology*, 4(12):2184–2191, 2019.

[255] Chun-Hui Gao, Hui Cao, Peng Cai, and Søren J Sørensen. The initial inoculation ratio regulates bacterial coculture interactions and metabolic capacity. *The ISME Journal*, 15(1):29–40, 2021.

[256] Gianalberto Losapio, Christian Schöb, Phillip PA Staniczenko, Francesco Carrara, Gian Marco Palamara, Consuelo M De Moraes, Mark C Mescher, Rob W Brooker, Bradley J Butterfield, Ragan M Callaway, et al. Network motifs involving both competition and facilitation predict biodiversity in alpine plant communities. *Proceedings of the National Academy of Sciences*, 118(6), 2021.

[257] Ilana L Brito, S Yilmaz, K Huang, Liyi Xu, Stacy D Jupiter, Aaron P Jenkins, Waisea Naisilisili, M Tamminen, CS Smillie, Jennifer R Wortman, et al. Mobile genes in the human microbiome are structured from global to individual scales. *Nature*, 535(7612):435–439, 2016.

[258] Simon P Hart, Martin M Turcotte, and Jonathan M Levine. Effects of rapid evolution on species coexistence. *Proceedings of the National Academy of Sciences*, 116(6):2112–2117, 2019.

[259] Jared Kehe, Anthony Kulesa, Anthony Ortiz, Cheri M Ackerman, Sri Gowtham Thakku, Daniel Sellers, Seppe Kuehn, Jeff Gore, Jonathan Friedman, and Paul C Blainey. Massively parallel screening of synthetic microbial communities. *Proceedings of the National Academy of Sciences*, 116(26):12804–12809, June 2019.

[260] Sara Mitri and Kevin Richard Foster. The genotypic view of social interactions in microbial communities. *Annual Review of Genetics*, 47:247–273, August 2013.

[261] Daniel M Cornforth and Kevin R Foster. Competition sensing: the social side of bacterial stress responses. *Nat. Rev. Microbiol.*, 11(4):285–293, April 2013.

[262] Allan Konopka, Stephen Lindemann, and Jim Fredrickson. Dynamics in microbial communities: unraveling mechanisms to identify principles. *The ISME Journal*, 9(7):1488–1495, 2015.

[263] Eric D Kelsic, Jeffrey Zhao, Kalin Vetsigian, and Roy Kishony. Counteraction of antibiotic production and degradation stabilizes microbial communities. *Nature*, 521(7553):516–519, May 2015.

[264] Babak Momeni, Li Xie, and Wenying Shou. Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. *Elife*, 6, March 2017.

[265] Lea Benedicte Skov Hansen, Dawei Ren, Mette Burmølle, and Søren J Sørensen. Distinct gene expression profile of xanthomonas retroflexus engaged in synergistic multispecies biofilm formation. *The ISME Journal*, 11(1):300–303, January 2017.

[266] Margaret M Mayfield and Daniel B Stouffer. Higher-order interactions capture unexplained complexity in diverse communities. *Nat Ecol Evol*, 1(3):62, February 2017.

[267] Alicia Sanchez-Gorostiaga, Djordje Bajić, Melisa L Osborne, Juan F Poyatos, and Alvaro Sanchez. High-order interactions distort the functional landscape of microbial consortia. *PLoS Biology*, 17(12):e3000550, December 2019.

[268] Glenn E Croston. Functional cell-based uHTS in chemical genomic drug discovery. *Trends Biotechnol.*, 20(3):110–115, March 2002.

[269] Christophe J Echeverri and Norbert Perrimon. High-throughput RNAi screening in cultured cells: a user's guide. *Nat. Rev. Genet.*, 7(5):373–384, 2006.

[270] S A Sundberg. High-throughput and ultra-high-throughput screening: solution- and cell-based approaches. *Curr. Opin. Biotechnol.*, 11(1):47–53, February 2000.

[271] Corrado Nai and Vera Meyer. From axenic to mixed cultures: technological advances accelerating a paradigm shift in microbiology. *Trends in microbiology*, 26(6):538–554, 2018.

[272] M T Mee, J J Collins, G M Church, and H H Wang. Syntrophic exchange in synthetic microbial communities. *Proceedings of the National Academy of Sciences*, 111(20):E2149–E2156, 2014.

[273] Ophelia S Venturelli, Alex C Carr, Garth Fisher, Ryan H Hsu, Rebecca Lau, Benjamin P Bowen, Susan Hromada, Trent Northen, and Adam P Arkin. Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.*, 14(6):e8157, June 2018.

[274] R A Long and F Azam. Antagonistic interactions among marine pelagic bacteria. *Appl. Environ. Microbiol.*, 67(11):4975–4983, November 2001.

[275] Anthony Kulesa, Jared Kehe, Juan E Hurtado, Prianca Tawde, and Paul C Blainey. Combinatorial drug discovery in nanoliter droplets. *Proceedings of the National Academy of Sciences*, 115(26):6685–6690, 2018.

[276] Colin J Ingham, Ad Sprenkels, Johan Bomer, Douwe Molenaar, Albert van den Berg, Johan E T van Hylckama Vlieg, and Willem M de Vos. The micro-petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. *Proceedings of the National Academy of Sciences*, 104(46):18217–18222, November 2007.

[277] Lifeng Kang, Matthew J Hancock, Mark D Brigham, and Ali Khademhosseini. Cell confinement in patterned nanoliter droplets in a microwell array by wiping. *J. Biomed. Mater. Res. A*, 9999A:NA–NA, 2009.

[278] Samaneh Mashaghi, Alireza Abbaspourrad, David A Weitz, and Antoine M van Oijen. Droplet microfluidics: A tool for biology, chemistry and nanotechnology. *Trends Analyt. Chem.*, 82:118–125, 2016.

[279] Tomasz S Kaminski, Ott Scheler, and Piotr Garstecki. Droplet microfluidics for microbiology: techniques, applications and challenges. *Lab Chip*, 16(12):2168–2187, June 2016.

[280] Rainer U Meckenstock, Frederick von Netzer, Christine Stumpp, Tillmann Lueders, Anne M Himmelberg, Norbert Hertkorn, Philipp Schmitt-Kopplin, Mourad Harir, Riad Hosein, Shirin Haque, and Dirk Schulze-Makuch. Oil biodegradation. water droplets in oil are microhabitats for microbial life. *Science*, 345(6197):673–676, August 2014.

[281] Otto X Cordero and Manoshi S Datta. Microbial interactions and community assembly at microscales. *Curr. Opin. Microbiol.*, 31:227–234, June 2016.

[282] Reed M Stubbendieck, Carol Vargas-Bautista, and Paul D Straight. Bacterial communities: Interactions to scale. *Front. Microbiol.*, 7, 2016.

[283] Margaret McFall-Ngai, Michael G Hadfield, Thomas C G Bosch, Hannah V Carey, Tomislav Domazet-Lošo, Angela E Douglas, Nicole Dubilier, Gerard Eberl, Tadashi Fukami, Scott F Gilbert, Ute Hentschel, Nicole King, Staffan Kjelleberg, Andrew H Knoll, Natacha Kremer, Sarkis K Mazmanian, Jessica L Metcalf, Kenneth Nealson, Naomi E Pierce, John F Rawls, Ann Reid, Edward G Ruby, Mary Rumpho, Jon G Sanders, Diethard Tautz, and Jennifer J Wernegreen. Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, 110(9):3229–3236, 2013.

[284] Lynn Margulis, Elso S Barghoorn, Debra Ashendorf, Sumana Banerjee, David Chase, Susan Francis, Stephen Giovannoni, and John Stolz. The microbial community in the layered sediments at laguna figueroa, baja california, mexico: Does it have precambrian analogues? *Precambrian Research*, 11(2):93–123, 1980.

[285] Lynn Sagan. On the origin of mitosing cells. *Journal of theoretical biology*, 14(3):225–IN6, 1967.

[286] Jesse R Zaneveld, Ryan McMinds, and Rebecca Vega Thurber. Stress and stability: applying the anna karenina principle to animal microbiomes. *Nature microbiology*, 2(9):1–8, 2017.

[287] Lei Dai, Daan Vorselen, Kirill S Korolev, and Jeff Gore. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*, 336(6085):1175–1177, 2012.

[288] Matthew K Waldor, Gene Tyson, Elhanan Borenstein, Howard Ochman, Andrew Moeller, B Brett Finlay, Heidi H Kong, Jeffrey I Gordon, Karen E Nelson, Karim Dabbagh, et al. Where next for microbiome research? *PLoS Biology*, 13(1):e1002050, 2015.

[289] Christian Rinke, Patrick Schwientek, Alexander Sczyrba, Natalia N Ivanova, Iain J Anderson, Jan-Fang Cheng, Aaron Darling, Stephanie Malfatti, Brandon K Swan, Esther A Gies, et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, 499(7459):431–437, 2013.

[290] Donovan H Parks, Christian Rinke, Maria Chuvochina, Pierre-Alain Chaumeil, Ben J Woodcroft, Paul N Evans, Philip Hugenholtz, and Gene W Tyson. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nature microbiology*, 2(11):1533–1542, 2017.

[291] Robin Tecon, Sara Mitri, Davide Ciccarese, Dani Or, Jan Roelof van der Meer, and David R Johnson. Bridging the holistic-reductionist divide in microbial ecology. *Msystems*, 4(1):e00265–18, 2019.

[292] Ji-Hyun Yun, Seong Woon Roh, Tae Woong Whon, Mi-Ja Jung, Min-Soo Kim, Doo-Sang Park, Changmann Yoon, Young-Do Nam, Yun-Ji Kim, Jung-Hye Choi, Joon-Yong Kim, Na-Ri Shin, Sung-Hee Kim, Won-Jae Lee, and Jin-Woo Bae. Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Appl. Environ. Microbiol.*, 80(17):5254–5264, September 2014.

[293] Ruth E Ley, Peter J Turnbaugh, Samuel Klein, and Jeffrey I Gordon. Human gut microbes associated with obesity. *Nature*, 444(7122):1022–1023, 2006.

[294] Rachel N Carmody, Georg K Gerber, Jesus M Luevano Jr, Daniel M Gatti, Lisa Somes, Karen L Svenson, and Peter J Turnbaugh. Diet dominates host genotype in shaping the murine gut microbiota. *Cell host & microbe*, 17(1):72–84, 2015.

[295] Benjamin E Wolfe. Using cultivated microbial communities to dissect microbiome assembly: challenges, limitations, and the path ahead. *Msystems*, 3(2):e00161–17, 2018.

[296] Andrew B Shreiner, John Y Kao, and Vincent B Young. The gut microbiome in health and in disease. *Current opinion in gastroenterology*, 31(1):69, 2015.

[297] Carlotta De Filippo, Duccio Cavalieri, Monica Di Paola, Matteo Ramazzotti, Jean Baptiste Poullet, Sebastien Massart, Silvia Collini, Giuseppe Pieraccini, and Paolo Lionetti. Impact of diet in shaping gut microbiota revealed by a comparative study in children from europe and rural africa. *Proceedings of the National Academy of Sciences*, 107(33):14691–14696, 2010.

[298] Lawrence A David, Corinne F Maurice, Rachel N Carmody, David B Gootenberg, Julie E Button, Benjamin E Wolfe, Alisha V Ling, A Sloan Devlin, Yug Varma, Michael A Fischbach, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*, 505(7484):559–563, 2014.

[299] Elham Hosseini, Charlotte Grootaert, Willy Verstraete, and Tom Van de Wiele. Propionate as a health-promoting microbial metabolite in the human gut. *Nutrition reviews*, 69(5):245–258, 2011.

[300] Serena Sanna, Natalie R van Zuydam, Anubha Mahajan, Alexander Kurilshikov, Arnau Vich Vila, Urmo Võsa, Zlatan Mujagic, Ad AM Masclee, Daisy MAE Jonkers, Marije Oosting, et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nature genetics*, 51(4):600–605, 2019.

[301] Lawrence A David, Ana Weil, Edward T Ryan, Stephen B Calderwood, Jason B Harris, Fahima Chowdhury, Yasmin Begum, Firdausi Qadri, Regina C LaRocque, and Peter J Turnbaugh. Gut microbial succession follows acute secretory diarrhea in humans. *MBio*, 6(3):e00381–15, 2015.

[302] Guanxiang Liang, Chunyu Zhao, Huanjia Zhang, Lisa Mattei, Scott Sherrill-Mix, Kyle Bittinger, Lyanna R Kessler, Gary D Wu, Robert N Baldassano, Patricia DeRusso, et al. The stepwise assembly of the neonatal virome is modulated by breastfeeding. *Nature*, 581(7809):470–474, 2020.

[303] Yanjiao Zhou, Daniel Jackson, Leonard B Bacharier, David Mauger, Homer Boushey, Mario Castro, Juliana Durack, Yvonne Huang, Robert F Lemanske, Gregory A Storch, et al. The upper-airway microbiota and loss of asthma control among asthmatic children. *Nature communications*, 10(1):1–10, 2019.

[304] Ilke De Boeck, Marianne FL van den Broek, Camille N Allonsius, Irina Spacova, Stijn Wittouck, Katleen Martens, Sander Wuyts, Eline Cauwenberghs, Katarina Jokicevic, Dieter Vandenheuvel, et al. Lactobacilli have a niche in the human nose. *Cell Reports*, 31(8):107674, 2020.

[305] Ilseung Cho and Martin J Blaser. The human microbiome: at the interface of health and disease. *Nature Reviews Genetics*, 13(4):260–270, 2012.

[306] Pernilla Lif Holgerson, L Harnevik, Olle Hernell, ACR Tanner, and Ingegerd Johansson. Mode of birth delivery affects oral microbiota in infants. *Journal of dental research*, 90(10):1183–1188, 2011.

[307] Juan Miguel Rodríguez, Kiera Murphy, Catherine Stanton, R Paul Ross, Olivia I Kober, Nathalie Juge, Ekaterina Avershina, Knut Rudi, Arjan Narbad, Maria C Jenmalm, et al. The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial ecology in health and disease*, 26(1):26050, 2015.

[308] Shijie Zhao, Tami D Lieberman, Mathilde Poyet, Kathryn M Kauffman, Sean M Gibbons, Mathieu Groussin, Ramnik J Xavier, and Eric J Alm. Adaptive evolution within gut microbiomes of healthy people. *Cell host & microbe*, 25(5):656–667, 2019.

[309] Henning Seedorf, Nicholas W Griffin, Vanessa K Ridaura, Alejandro Reyes, Jiye Cheng, Federico E Rey, Michelle I Smith, Gabriel M Simon, Rudolf H Scheffrahn, Dagmar Woebken, Alfred M Spormann, William Van Treuren, Luke K Ursell, Megan Pirrung, Adam Robbins-Pianka, Brandi L Cantarel, Vincent Lombard, Bernard Henrissat, Rob Knight, and Jeffrey I Gordon. Bacteria from diverse habitats colonize and compete in the mouse gut. *Cell*, 159(2):253–266, October 2014.

[310] Mahesh S Desai, Anna M Seekatz, Nicole M Koropatkin, Nobuhiko Kamada, Christina A Hickey, Mathis Wolter, Nicholas A Pudlo, Sho Kitamoto, Nicolas Terrapon, Arnaud Muller, et al. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell*, 167(5):1339–1353, 2016.

[311] Carolina Tropini, Eli Lin Moss, Bryan Douglas Merrill, Katharine Michelle Ng, Steven Kyle Higginbottom, Ellen Pun Casavant, Carlos Gutierrez Gonzalez, Brayon Fremin, Donna Michelle Bouley, Joshua Eric Elias, et al. Transient osmotic perturbation causes long-term alteration to the gut microbiota. *Cell*, 173(7):1742–1754, 2018.

[312] Gina R Lewin, Apollo Stacy, Kelly L Michie, Richard J Lamont, and Marvin Whiteley. Large-scale identification of pathogen essential genes during coinfection with sympatric and allopatric microbes. *Proceedings of the National Academy of Sciences*, 116(39):19685–19694, 2019.

[313] Ann K Corsi, Bruce Wightman, and Martin Chalfie. A transparent window into biology: a primer on caenorhabditis elegans. *Genetics*, 200(2):387–407, 2015.

[314] Marie-Anne Félix and Christian Braendle. The natural history of caenorhabditis elegans. *Current biology*, 20(22):R965–R969, 2010.

[315] L Byerly, RC Cassada, and RL Russell. The life cycle of the nematode caenorhabditis elegans: I. wild-type growth and reproduction. *Developmental biology*, 51(1):23–33, 1976.

[316] James D McGhee. The caenorhabditis elegans intestine. *Wiley Interdisciplinary Reviews: Developmental Biology*, 2(3):347–367, 2013.

[317] Erik Allman, David Johnson, and Keith Nehrke. Loss of the apical V-ATPase a-subunit VHA-6 prevents acidification of the intestinal lumen during a rhythmic behavior in c. elegans. *American Journal of Physiology-Cell Physiology*, 297(5):C1071–C1081, 2009.

[318] James McGhee and Shervin Ghafouri. Bacterial residence time in the intestine of caenorhabditis elegans. *Nematology*, 9(1):87–91, 2007.

[319] Veeren M Chauhan, Gianni Orsi, Alan Brown, David I Pritchard, and Jonathan W Aylott. Mapping the pharyngeal and intestinal ph ofcaenorhabditis elegansand Real-Time luminal ph oscillations using extended dynamic range pH-Sensitive nanosensors. *ACS Nano*, 7(6):5577–5587, 2013.

[320] Marie-Anne Félix and Fabien Duveau. Population dynamics and habitat sharing of natural populations of caenorhabditis elegans and c. briggsae. *BMC biology*, 10(1):1–19, 2012.

[321] Marina Kniazeva and Gary Ruvkun. Rhizobium induces dna damage in caenorhabditis elegans intestinal cells. *Proceedings of the National Academy of Sciences*, 116(9):3784–3792, 2019.

[322] Filipe Cabreiro and David Gems. Worms need microbes too: microbiota, health and aging in caenorhabditis elegans. *EMBO Mol. Med.*, 5(9):1300–1310, September 2013.

[323] Arun Kumar, Aiswarya Baruah, Masahiro Tomioka, Yuichi Iino, Mohan C Kalita, and Mojibur Khan. Caenorhabditis elegans: a model to understand host–microbe interactions. *Cellular and Molecular Life Sciences*, 77(7):1229–1249, 2020.

[324] Man-Wah Tan, Shalina Mahajan-Miklos, and Frederick M Ausubel. Killing of caenorhabditis elegans by pseudomonas aeruginosa used to model mammalian bacterial pathogenesis. *Proceedings of the National Academy of Sciences*, 96(2):715–720, 1999.

[325] C Léopold Kurz, Sophie Chauvet, Emmanuel Andrès, Marianne Aurouze, Isabelle Vallet, Gérard PF Michel, Mitch Uh, Jean Celli, Alain Filloux, Sophie De Bentzmann, et al. Virulence factors of the human opportunistic pathogen serratia marcescens identified by in vivo screening. *The EMBO journal*, 22(7):1451–1460, 2003.

[326] Hanna Fares and Iva Greenwald. Genetic analysis of endocytosis in caenorhabditis elegans: coelomocyte uptake defective mutants. *Genetics*, 159(1):133–145, 2001.

[327] Mengzhou Zhou, Hai Yu, Xianhua Yin, Parviz M Sabour, Wei Chen, and Joshua Gong. Lactobacillus zeae protects caenorhabditis elegans from enterotoxigenic escherichia coli-caused death by inhibiting enterotoxin gene expression of the pathogen. *PloS One*, 9(2):e89004, 2014.

[328] Alejandro Aballay, Peter Yorgey, and Frederick M Ausubel. Salmonella typhimurium proliferates and establishes a persistent infection in the intestine of caenorhabditis elegans. *Current Biology*, 10(23):1539–1542, 2000.

[329] Danielle A Garsin, Costi D Sifri, Eleftherios Mylonakis, Xiang Qin, Kavindra V Singh, Barbara E Murray, Stephen B Calderwood, and Frederick M Ausubel. A simple model host for identifying gram-positive virulence factors. *Proceedings of the National Academy of Sciences*, 98(19):10892–10897, 2001.

[330] Delia O'Rourke, Dilair Baban, Maria Demidova, Richard Mott, and Jonathan Hodgkin. Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of c. elegans with m. nematophilum. *Genome research*, 16(8):1005–1016, 2006.

[331] Eleftherios Mylonakis, Frederick M Ausubel, John R Perfect, Joseph Heitman, and Stephen B Calderwood. Killing of caenorhabditis elegans by cryptococcus neoformans as a model of yeast pathogenesis. *Proceedings of the National Academy of Sciences*, 99(24):15675–15680, 2002.

[332] Anton Y Peleg, Emmanouil Tampakakis, Beth Burgwyn Fuchs, George M Eliopoulos, Robert C Moellering, and Eleftherios Mylonakis. Prokaryote–eukaryote interactions identified by using caenorhabditis elegans. *Proceedings of the National Academy of Sciences*, 105(38):14585–14590, 2008.

[333] Marcelo Ortega-Riveros, Iker De-la Pinta, Cristina Marcos-Arias, Guillermo Ezpeleta, Guillermo Quindós, and Elena Eraso. Usefulness of the non-conventional caenorhabditis elegans model to assess candida virulence. *Mycopathologia*, 182(9):785–795, 2017.

[334] Emily R Troemel, Marie-Anne Félix, Noah K Whiteman, Antoine Barrière, and Frederick M Ausubel. Microsporidia are natural intracellular parasites of the nematode caenorhabditis elegans. *PLoS Biology*, 6(12):e309, 2008.

[335] Keir M Balla and Emily R Troemel. C aenorhabditis elegans as a model for intracellular pathogen infection. *Cellular microbiology*, 15(8):1313–1322, 2013.

[336] Costi D Sifri, Jakob Begun, and Frederick M Ausubel. The worm has turned–microbial virulence modeled in caenorhabditis elegans. *Trends in microbiology*, 13(3):119–127, 2005.

[337] Emily R Troemel, Stephanie W Chu, Valerie Reinke, Siu Sylvia Lee, Frederick M Ausubel, and Dennis H Kim. p38 mapk regulates expression of immune response genes and contributes to longevity in c. elegans. *PLoS Genetics*, 2(11):e183, 2006.

[338] Robert P Shivers, Tristan Kooistra, Stephanie W Chu, Daniel J Pagano, and Dennis H Kim. Tissue-specific activities of an immune signaling module regulate physiological responses to pathogenic and nutritional bacteria in c. elegans. *Cell host & microbe*, 6(4):321–330, 2009.

[339] Laura C Clark and Jonathan Hodgkin. Commensals, probiotics and pathogens in the c aenorhabditis elegans model. *Cellular microbiology*, 16(1):27–38, 2014.

[340] Olivier Zugasti and Jonathan J Ewbank. Neuroimmune regulation of antimicrobial peptide expression by a noncanonical tgf-β signaling pathway in caenorhabditis elegans epidermis. *Nature immunology*, 10(3):249–256, 2009.

[341] Javier E Irazoqui, Jonathan M Urbach, and Frederick M Ausubel. Evolution of host innate defence: insights from caenorhabditis elegans and primitive invertebrates. *Nature Reviews Immunology*, 10(1):47–58, 2010.

[342] Hannah R Nicholas and Jonathan Hodgkin. Responses to infection and possible recognition strategies in the innate immune system of caenorhabditis elegans. *Molecular immunology*, 41(5):479–493, 2004.

[343] Hinrich Schulenburg, Marc P Hoeppner, January Weiner, and Erich Bornberg-Bauer. Specificity of the innate immune system and diversity of c-type lectin domain (CTLD) proteins in the nematode caenorhabditis elegans. *Immunobiology*, 213(3-4):237–250, 2008.

[344] Lien Callewaert and Chris W Michiels. Lysozymes in the animal kingdom. *Journal of biosciences*, 35(1):127–160, 2010.

[345] Jonathan J Ewbank and Olivier Zugasti. C. elegans: model host and tool for antimicrobial drug discovery. *Disease models & mechanisms*, 4(3):300–304, 2011.

[346] Wentao Yang, Katja Dierking, Daniela Esser, Andreas Tholey, Matthias Leippe, Philip Rosenstiel, and Hinrich Schulenburg. Overlapping and unique signatures in the proteomic and transcriptomic responses of the nematode caenorhabditis elegans toward pathogenic bacillus thuringiensis. *Developmental & Comparative Immunology*, 51(1):1–9, 2015.

[347] Evan L Ardiel and Catharine H Rankin. An elegant mind: learning and memory in caenorhabditis elegans. *Learning & memory*, 17(4):191–201, 2010.

[348] Boris Borisovich Shtonda and Leon Avery. Dietary choice behavior in caenorhabditis elegans. *Journal of experimental biology*, 209(1):89–102, 2006.

[349] Joshua D Meisel and Dennis H Kim. Behavioral avoidance of pathogenic bacteria by caenorhabditis elegans, 2014.

[350] Mario De Bono. Molecular approaches to aggregation behavior and social attachment. *Journal of neurobiology*, 54(1):78–92, 2003.

[351] Michele Sammut, Steven J Cook, Ken CQ Nguyen, Terry Felton, David H Hall, Scott W Emmons, Richard J Poole, and Arantza Barrios. Glia-derived neurons are required for sex-specific learning in c. elegans. *Nature*, 526(7573):385–390, 2015.

[352] Jeffrey L Rhoades, Jessica C Nelson, Ijeoma Nwabudike, K Yu Stephanie, Ian G McLachlan, Gurrein K Madan, Eden Abebe, Joshua R Powers, Daniel A Colón-Ramos, and Steven W Flavell. Asics mediate food responses in an enteric serotonergic neuron that controls foraging behaviors. *Cell*, 176(1-2):85–97, 2019.

[353] Malin EV Johansson, Mia Phillipson, Joel Petersson, Anna Velcich, Lena Holm, and Gunnar C Hansson. The inner of the two muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proceedings of the National Academy of Sciences*, 105(39):15064–15069, 2008.

[354] Tiange Lang, Gunnar C Hansson, and Tore Samuelsson. Gel-forming mucins appeared early in metazoan evolution. *Proceedings of the National Academy of Sciences*, 104(41):16209–16214, 2007.

[355] Casandra L Hoffman, Jonathan Lalsiamthara, and Alejandro Aballay. Host mucin is exploited by pseudomonas aeruginosa to provide monosaccharides required for a successful infection. *MBio*, 11(2), March 2020.

[356] Younghoon Kim and Eleftherios Mylonakis. Caenorhabditis elegans immune conditioning with the probiotic bacterium lactobacillus acidophilus strain ncfm enhances gram-positive immune responses. *Infection and immunity*, 80(7):2500–2508, 2012.

[357] Levi T Morran, Olivia G Schmidt, Ian A Gelarden, Raymond C Parrish, and Curtis M Lively. Running with the red queen: host-parasite coevolution selects for biparental sex. *Science*, 333(6039):216–218, 2011.

[358] Kayla C King, Michael A Brockhurst, Olga Vasieva, Steve Paterson, Alex Betts, Suzanne A Ford, Crystal L Frost, Malcolm J Horsburgh, Sam Haldenby, and Gregory D D Hurst. Rapid evolution of microbe-mediated protection against pathogens in a worm host. *The ISME Journal*, 10(8):1915–1924, 2016.

[359] Suzanne A Ford, David Williams, Steve Paterson, and Kayla C King. Co-evolutionary dynamics between a defensive microbe and a pathogen driven by fluctuating selection. *Molecular ecology*, 26(7):1778–1789, 2017.

[360] Fan Zhang, Maureen Berg, Katja Dierking, Marie-Anne Félix, Michael Shapira, Buck S Samuel, and Hinrich Schulenburg. as a model for microbiome research. *Front. Microbiol.*, 8:485, March 2017.

[361] Sirena Montalvo-Katz, Hao Huang, Michael David Appel, Maureen Berg, and Michael Shapira. Association with soil bacteria enhances p38-dependent infection resistance in caenorhabditis elegans. *Infection and Immunity*, 81(2):514–520, February 2013.

[362] Carola Petersen, Philipp Dirksen, Swantje Prahl, Eike Andreas Strathmann, and Hinrich Schulenburg. The prevalence of caenorhabditis elegans across 1.5 years in selected north german locations: the importance of substrate type, abiotic parameters, and caenorhabditis competitors. *BMC ecology*, 14(1):1–10, 2014.

[363] Kohar A B Kissoyan, Moritz Drechsler, Eva-Lena Stange, Johannes Zimmermann, Christoph Kaleta, Helge B Bode, and Katja Dierking. Natural c. elegans microbiota protects against infection via production of a cyclic lipopeptide of the viscosin group. *Curr. Biol.*, 29(6):1030–1037.e5, March 2019.

[364] Buck S Samuel, Holli Rowedder, Christian Braendle, Marie-Anne Félix, and Gary Ruvkun. Caenorhabditis elegans responses to bacteria from its natural habitats. *Proceedings of the National Academy of Sciences*, 113(27):E3941–E3949, 2016.

[365] Philipp Dirksen, Sarah Arnaud Marsh, Ines Braker, Nele Heitland, Sophia Wagner, Rania Nakad, Sebastian Mader, Carola Petersen, Vienna Kowallik, Philip Rosenstiel, Marie-Anne Félix, and Hinrich Schulenburg. The native microbiome of the nematode caenorhabditis elegans: gateway to a new host-microbiome model. *BMC Biology*, 14(1), 2016.

[366] Anthony Ortiz, Nicole M Vega, Christoph Ratzke, and Jeff Gore. Interspecies bacterial competition regulates community assembly in the c. elegans intestine. *The ISME Journal*, pages 1–15, 2021.

[367] Robert B Fierer & Jackson. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, 103(3):626–631, January 2006.

[368] Angela E Douglas. Multiorganismal insects: diversity and function of resident microorganisms. *Annual Review of Entomology*, 60:17–34, January 2015.

[369] J K Nicholson, E Holmes, J Kinross, R Burcelin, G Gibson, W Jia, and S Pettersson. Host-Gut microbiota metabolic interactions. *Science*, 336(6086):1262–1267, 2012.

[370] Les Dethlefsen, Margaret McFall-Ngai, and David A Relman. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature*, 449(7164):811–818, 2007.

[371] Gil Sharon, Timothy R Sampson, Daniel H Geschwind, and Sarkis K Mazmanian. The central nervous system and the gut microbiome. *Cell*, 167(4):915–932, November 2016.

[372] Philipp Engel and Nancy A Moran. The gut microbiota of insects–diversity in structure and function. *FEMS microbiology reviews*, 37(5):699–735, 2013.

[373] Guus Roeselers, Erika K Mittge, W Zac Stephens, David M Parichy, Colleen M Cavanaugh, Karen Guillemin, and John F Rawls. Evidence for a core gut microbiota in the zebrafish. *The ISME Journal*, 5(10):1595–1608, October 2011.

[374] Adam C-N Wong, John M Chaston, and Angela E Douglas. The inconstant gut microbiota of drosophila species revealed by 16S rRNA gene analysis. *The ISME Journal*, 7(10):1922–1932, October 2013.

[375] Karen L Adair and Angela E Douglas. Making a microbiome: the many determinants of host-associated microbial community composition. *Curr. Opin. Microbiol.*, 35:23–29, February 2017.

[376] Seth R Bordenstein and Kevin R Theis. Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS Biology*, 13(8):e1002226, August 2015.

[377] Nancy A Moran and Daniel B Sloan. The hologenome concept: Helpful or hollow? *PLoS Biology*, 13(12):e1002311, 2015.

[378] Xicotencatl Gracida and ChristianÂ R. Eckmann. Fertility and germline stem cell maintenance under different diets requires nhr-114/hnf4 in *C. elegans*. *Current Biology*, 23(7):607–613, 2013.

[379] L Avery. Food transport in the c. elegans pharynx, 2003.

[380] Hasan Celiker and Jeff Gore. Clustering in community structure across replicate ecosystems following a long-term bacterial evolution experiment. *Nat. Commun.*, 5:4643, August 2014.

[381] B Sinervo and C M Lively. The rock–paper–scissors game and the evolution of alternative male strategies. *Nature*, 380(6571):240–243, 1996.

[382] Naoki Narisawa, Shin Haruta, Hiroyuki Arai, Masaharu Ishii, and Yasuo Igarashi. Coexistence of Antibiotic-Producing and Antibiotic-Sensitive bacteria in biofilms is mediated by resistant bacteria. *Applied and Environmental Microbiology*, 74(12):3887–3894, 2008.

[383] Santiago Soliveres, Anika Lehmann, Steffen Boch, Florian Altermatt, Francesco Carrara, Thomas W Crowther, Manuel Delgado-Baquerizo, Anne Kempel, Daniel S Maynard, Matthias C Rillig, Brajesh K Singh, Pankaj Trivedi, and Eric Allan. Intransitive competition is common across five major taxonomic groups and is driven by productivity, competitive rank and functional traits. *Journal of Ecology*, 106(3):852–864, 2018.

[384] Emily R Troemel, Stephanie W Chu, Valerie Reinke, Siu Sylvia Lee, Frederick M Ausubel, and Dennis H Kim. p38 MAPK regulates expression of immune response genes and contributes to longevity in c. elegans, 2005.

[385] Michael TeKippe and Alejandro Aballay. C. elegans germline-deficient mutants respond to pathogen infection using shared and distinct mechanisms. *PLoS One*, 5(7):e11777, July 2010.

[386] M J Beanan and S Strome. Characterization of a germ-line proliferation mutation in c. elegans. *Development*, 116(3):755–766, November 1992.

[387] Jarrett F Lebov, Brandon H Schlomann, Catherine D Robinson, and Brendan J M Bohannan. Phenotypic parallelism during experimental adaptation of a Free-Living bacterium to the zebrafish gut. *MBio*, 11(4), August 2020.

[388] Norma M Morella, Francis Cheng-Hsuan Weng, Pierre M Joubert, C Jessica E Metcalf, Steven Lindow, and Britt Koskella. Successive passaging of a plant-associated microbiome reveals robust habitat and host genotype-dependent selection. *Proceedings of the National Academy of Sciences*, 117(2):1148–1159, January 2020.

[389] Waldan K Kwong and Nancy A Moran. Gut microbial communities of social bees. *Nat. Rev. Microbiol.*, 14(6):374–384, June 2016.

[390] Philipp Dirksen, Adrien Assié, Johannes Zimmermann, Fan Zhang, Adina-Malin Tietje, Sarah Arnaud Marsh, Marie-Anne Félix, Michael Shapira, Christoph Kaleta, Hinrich Schulenburg, and Buck S Samuel. CeMbio - the microbiome resource. *G3*, 10(9):3025–3039, September 2020.

[391] Christoph Ratzke, Julien Barrere, and Jeff Gore. Strength of species interactions determines biodiversity and stability in microbial communities. *Nat Ecol Evol*, 4(3):376–383, March 2020.

[392] Christoph Ratzke, Jonas Denk, and Jeff Gore. Ecological suicide in microbes. *Nat Ecol Evol*, 2(5):867–872, May 2018.

[393] Benjamin Obadia, Z T Güvener, Vivian Zhang, Javier A Ceja-Navarro, Eoin L Brodie, William W Ja, and William B Ludington. Probabilistic invasion underlies natural gut microbiome stability, 2017.

[394] Matthew Jemielita, Michael J Taormina, Adam R Burns, Jennifer S Hampton, Annah S Rolig, Karen Guillemin, and Raghuveer Parthasarathy. Spatial and temporal features of the growth of a bacterial species colonizing the zebrafish gut. *MBio*, 5(6), December 2014.

[395] Brandon H Schlomann, Travis J Wiles, Elena S Wall, Karen Guillemin, and Raghuveer Parthasarathy. Bacterial cohesion predicts spatial distribution in the larval zebrafish intestine. *Biophys. J.*, 115(11):2271–2277, December 2018.

[396] Theresa Stiernagle. Maintenance of c. elegans. *WormBook*, pages 1–11, February 1999.

[397] Sihui Zhang, Diya Banerjee, and Jeffrey R Kuhn. Isolation and culture of larval cells from c. elegans. *PloS One*, 6(4):e19505, 2011.

[398] M A Larkin, G Blackshields, N P Brown, R Chenna, P A McGettigan, H McWilliam, F Valentin, I M Wallace, A Wilm, R Lopez, J D Thompson, T J Gibson, and D G Higgins. Clustal W and clustal X version 2.0. *Bioinformatics*, 23(21):2947–2948, 2007.

[399] Vincent Lefort, Jean-Emmanuel Longueville, and Olivier Gascuel. SMS: Smart model selection in PhyML. *Mol. Biol. Evol.*, 34(9):2422–2424, September 2017.

[400] Jared Kehe, Anthony Ortiz, Anthony Kulesa, Jeff Gore, Paul Blainey, and Jonathan Friedman. Positive interactions are common among culturable bacteria. *bioRxiv*, 2020.

[401] John F Bruno, John J Stachowicz, and Mark D Bertness. Inclusion of facilitation into ecological theory. *Trends in ecology & evolution*, 18(3):119–125, 2003.

[402] Michael Rothballer, Barbara Eckert, Michael Schmid, Agnes Fekete, Michael Schloter, Angelika Lehner, Stephan Pollmann, and Anton Hartmann. Endophytic root colonization of gramineous plants by herbaspirillum frisingense. *FEMS microbiology ecology*, 66(1):85–95, 2008.

[403] Daniel Straub, Michael Rothballer, Anton Hartmann, and Uwe Ludewig. The genome of the endophytic bacterium h. frisingense GSF30T identifies diverse strategies in the herbaspirillum genus to interact with plants. *Front. Microbiol.*, 4, 2013.

[404] Joshua Schimel and Sean Michael Schaeffer. Microbial control over carbon cycling in soil. *Frontiers in microbiology*, 3:348, 2012.

[405] Kevin R Arrigo. Marine microorganisms and global nutrient cycles. *Nature*, 437(7057):349–355, 2005.

[406] Richard D Bardgett, Chris Freeman, and Nicholas J Ostle. Microbial contributions to climate change through carbon cycle feedbacks. *The ISME Journal*, 2(8):805–814, 2008.

[407] Jizhong Zhou, Kai Xue, Jianping Xie, Ye Deng, Liyou Wu, Xiaoli Cheng, Shenfeng Fei, Shiping Deng, Zhili He, Joy D Van Nostrand, et al. Microbial mediation of carbon-cycle feedbacks to climate warming. *Nature Climate Change*, 2(2):106–110, 2012.

[408] David M Weller, Jos M Raaijmakers, Brian B McSpadden Gardener, and Linda S Thomashow. Microbial populations responsible for specific soil suppressiveness to plant pathogens. *Annual review of phytopathology*, 40(1):309–348, 2002.

[409] Rodrigo Mendes, Marco Kruijt, Irene De Bruijn, Ester Dekkers, Menno van der Voort, Johannes HM Schneider, Yvette M Piceno, Todd Z DeSantis, Gary L Andersen, Peter AHM Bakker, et al. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. *Science*, 332(6033):1097–1100, 2011.

[410] Rifat Hayat, Safdar Ali, Ummay Amara, Rabia Khalid, and Iftikhar Ahmed. Soil beneficial bacteria and their role in plant growth promotion: a review. *Annals of microbiology*, 60(4):579–598, 2010.

[411] Susan V Lynch and Oluf Pedersen. The human intestinal microbiome in health and disease. *New England Journal of Medicine*, 375(24):2369–2379, 2016.

[412] Tim A Hoek, Kevin Axelrod, Tommaso Biancalani, Eugene A Yurtsev, Jinghui Liu, and Jeff Gore. Resource availability modulates the cooperative and competitive nature of a microbial cross-feeding mutualism. *PLoS Biology*, 14(8):e1002540, 2016.

[413] Niels Klitgord and Daniel Segrè. Environments that induce synthetic microbial ecosystems. *PLoS Computational Biology*, 6(11):e1001002, 2010.

[414] Stefanía Magnúsdóttir, Almut Heinken, Laura Kutt, Dmitry A Ravcheev, Eugen Bauer, Alberto Noronha, Kacy Greenhalgh, Christian Jäger, Joanna Baginska, Paul Wilmes, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature biotechnology*, 35(1):81–89, 2017.

[415] Alan R Pacheco, Mauricio Moel, and Daniel Segrè. Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nature communications*, 10(1):1–12, 2019.

[416] J Jeffrey Morris, Richard E Lenski, and Erik R Zinser. The black queen hypothesis: evolution of dependencies through adaptive gene loss. *MBio*, 3(2), 2012.

[417] Robert Marsland, Wenping Cui, and Pankaj Mehta. A minimal model for microbial biodiversity can reproduce experimentally observed ecological patterns. *Scientific reports*, 10(1):1–17, 2020.

[418] Katharine Z Coyte, Jonas Schluter, and Kevin R Foster. The ecology of the microbiome: networks, competition, and stability. *Science*, 350(6261):663–666, 2015.

[419] Guy Bunin. Ecological communities with lotka-volterra dynamics. *Physical Review E*, 95(4):042414, 2017.

[420] Monu Jariyal, Vikas Jindal, Kousik Mandal, Virash Kamal Gupta, and Balwinder Singh. Bioremediation of organophosphorus pesticide phorate in soil by microbial consortia. *Ecotoxicology and environmental safety*, 159:310–316, 2018.

[421] M Kavino and SK Manoranjitham. In vitro bacterization of banana (musa spp.) with native endophytic and rhizospheric bacterial isolates: novel ways to combat fusarium wilt. *European journal of plant pathology*, 151(2):371–387, 2018.

[422] Bret D Wallace and Matthew R Redinbo. The human microbiome is a source of therapeutic drug targets. *Current opinion in chemical biology*, 17(3):379–384, 2013.

[423] Sophie Weiss, Will Van Treuren, Catherine Lozupone, Karoline Faust, Jonathan Friedman, Ye Deng, Li Charlie Xia, Zhenjiang Zech Xu, Luke Ursell, Eric J Alm, et al. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *The ISME Journal*, 10(7):1669–1681, 2016.

[424] Hong-Tai Cao, Travis E Gibson, Amir Bashan, and Yang-Yu Liu. Inferring human microbial dynamics from temporal metagenomics data: Pitfalls and lessons. *BioEssays*, 39(2):1600188, 2017.

[425] Cheri M Ackerman, Cameron Myhrvold, Sri Gowtham Thakku, Catherine A Freije, Hayden C Metsky, David K Yang, H Ye Simon, Chloe K Boehm, Tinna-Sólveig F Kosoko-Thoroddsen, Jared Kehe, et al. Massively multiplexed nucleic acid detection with cas13. *Nature*, 582(7811):277–282, 2020.

[426] Aleksej Zelezniak, Sergej Andrejev, Olga Ponomarova, Daniel R Mende, Peer Bork, and Kiran Raosaheb Patil. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences*, 112(20):6449–6454, 2015.

[427] Yasuhiro Tanaka, Satoshi Hanada, Akira Manome, Takayasu Tsuchida, Ryuichiro Kurane, Kazunori Nakamura, and Yoichi Kamagata. Catellibacterium nectariphilum gen. nov., sp. nov., which requires a diffusible compound from a strain related to the genus sphingomonas for vigorous growth. *International journal of systematic and evolutionary microbiology*, 54(3):955–959, 2004.

[428] Kevin Gross. Positive interactions among competitors can produce species-rich communities. *Ecology letters*, 11(9):929–936, 2008.

[429] Roberto Kolter, Deborah A Siegele, and Antonio Tormo. The stationary phase of the bacterial life cycle. *Annual review of microbiology*, 47:855–875, 1993.

[430] Markus Basan, Sheng Hui, Hiroyuki Okano, Zhongge Zhang, Yang Shen, James R Williamson, and Terence Hwa. Overflow metabolism in escherichia coli results from efficient proteome allocation. *Nature*, 528(7580):99–104, 2015.

[431] Karsten Zengler and Livia S Zaramela. The social network of microorganisms—how auxotrophies shape complex communities. *Nature Reviews Microbiology*, 16(6):383–390, 2018.

[432] Rudolf O Schlechter, Hyunwoo Jun, Michał Bernach, Simisola Oso, Erica Boyd, Dian A Muñoz-Lintz, Renwick CJ Dobson, Daniela M Remus, and Mitja NP Remus-Emsermann. Chromatic bacteria–a broad host-range plasmid and chromosomal insertion toolbox for fluorescent protein expression in bacteria. *Frontiers in microbiology*, 9:3052, 2018.

[433] Andrew W Brooks, Kevin D Kohl, Robert M Brucker, Edward J van Opstal, and Seth R Bordenstein. Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biology*, 14(11):e2000225, 2016.

[434] Nittay Meroz, Nesli Tovi, Yael Sorokin, and Jonathan Friedman. Community composition of microbial microcosms follows simple assembly rules at evolutionary timescales. *Nature communications*, 12(1):1–9, 2021.

[435] Pablo Vinuesa, Luz E Ochoa-Sánchez, and Bruno Contreras-Moreira. Get_phylomarkers, a software package to select optimal orthologous clusters for phylogenomics and inferring pan-genome phylogenies, used for a critical geno-taxonomic revision of the genus stenotrophomonas. *Frontiers in microbiology*, 9:771, 2018.

[436] James Q Boedicker, Liang Li, Timothy R Kline, and Rustem F Ismagilov. Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics. *Lab Chip*, 8(8):1265–1272, August 2008.

[437] Stanislav S Terekhov, Ivan V Smirnov, Maja V Malakhova, Andrei E Samoilov, Alexander I Manolov, Anton S Nazarov, Dmitry V Danilov, Svetlana A Dubiley, Ilya A Osterman, Maria P Rubtsova, Elena S Kostryukova, Rustam H Ziganshin, Maria A Kornienko, Anna A Vanyushkina, Olga N Bukato, Elena N Ilina, Valentin V Vlasov, Konstantin V Severinov, Alexander G Gabibov, and Sidney Altman. Ultrahigh-throughput functional profiling of microbiota communities. *Proceedings of the National Academy of Sciences*, 115(38):9551–9556, September 2018.

[438] Hasan Celiker and Jeff Gore. Cellular cooperation: insights from microbes. *Trends Cell Biol.*, 23(1):9–15, 2013.

[439] Alexander Eng and Elhanan Borenstein. An algorithm for designing minimal microbial communities with desired metabolic capacities. *Bioinformatics*, 32(13):2008–2016, July 2016.

[440] Nathan I Johns, Tomasz Blazejewski, Antonio Lc Gomes, and Harris H Wang. Principles for designing synthetic microbial communities. *Curr. Opin. Microbiol.*, 31:146–153, June 2016.

[441] Erik F Y Hom and Andrew W Murray. Plant-fungal ecology. niche engineering demonstrates a latent capacity for fungal-algal mutualism. *Science*, 345(6192):94–98, July 2014.

[442] Bo Hu, Jin Du, Rui-Yang Zou, and Ying-Jin Yuan. An environment-sensitive synthetic microbial ecosystem. *PLoS One*, 5(5):e10619, May 2010.

[443] Thomas J Borody and Alexander Khoruts. Fecal microbiota transplantation and emerging applications. *Nat. Rev. Gastroenterol. Hepatol.*, 9(2):88–96, December 2011.

[444] Jie Hu, Zhong Wei, Ville-Petri Friman, Shao-Hua Gu, Xiao-Fang Wang, Nico Eisenhauer, Tian-Jie Yang, Jing Ma, Qi-Rong Shen, Yang-Chun Xu, and Alexandre Jousset. Probiotic diversity enhances rhizosphere microbiome function and plant disease suppression. *MBio*, 7(6), December 2016.

[445] P N Bhattacharyya and D K Jha. Plant growth-promoting rhizobacteria (PGPR): emergence in agriculture. *World J. Microbiol. Biotechnol.*, 28(4):1327–1350, April 2012.

[446] E Benizri, E Baudoin, and A Guckert. Root colonization by inoculated plant Growth-Promoting rhizobacteria. *Biocontrol Sci. Technol.*, 11(5):557–574, 2001.

[447] Trevor D Lawley, Simon Clare, Alan W Walker, Mark D Stares, Thomas R Connor, Claire Raisen, David Goulding, Roland Rad, Fernanda Schreiber, Cordelia Brandt, Laura J Deakin, Derek J Pickard, Sylvia H Duncan, Harry J Flint, Taane G Clark, Julian Parkhill, and Gordon Dougan. Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing clostridium difficile disease in mice. *PLoS Pathogens*, 8(10):e1002995, October 2012.

[448] Sandrine Brugiroux, Markus Beutler, Carina Pfann, Debora Garzetti, Hans-Joachim Ruscheweyh, Diana Ring, Manuel Diehl, Simone Herp, Yvonne Lötscher, Saib Hussain, Boyke Bunk, Rüdiger Pukall, Daniel H Huson, Philipp C Münch, Alice C McHardy, Kathy D McCoy, Andrew J Macpherson, Alexander Loy, Thomas Clavel, David Berry, and Bärbel Stecher. Genome-guided design of a defined mouse microbiota that confers colonization resistance against salmonella enterica serovar typhimurium. *Nat Microbiol*, 2:16215, November 2016.

[449] Silvia Caballero, Sohn Kim, Rebecca A Carter, Ingrid M Leiner, Bože Sušac, Liza Miller, Grace J Kim, Lilan Ling, and Eric G Pamer. Cooperating commensals restore colonization resistance to Vancomycin-Resistant enterococcus faecium. *Cell Host Microbe*, 21(5):592–602.e4, May 2017.

[450] Muya Shu, Yanhan Wang, Jinghua Yu, Sherwin Kuo, Alvin Coda, Yong Jiang, Richard L Gallo, and Chun-Ming Huang. Fermentation of propionibacterium acnes, a commensal bacterium in the human skin microbiome, as skin probiotics against Methicillin-Resistant staphylococcus aureus. *PLoS One*, 8(2):e55380, 2013.

[451] Mohammad R Seyedsayamdost, Matthew F Traxler, Jon Clardy, and Roberto Kolter. Old meets new: using interspecies interactions to detect secondary metabolite production in actinomycetes. *Methods Enzymol.*, 517:89–109, 2012.

[452] Jacopo Grilli, György Barabás, Matthew J Michalska-Smith, and Stefano Allesina. Higher-order interactions stabilize dynamics in competitive network models. *Nature*, 548(7666):210–213, 2017.

[453] Felix Sommer, Jacqueline Moltzau Anderson, Richa Bharti, Jeroen Raes, and Philip Rosenstiel. The resilience of the intestinal microbiota influences health and disease. *Nature Reviews Microbiology*, 15(10):630–638, 2017.

[454] Clare Abreu, Anthony Ortiz Lopez, and Jeff Gore. Pairing off: a bottom-up approach to the human gut microbiome. *Molecular systems biology*, 14(6):e8425, 2018.