Host-microbiome interactions impacting pathogen and mutualist colonization within defensive symbioses

By

Jennifer R. Bratburd

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Microbiology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2020

Date of final oral examination: 08/04/2020

The dissertation is approved by the following members of the Final Oral Committee:

Cameron R. Currie, Professor, Bacteriology

Federico E. Rey, Associate Professor, Bacteriology

David R. Andes, Professor, Medical Microbiology and Immunology

Caitlin S. Pepperell, Associate Professor, Medical Microbiology and Immunology

Lingjun Li, Professor, Pharmacy

Dissertation Abstract

Host-microbiome interactions impacting microbial colonization in defensive symbioses

By

Jennifer R. Bratburd

Under the supervision of Professor Cameron R. Currie

at the University of Wisconsin-Madison

Microbial interactions shape the world around us. One major determinant of the effect a microbe will have on its environment is the microbe's ability to colonize. For pathogens, colonization directly impacts the host's health and many hosts have mechanisms to limit or otherwise control microbial colonization. These limitations may also prove challenging for commensal or mutualistic microbes, which themselves may be critical many aspects of host health, including defending the host against pathogens. In this dissertation, I explore a spectrum of host-microbiome interactions, ranging from an individual mutualistic bacterial strains of the fungus-growing ant system to whole human gut microbial communities using several approaches to better understand defensive mutualisms. In Chapter 1, I discuss how understanding defensive symbiosis of social animal models, in particular insect systems, may help in understanding with human problems with controlling pathogens in large social populations. In Chapter 2, I present experimental colonization data and comparative genomics that suggests the lack of specificity from the symbiont in the fungus-growing ant and *Pseudonocardia* mutualism. To explore

pathogen interactions with a more complex microbial community, in Chapter 3, I investigate how human gut microbial community responds to infection of the host in a gnotobiotic mouse model with metagenomics and metabolomics approaches. I contrast host without microbiota and hosts with microbiota but no infection to find that infection greatly perturbs the communities and I find particular metabolites in abundance on in the presence of both microbial community and pathogen. Expanding on human gut microbiome and germ-free mouse model approach, in Chapter 4, I use human stool samples as donors to inoculate germ-free in order to identify variability in the microbiome resistance to infection and apply metagenomic techniques to examine commonalities of resistant microbiomes. I find limited evidence of shared taxonomic groups in resistant microbes, but some indication of shared functional genes in the metagenomes associated with pathogen resistance. Together, these approaches provide insight into the complexity of host interplay with defensive microbes.

Acknowledgements

I am grateful and hugely thankful for the community that has supported me and contributed to my success on the path to obtaining a PhD. So many people have enabled amazing opportunities and I cannot describe joy and enrichment I have had with my wonderful community.

This work would not be possible without the funding by the Microbes in Health and Disease Training Grant, the Michael and Winona Foster Bacteriology Fellowship, the National Institutes of Health with the ICBG and CETR.

I am very appreciative of my advisor, Dr. Cameron Currie who has built an extraordinary network of talented students and post-docs. I am thankful for Cameron's expertise, for giving me the freedom to exploring new areas of research. I am indebted to the guidance of many of my labmates who were always willing to help, provided thoughtful feedback, ideas, discussion and encouragement. I am inspired by students who graduated: Bradon, Gina, Lily, Eric, Heidi, and Marc, as well as the hard-working post-docs: Reed, Robb, Adam, Camila, Margaret, Huan, Hongjie, Weilan and Dan. I have enjoyed the enthusiasm of current and former students: Donny, Soleil, Kirsten, JoJo, Alex Mundy, and Alex Cheong, and Charlotte (who I owe yet more gratitude for taking care of me when I became violently ill in Costa Rica). I am thankful for Caitlin Carlson who made the lab run smoothly, along with Laura, Evelyn, and Julian. I am thankful for the many undergraduates I have the opportunity to interact with, especially KM Barnett and Lexis Wedell who I mentored and were incredibly positive workers. I am also grateful to have worked with Rachel Arango and interacted with many other Currie-lab collaborators.

The Microbiology Doctoral Training Program and the community associated have been incredibly helpful over the years. I am grateful for our coordinators, Cathy and Terra, as well the administrative staff including Tracy, Kari, Kim, and Janet. My cohort has been an amazing source of support during the PhD process, and helping with ideas, practicing talks, and improving resumes.

My committee members have provided helpful feedback over the years, and I am very appreciative of their key insights, which shaped the direction of my research and enabled collaborations.

I am extremely appreciative of my collaborators. The Rey lab has been incredibly helpful over the years and I have enjoyed interacting with them and learning from their expertise. I am also grateful especially for the mice themselves. I am also very thankful for the expertise of the Li lab, especially Caitlin and Jericha. I am very appreciative of the time I had to work with the Pinto lab, and especially for help in the hard labor of digging up ants out of the ground or occasionally out of trees. Likewise, I am appreciative of the staff of La Selva for making field work possible. I am also grateful for the opportunity to work with Darin Wiesner of the Klein lab.

I am very glad to have had the opportunity to work with many groups on campus, especially Catalysts for Science Policy, where I had enjoyed the opportunity to write and publish policy

memos and organize and lead events. I am also grateful for the funding to attend conferences with CaSP.

Finally, I would not have made it this far without the support of my family and friends, who have encouraged me all along the way. In particular I thank my parents, my sister, my step-family, and Dr. Kelsey Florek and Dr. Kurt Throckmorton who I do not have the words to express my love and appreciation.

Table of Contents

Chapter 1: Defensive Symbioses in Social Insects Can Inform Human Health and Agriculture

Jennifer R. Bratburd, Rachel A. Arango, Heidi A. Horn

Reprinted from Bratburd JR, Arango RA and Horn HA (2020) Defensive Symbioses in Social Insects Can Inform Human Health and Agriculture. Front. Microbiol. 11:76. doi: 10.3389/fmicb.2020.00076

1.1 Abstract

Social animals are among the most successful organisms on the planet and derive many benefits from living in groups, including facilitating the evolution of agriculture. However, living in groups increases the risk of disease transmission in social animals themselves and the cultivated crops upon which they obligately depend. Social insects offer an interesting model to compare to human societies, in terms of how insects manage disease within their societies and with their agricultural symbionts. As living in large groups can help the spread of beneficial microbes as well as pathogens, we examine the role of defensive microbial symbionts in protecting the host from pathogens. We further explore how beneficial microbes may influence other pathogen defenses including behavioral and immune responses, and how we can use insect systems as models to inform on issues relating to human health and agriculture.

1.2 Introduction

Some of the most successful species on the planet in terms of number of species generated over time, ability to inhabit diverse ecosystems, and maintenance of high population densities are social animals (Wilson, 1987). Social lifestyles, however, come at the cost of increased exposure to pathogens. Both modeling and experimental results indicate that population size and density correlate with pathogen prevalence and diversity (Anderson and

May, 1979, 1982; Altizer et al., 2003; Schmid-Hempel, 2017). The 10-fold expansion of the human population in the last 200 years with similar population density increases has caused concerns around the risk of spreading infectious diseases (Cohen, 2003). Social insects have faced the same challenges successfully, maintaining high population densities over millions of years and are simple models to gain a better understanding of how to mitigate pathogen burden and spread (Figure 1).

While social living may enhance pathogen spread, social living also enables the spread of beneficial microbes (Biedermann and Rohlfs, 2017). For instance, after termites molt, they must replace their gut symbionts from other nest mates through trophallaxis and coprophagy. This "social gut" is suggested to contribute to nestmate recognition as well as development, nutrition, and defense (Breznak and Brune, 1994; Matsuura, 2001; Nakashima et al., 2002; Adams and Boopathy, 2005). Many microbes benefit the host by providing protection against predators, parasites, pathogens, or environmental stresses, also known as defensive symbiosis (White and Torres, 2009). In a mutualistic relationship, the host provides shelter and/or nutrients in exchange for defense. Understanding interactions between hosts, pathogens, and beneficial microbes can inform on the potential use of beneficial symbionts in systematically targeting certain pathogens.

In interactions between social animals, their microbial defensive symbionts and pathogens, many different selective pressures may be operating simultaneously. Pathogen pressures can impact host and symbiont (King and Bonsall, 2017; Engl et al., 2018). Beneficial symbionts may influence social behavior to facilitate their horizontal transmission, but core microbiota may be influenced by diet or other factors (Sherwin et al., 2019). The evolutionary and ecological dynamics of microbial symbiont relationships with social animals are not well understood. To deconvolute these interactions, social insects are interesting models to compare

social and solitary relatives (e.g., bees, discussed below) or comparing changes in microbiota of species that alternate between gregarious and solitary lifestages may also be useful (Lavy et al., 2018).

In this review, we discuss the role of microbial defensive symbionts in pathogen mitigation within social communities and their associated agricultural systems. We also consider how defensive symbiosis intersects with immunological and behavioral defenses. We compare examples from insects with defensive symbionts in humans and highlight how insect models can advance understanding the social impacts of defensive symbionts.

1.3 Insect Defenses Against Pathogens

While defensive symbionts can benefit both social and solitary animals, social living may better enable sharing defensive symbionts than solitary lifestyles. For example, eusocial bees (e.g., *Apis mellifera* and *Bombus* spp.), have a consistent core microbiota that defends against the trypanosome gut parasite *Crithidia bombi*, whereas solitary bees do not have a consistent core community (Koch and Schmid-Hempel, 2011). Several core microbiome members, including *Gilliamella apicola* and *Lactobacillus* spp., correlate with decreased susceptibility to *C. bombi* (Cariveau et al., 2014; Mockler et al., 2018; Näpflin and Schmid-Hempel, 2018). Additionally, experiments disrupting the core bee microbiota support the hypothesis that the gut microbiota plays a role in protecting against opportunistic pathogens (Raymann et al., 2017) and another common parasite, *Lotmaria passim* (Schwarz et al., 2016). Biofilm formation by the core strains is the suggested protective mechanism against this pathogen, as indicated by fluorescent *in situ* hybridization (FISH) imaging (Martinson et al., 2012) and the enrichment of secretion systems and surface proteins in bee gut metagenomes (Engel et al., 2012). As biofilm formation and colonization resistance are broad defensive mechanisms, it is unclear whether solitary bees have

microbes with similar functionality. Likewise, social bee gut microbes may confer other functions affecting fitness.

Social animals need to not only protect themselves from disease, but also their shared food sources. Three lineages of eusocial or subsocial insects demonstrate agricultural behavior: ants (Myrmicinae: Attini), termites (Macrotermitinae), and ambrosia beetles (Xyleborinae and others). All of these insects live in gregarious communities supporting the hypothesis that sociality allowed for evolution of insect agriculture (Mueller et al., 2005). Fungus farming termites cultivate basidiomycete fungi, *Termitomyces* spp. as a food source that are either vertically or horizontally acquired depending on termite species (Johnson and Hagen, 1981; Korb and Aanen, 2003). Some termites (*Macrotermes natalensis*) harbor *Bacillus* sp. that produce bacillaene which has antifungal activity and helps protect the fungal cultivar (Um et al., 2013). Xyleborine ambrosia beetles cultivate an assemblage of fungi, rather than a single fungal cultivar, which comprises mycelial fungi, yeasts, and bacteria (Norris, 1965; Hulcr and Stelinski, 2017). A cycloheximide-producing *Streptomyces* phylotype has been isolated from two species of ambrosia beetles as a possible defensive symbiont (Grubbs et al., 2019).

In the fungus-growing ants, microbial associations range from mutualistic to parasitic and are well-described. The ants grow a fungal cultivar as their primary food source in a monoculture, which makes it highly susceptible to the specialized fungal pathogen *Escovopsis* (Ascomycete; Hypocreales). To protect their food source, the ants evolved several defense mechanisms, including a mutualism with *Pseudonocardia* spp. (Currie et al., 1999b, 2003). *Pseudonocardia* produces antimicrobial molecules that are active against *Escovopsis* (Currie et al., 1999b, 2003; Poulsen et al., 2010). Growing *Pseudonocardia* and *Escovopsis* together reveals patterns of inhibition and resistance between the two organisms suggesting population

and interaction dynamics at fine phylogenetic scales (Poulsen et al., 2010; Cafaro et al., 2011). Several of the antibiotics produced by *Pseudonocardia* have been characterized (Oh et al., 2009; Carr et al., 2012; Van Arnam et al., 2016) although the full diversity of antibiotics used is unknown.

1.4 Interactions of Defensive Symbionts With Host Defenses in Insects

Other methods of pathogen resistance, such as behavior and immunity, aid in disease resistance and can be influenced by microbes (Nyholm and Graf, 2012; Lizé et al., 2014; Flórez et al., 2015). Host and symbionts may adapt to each other in different ways: symbionts may avoid triggering immune function (Trappeniers et al., 2019); hosts may diversify immune pathways (Maire et al., 2019) or hosts may potentially reduce immune function (International Aphid Genomics Consortium, 2010; Douglas et al., 2011). Further examples of innate immunity in social insects can be found in the following review (Otani et al., 2016).

Social insects can coordinate defensive behaviors, some of which may be triggered or helped by beneficial microbes. Many of the defensive behaviors in social insects are aimed at maintaining sanitation of the nest as well as the individuals within the nest. This phenomenon of collective actions to mitigate pathogen spread/exposure is known as social immunity, which is defined as the control or elimination of potential pathogens by cooperation of individuals through behavioral, physiological, and/or organizational means (Cremer et al., 2007; Meunier, 2015). For example, subsocial aphid *Nipponaphis monzeni* soldiers respond to attacks on their colonies by swarming and exploding their abdomens. Their abdomens are swollen with hemocytes and tyrosine that seal and protect the colony. The endosymbiotic bacterium, *Buchnera*, regulated by aphid host genes, helps overproduce tyrosine (Kutsukake et al., 2019). This example highlights the complex interplay occurring between host, beneficial symbionts,

immune system, and social structure of an organism. Other examples of social immunity include grooming, removing waste material and weeding nests and fungal gardens. Further experimentation using antibiotics or probiotics could explore the manner in which microbes may influence behavior and fitness (Alberoni et al., 2018).

Defensive behaviors can also be facilitated by the microbial production of chemical signals or chemical defenses. Social insects participate in extensive grooming behaviors categorized as autogrooming (i.e., self-grooming) and allogrooming (i.e., grooming among nestmates), which serve not only to remove foreign substances from the body surface, but can also provide lasting antimicrobial defenses (Zhukovskaya et al., 2013). In terms of using microbes for production of chemical defenses, many examples in the above defensive symbioses fit this description (e.g., antimicrobial phenols from locust symbionts, antibiotics from fungus-farming ant symbionts). Microbes are also capable of producing chemical signals, such as the intestinal microbes of subterranean termites (*Reticulitermes speratus*), which allow recognition of nestmates from non-nestmate intruders (Matsuura, 2001). The diversity of interactions between defensive microbes and host behavior remains an open area of exploration.

1.5 Human Defenses Against Pathogens

As in insects, the microbiota provides defense against various pathogens in humans, but is more complex than insect microbiomes. While different sites, such as the vagina and nasal cavity can support symbionts with abilities to produce defensive compounds (Donia et al., 2014; Zipperer et al., 2016), most of the potential defensive microbes described reside in the gut. Unlike many insect gut microbiotas, the human gut microbiota may contain hundreds of species (Qin et al., 2010). Adding further complication, whereas in bees and other hosts a core community is evident, a consistent core community has not been identified in humans, although

a core functionality appears more conserved than particular strains (Turnbaugh and Gordon, 2009; Human Microbiome Project Consortium, 2012). Although humans lack an equivalent solitary lifestyle to insects, evidence suggests that humans in close social relationships may share a variety of bacteria with one another and have greater richness and diversity than humans living alone (Dill-McFarland et al., 2019).

Many different mechanisms for microbial defense exist and understanding the microbiota's functions may lead to improved therapies. For example, fecal microbiota transplants for treating *Clostridium difficile* infections that are non-responsive to antibiotics have cure rates of 90% (Bakken et al., 2011; Youngster et al., 2016). Several mechanisms have been suggested including that the microbiota outcompete the pathogen for nutrients, microbially produced antibiotics target *C. difficile*, microbially produced secondary bile acids inhibit *C. difficile*, and microbial interactions with the immune system help repair the gut barrier (Khoruts and Sadowsky, 2016). Human gut microbes have also been linked to defense against *Vibrio cholerae*, where correlations have been found between microbiota taxa present in the gut and resistance to cholera (Hsiao et al., 2014; Midani et al., 2018). Likewise, human microbiota strains compete with *Salmonella* for nutrients and produce metabolites that potentially inhibit *Salmonella* (Antunes et al., 2014; Bratburd et al., 2018; Zhang et al., 2018). Although many interactions and correlations have been suggested between defensive symbiotic bacteria and pathogens in humans, the challenge remains to explore these symbionts on a society-wide scale to understand the benefits not only to individuals but to public health.

Although humans do not have ancient history (on an evolutionary time scale) with agriculture, many crops used by humans associate with defensive microbes against certain pathogens. One example of an agricultural defensive symbiont is *Pseudomonas fluorescens*, a

bacterium that produces the antibiotic 2,4-diacetylphloroglucinol, which can inhibit the causative agent of take-all disease in wheat (Keel et al., 1992). This bacterium can be found naturally in soils and is a prominent example of suppressive soils, where soil harbors a community or certain strains that inhibit plant pathogens, analogous to the idea of colonization resistance in animals. Beneficial microbes may provide an environmentally sustainable alternative to chemical control of pathogens and vectors, but will require maintaining beneficial microbes in agricultural settings and consideration of microbial interactions in plant breeding beyond the host's pathogen resistance (see the following review for more detail (Syed Ab Rahman et al., 2018).

1.6 Interactions of Defensive Symbionts With Host Defenses in Humans

The role of the immune system and behaviors is increasingly recognized as not only defending against harmful microbes, but also fostering the establishment and maintenance of bacterial symbionts. We direct the reader to other reviews for further exploration of the numerous interactions between the microbiota and the immune system (Belkaid and Harrison, 2017) and behavior (Vuong et al., 2017; Johnson and Foster, 2018).

Humans have been practicing their own social immunity with hygienic behaviors throughout history. This includes early ritualistic behaviors, quarantine and sanitation, and after the rise of the germ-theory of disease, water treatment, vaccinations, and vector control (Institute of Medicine (US) Committee for the Study of the Future of Public Health, 1988; Curtis, 2007). While humans have taken advantage of antimicrobial compounds from a variety of sources for hundreds of years (Aminov, 2010; Harrison et al., 2015), large scale antibiotic discovery, often microbially derived, took off in the 1900's and enabled treating a wide variety of pathogens in people as well as in agriculture (Aminov, 2010). Unfortunately, broad-spectrum antibiotics can have lasting impacts on the microbiota affecting the many interactions discussed above (Jernberg

et al., 2007). While efforts to eliminate pathogens have substantial impacts, most notably with vaccines eliminating smallpox and reducing other disease to 99% fewer cases (Orenstein and Ahmed, 2017), practices for sharing beneficial microbes could also be valuable for medicine and agriculture. These practices may include fecal microbiota transplants, probiotic and prebiotic supplementation (George Kerry et al., 2018; Sonnenburg and Sonnenburg, 2019), creating built environments that favor beneficial microbes (Kembel et al., 2012); however, besides perhaps fecal microbiota transplants for treating *C. difficile*, these practices currently lack substantial evidence of efficacy.

1.7 What Can We Learn From Insects?

Insects are useful models to address societal-wide impacts of defensive symbionts (Table 1). Given the vast complexity in the human gut, insects can be a simple model to dissect various mechanisms of microbial defenses since insects tend to have simplified microbiomes relative to humans. Comparisons between social and solitary insects (whether in different life stages as described above with locusts, or among related social and solitary members as described with bees) can shed light on what roles, if any, defensive symbionts have played in the evolution of sociality. Insect colonies are well-defined social units for replication, tend to have limited within colony genetic variation, and can be reared in controlled conditions. The insects themselves often have relatively fast life cycles, which is useful for examining fitness and intergeneration effects defensive microbes may have. Social insects also engage in behaviors of interest, like farming. In the most direct sense, natural products from insect symbioses may be useful as leads for new antibiotics themselves (Stow and Beattie, 2008; Ramadhar et al., 2014; Chevrette et al., 2019) and insects have inherent practical value as many species are important pollinators or pests; however, we also want to highlight using insect models to explore the societal impact of gaining

or losing beneficial symbionts. We detailed many benefits of insect models above, but these models come with drawbacks. The simplicities of social insect models limit conclusions relevant for humans to basic ecological dynamics. Insect models lack many features that mediate host-microbe interactions in humans, including an adaptive immune system or complex nervous systems. While much microbiome research has focused on the impact to the individual host, social insects can be used to address basic ecological and evolutionary dynamics including (i) how resilient societies transmit beneficial microbes to other individuals; and (ii) the larger impact of beneficial microbes at the population level.

Social insect models can address how social animals maximize beneficial microbe transmission while minimizing pathogen spread. Disrupting transmission of beneficial microbes can render hosts more susceptible to disease (Bohnhoff et al., 1954; Currie et al., 1999a; Raymann et al., 2017). In some human societies, transmission and maintenance of microbes has changed dramatically with the introduction of antibiotics, hygiene practices, and diet changes (Bokulich et al., 2016; Vangay et al., 2018). Disruptions in microbiota transmission are hypothesized to have health impacts, including obesity (Principi and Esposito, 2016). In both social insects and humans we have limited understanding of how beneficial microbes are effectively transmitted. In the leaf-cutter ant system, we know that the defensive symbiont *Pseudonocardia* is generally vertically transmitted, acquired during a narrow time window (Marsh et al., 2014) and may use certain host structures (Li et al., 2018), but we do not know what limits bacterial acquisition to certain strains and microbial adaptations to the host. Analogously in humans, we know microbial acquisition begins at birth but the roles and extent of vertically versus horizontally acquired microbes is still debated (Ferretti et al., 2018; Korpela and de Vos, 2018; Moeller et al., 2018; Brito et al., 2019). One drawback of insect models is that

specific mechanisms enabling transmission and colonization of beneficial microbes likely differ considerably between insects and humans (e.g., coprophagy is normal behavior for all termite colony members, while fecal microbiota transplant in humans is a medical procedure for the sick). Similarly, humans may travel further and interact with other communities introducing complicated interactions that may not be captured with insect models. However, the defined social structures of eusocial insects may be useful for understanding and manipulating microbial transmission later in life. Reproductive queens have limited contact with other adult workers, for instance, and understanding when and how they share microbes with other castes could illuminate the social elements of microbial transmission (Otani et al., 2019). Microbiomes of distinct nest structures provide an interesting comparison to the idea of built environments (Sharma and Gilbert, 2018).

Additionally, social insect models may address how environmental perturbations such as diet or temperature change the overall community response to pathogens and illuminate fitness effects in different contexts. For example, different substrates used in leafcutter ant fungal gardens impacts overall colony survivorship (Khadempour et al., 2016). While some leafcutter ants associate with defensive symbionts as described above, others rely on their own chemical defenses (Fernández-Marín et al., 2009). The leafcutting ant model could be used to explore how resilient different defensive strategies (chemical or biological control) are to perturbations such as the availability of different substrates. Fisher et al. (2019) predict how other social insect characteristics (including degree of specialization and nest architecture) may enhance susceptibility or resilience to various climate perturbations. The relative simplicity of insect models could help test and reveal basic principles to understand how microbial defenses change in different contexts.

1.8 Conclusion

How societies effectively address risk of pathogen exposure is of increasing concern, especially as the human population size and density rises. Social insects provide a window to explore disease management on a society-wide scale. Increasingly, defensive symbionts are recognized for their valuable role in mitigating pathogens, in insects as well as in humans. Social insects can act as useful models to address the role of defensive symbionts in societies and their interactions with physiological, chemical, and behavioral defenses. Examples from insects provide insight for microbiome-based therapies and agricultural products, as well as help address basic questions on how beneficial microbes are transmitted, maintained, and perturbed in social animals.

1.9 Author Contributions

JB, HH, and RA wrote the manuscript. All authors contributed to the manuscript revision and approved the submitted version.

1.10 Funding

Support for this project was provided through National Institutes of Health (NIH) U19 Al109673 and NIH U19 TW009872. Funding for JB provided through the University of Wisconsin-Madison Department of Bacteriology Michael and Winona Foster fellowship and NIH T32 AI55397. Funding for HH was provided through the Department of Energy Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award DE-SC0018409. Funding for RA was provided through the USDA Forest Products Laboratory.

1.11 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

1.12 Acknowledgments

We thank Dr. Cameron Currie, Dr. Reed Stubbendieck, Dr. Margaret Thairu, and Dr. Marc Chevrette for their insightful comments and critical appraisal on the manuscript.

1.13 Tables and Figures

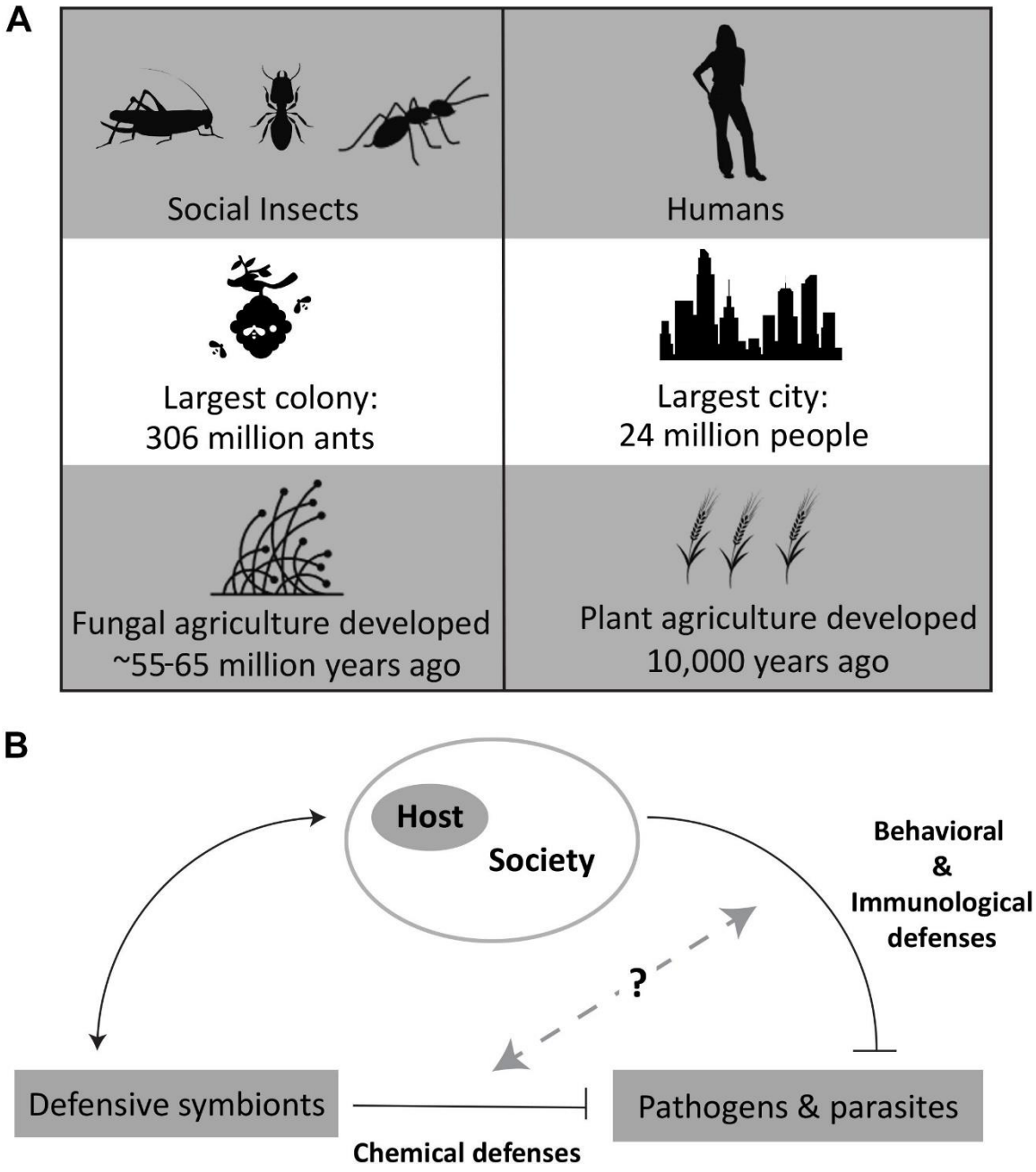


Figure 1. (A) Comparisons of human and insect societies, based on social grouping sizes (Burchill and Moreau, 2016; Sawe, 2018) and history with agriculture (Pringle, 1998; Schultz and Brady, 2008). **(B)** Overview of the relationship of defensive symbionts with host and pathogens. Specific image credit from the Noun Project (https://thenounproject.com/): Woman

by Lluisa Iborra, Locust by OCHA Visual, Termite by Heberti Almeida, Ant by Jacob Eckert, City by sumhi_icon, Beehive by Juraj Sedlák, Barley by Nathan Stang, and Fungi by CombineDesign. All images used and modified under the Creative Commons License, Attribution 3.0.

Advantages of insect models	Human alternatives
Control of variables (diet, environment, etc.)	Diets and environment generally not experimentally manipulated; metadata may be limited or subject to self-reporting inaccuracies
Defined units of replication for social group (e.g., one colony)	Units could be family, geographical region, etc.
Relatively simple microbiomes	Complex gut microbiomes, other sites varying complexity
Shorter life cycles	Long life cycles
Genetic variation within a colony lower than from a general population	Variable genetic variation
Lifestyle variation exists, including solitary, social, and eusocial members	Different types of social groupings, but all social

Table 1. Comparison of social insect and human models for defensive symbiosis.

1.14 References

Adams, L., and Boopathy, R. (2005). Isolation and characterization of enteric bacteria from the hindgut of Formosan termite. *Bioresour. Technol.* 96, 1592–1598. doi: 10.1016/J.BIORTECH.2004.12.020

Alberoni, D., Baffoni, L., Gaggìa, F., Ryan, P. M., Murphy, K., Ross, P. R., et al. (2018). Impact of beneficial bacteria supplementation on the gut microbiota, colony development and productivity of *Apis mellifera* L. *Benef. Microbes* 9, 269–278. doi: 10.3920/BM2017.0061

Altizer, S., Nunn, C. L., Thrall, P. H., Gittleman, J. L., Antonovics, J., Cunningham, A. A., et al. (2003). Social organization and parasite risk in mammals: integrating theory and empirical studies. *Annu. Rev. Ecol. Evol. Syst.* 34, 517–547. doi: 10.1146/annurev.ecolsys.34.030102.151725

Aminov, R. I. (2010). A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* 1:134. doi: 10.3389/fmicb.2010.00134

Anderson, R. M., and May, R. M. (1979). Population biology of infectious diseases: part I. *Nature* 280, 361–367. doi: 10.1038/280361a0

Anderson, R. M., and May, R. M. (1982). Coevolution of hosts and parasites. *Parasitology* 85, 411–426. doi: 10.1017/S0031182000055360

Antunes, L. C. M., McDonald, J. A. K., Schroeter, K., Carlucci, C., Ferreira, R. B. R., Wang, M., et al. (2014). Antivirulence activity of the human gut metabolome. *mBio* 5, e1183–e1114. doi: 10.1128/mBio.01183-14

Bakken, J. S., Borody, T., Brandt, L. J., Brill, J. V., Demarco, D. C., Franzos, M. A., et al. (2011). Treating *Clostridium difficile* infection with fecal microbiota transplantation. *Clin. Gastroenterol. Hepatol.* 9, 1044–1049. doi: 10.1016/j.cgh.2011.08.014

Belkaid, Y., and Harrison, O. J. (2017). Homeostatic immunity and the microbiota. *Immunity* 46, 562–576. doi: 10.1016/j.immuni.2017.04.008

Biedermann, P. H., and Rohlfs, M. (2017). Evolutionary feedbacks between insect sociality and microbial management. *Curr. Opin. Insect Sci.* 22, 92–100. doi: 10.1016/j.cois.2017.06.003

Bohnhoff, M., Drake, B. L., and Miller, C. P. (1954). Effect of streptomycin on susceptibility of intestinal tract to experimental *Salmonella* infection. *Proc. Soc. Exp. Biol. Med.* 86, 132–137. doi: 10.3181/00379727-86-21030

Bokulich, N. A., Chung, J., Battaglia, T., Henderson, N., Jay, M., Li, H., et al. (2016). Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Sci. Transl. Med.* 8:343ra82. doi: 10.1126/scitranslmed.aad7121

Bratburd, J. R., Keller, C., Vivas, E., Gemperline, E., Li, L., Rey, F. E., et al. (2018). Gut microbial and metabolic responses to *Salmonella enterica* Serovar Typhimurium and *Candida albicans*. *mBio* 9, e2032–e2018. doi: 10.1128/MBIO.02032-18

Breznak, J. A., and Brune, A. (1994). Role of microorganisms in the digestion of lignocellulose by termites. *Annu. Rev. Entomol.* 39, 453–487. doi: 10.1146/annurev.en.39.010194.002321

Brito, I. L., Gurry, T., Zhao, S., Huang, K., Young, S. K., Shea, T. P., et al. (2019). Transmission of human-associated microbiota along family and social networks. *Nat. Microbiol.* 4, 964–971. doi: 10.1038/s41564-019-0409-6

Burchill, A. T., and Moreau, C. S. (2016). Colony size evolution in ants: macroevolutionary trends. *Insectes Soc.* 63, 291–298. doi: 10.1007/s00040-016-0465-3

Cafaro, M. J., Poulsen, M., Little, A. E. F., Price, S. L., Gerardo, N. M., Wong, B., et al. (2011). Specificity in the symbiotic association between fungus-growing ants and protective *Pseudonocardia* bacteria. *Proc. R. Soc. B Biol. Sci.* 278, 1814–1822. doi: 10.1098/rspb.2010.2118

Cariveau, D. P., Elijah Powell, J., Koch, H., Winfree, R., and Moran, N. A. (2014). Variation in gut microbial communities and its association with pathogen infection in wild bumble bees (*Bombus*). *ISME J.* 8, 2369–2379. doi: 10.1038/ismej.2014.68

Carr, G., Derbyshire, E. R., Caldera, E., Currie, C. R., and Clardy, J. (2012). Antibiotic and antimalarial quinones from fungus-growing ant-associated *Pseudonocardia* sp. *J. Nat. Prod.* 75, 1806–1809. doi: 10.1021/np300380t

Chevrette, M. G., Carlson, C. M., Ortega, H. E., Thomas, C., Ananiev, G. E., Barns, K. J., et al. (2019). The antimicrobial potential of *Streptomyces* from insect microbiomes. *Nat. Commun.* 10:516. doi: 10.1038/s41467-019-08438-0

Cohen, J. E. (2003). Human population: the next half century. *Science* 302, 1172–1175. doi: 10.1126/science.1088665

Cremer, S., Armitage, S. A. O., and Schmid-Hempel, P. (2007). Social immunity. *Curr. Biol.* 17, R693–R702. doi: 10.1016/J.CUB.2007.06.008

Currie, C. R., Bot, A. N. M., and Boomsma, J. J. (2003). Experimental evidence of a tripartite mutualism: bacteria protect ant fungus gardens from specialized parasites. *Oikos* 101, 91–102. doi: 10.1034/j.1600-0706.2003.12036.x

Currie, C. R., Mueller, U. G., Malloch, D., Dowd, S. E., Hong, E., and Mueller, U. G. (1999a). The agricultural pathology of ant fungus gardens. *Proc. Natl. Acad. Sci. U.S.A.* 96, 7998–8002. doi: 10.1073/pnas.96.14.7998

Currie, C. R., Scott, J. A., Summerbell, R. C., and Malloch, D. (1999b). Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. *Nature* 398, 701–704. doi: 10.1038/19519

Curtis, V. A. (2007). Dirt, disgust and disease: a natural history of hygiene. *J. Epidemiol. Community Health* 61, 660–664. doi: 10.1136/jech.2007.062380

Dill-McFarland, K. A., Tang, Z.-Z., Kemis, J. H., Kerby, R. L., Chen, G., Palloni, A., et al. (2019). Close social relationships correlate with human gut microbiota composition. *Sci. Rep.* 9:703. doi: 10.1038/s41598-018-37298-9

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

Douglas, A. E., Bouvaine, S., and Russell, R. R. (2011). How the insect immune system interacts with an obligate symbiotic bacterium. *Proc. R. Soc. B Biol. Sci.* 278, 333–338. doi: 10.1098/rspb.2010.1563

Engel, P., Martinson, V. G., and Moran, N. A. (2012). Functional diversity within the simple gut microbiota of the honey bee. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11002–11007. doi: 10.1073/pnas.1202970109

Engl, T., Kroiss, J., Kai, M., Nechitaylo, T. Y., Svatoš, A., and Kaltenpoth, M. (2018). Evolutionary stability of antibiotic protection in a defensive symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2020–E2029. doi: 10.1073/pnas.1719797115

Fernández-Marín, H., Zimmerman, J. K., Nash, D. R., Boomsma, J. J., and Wcislo, W. T. (2009). Reduced biological control and enhanced chemical pest management in the evolution of fungus farming in ants. *Proc. Biol. Sci.* 276, 2263–2269. doi: 10.1098/rspb.2009.0184

Ferretti, P., Pasolli, E., Tett, A., Asnicar, F., Gorfer, V., Fedi, S., et al. (2018). Mother-to-infant microbial transmission from different body sites shapes the developing infant gut microbiome. *Cell Host. Microbe* 24, 133.e5–145.e.5 doi: 10.1016/J.CHOM.2018.06.005

Fisher, K., West, M., Lomeli, A. M., Woodard, S. H., and Purcell, J. (2019). Are societies resilient? Challenges faced by social insects in a changing world. *Insectes Soc.* 66, 5–13. doi: 10.1007/s00040-018-0663-2

Flórez, L. V., Biedermann, P. H. W., Engl, T., and Kaltenpoth, M. (2015). Defensive symbioses of animals with prokaryotic and eukaryotic microorganisms. *Nat. Prod. Rep.* 32, 904–936. doi: 10.1039/c5np00010f

George Kerry, R., Patra, J. K., Gouda, S., Park, Y., Shin, H. S., and Das, G. (2018). Benefaction of probiotics for human health: a review. *J. Food Drug Anal.* 26, 927–939. doi: 10.1016/j.jfda.2018.01.002

Grubbs, K. J., Surup, F., Biedermann, P. H. W., McDonald, B. R., Klassen, J., Carlson, C. M., et al. (2019). Cycloheximide-producing *Streptomyces* associated with *Xyleborinus saxesenii* and *Xyleborus affinis* fungus-farming ambrosia beetles. *bioRxiv* [preprint]. doi: 10.1101/511493

Harrison, F., Roberts, A. E. L., Gabrilska, R., Rumbaugh, K. P., Lee, C., and Diggle, S. P. (2015). A 1,000-Year-Old antimicrobial remedy with antistaphylococcal activity. *MBio* 6:e01129. doi: 10.1128/mBio.01129-15

Hsiao, A., Ahmed, A. M. S., Subramanian, S., Griffin, N. W., Drewry, L. L., Petri, W. A., et al. (2014). Members of the human gut microbiota involved in recovery from *Vibrio cholerae* infection. *Nature* 515, 423–426. doi: 10.1038/nature13738

Hulcr, J., and Stelinski, L. L. (2017). The ambrosia symbiosis: from evolutionary ecology to practical management. *Annu. Rev. Entomol.* 62, 285–303. doi: 10.1146/annurev-ento-031616-035105

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. doi: 10.1038/nature11234

Institute of Medicine (US) Committee for the Study of the Future of Public Health (1988). *The Future of Public Health.* Washington, D.C.: National Academies Press. doi: 10.17226/1091

International Aphid Genomics Consortium (2010). Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol.* 8:e1000313. doi: 10.1371/journal.pbio.1000313

Jernberg, C., Löfmark, S., Edlund, C., and Jansson, J. K. (2007). Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *ISME J.* 1, 56–66. doi: 10.1038/ismej.2007.3

Johnson, J. B., and Hagen, K. S. (1981). A neuropterous larva uses an allomone to attack termites. *Nature* 289, 506–507. doi: 10.1038/289506a0

Johnson, K. V.-A., and Foster, K. R. (2018). Why does the microbiome affect behaviour? *Nat. Rev. Microbiol.* 16, 647–655. doi: 10.1038/s41579-018-0014-3

Keel, C., Schnider, U., Maurhofer, M., Voisard, C., Laville, J., Burger, U., et al. (1992). Suppression of root diseases by *Pseudomonas fluorescens* CHA0: importance of the bacterial secondary metabolite 2,4-Diacetylphloroglucinol. *Mol. Plant-Microbe Interact.* 5:4. doi: 10.1094/MPMI-5-004

Kembel, S. W., Jones, E., Kline, J., Northcutt, D., Stenson, J., Womack, A. M., et al. (2012). Architectural design influences the diversity and structure of the built environment microbiome. *ISME J.* 6, 1469–1479. doi: 10.1038/ismej.2011.211

Khadempour, L., Burnum-Johnson, K. E., Baker, E. S., Nicora, C. D., Webb-Robertson, B.-J. M., White, R. A., et al. (2016). The fungal cultivar of leaf-cutter ants produces specific enzymes in response to different plant substrates. *Mol. Ecol.* 25, 5795–5805. doi: 10.1111/mec.13872

Khoruts, A., and Sadowsky, M. J. (2016). Understanding the mechanisms of faecal microbiota transplantation. *Nat. Rev. Gastroenterol. Hepatol.* 13, 508–516. doi: 10.1038/nrgastro.2016.98

King, K. C., and Bonsall, M. B. (2017). The evolutionary and coevolutionary consequences of defensive microbes for host-parasite interactions. *BMC Evol. Biol.* 17:190. doi: 10.1186/s12862-017-1030-z

Koch, H., and Schmid-Hempel, P. (2011). Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19288–19292. doi: 10.1073/pnas.1110474108

Korb, J., and Aanen, D. K. (2003). The evolution of uniparental transmission of fungal symbionts in fungus-growing termites (*Macrotermitinae*). *Behav. Ecol. Sociobiol.* 53, 65–71. doi: 10.1007/s00265-002-0559-y

Korpela, K., and de Vos, W. M. (2018). Early life colonization of the human gut: microbes matter everywhere. *Curr. Opin. Microbiol.* 44, 70–78. doi: 10.1016/J.MIB.2018.06.003

Kutsukake, M., Moriyama, M., Shigenobu, S., Meng, X.-Y., Nikoh, N., Noda, C., et al. (2019). Exaggeration and cooption of innate immunity for social defense. *Proc. Natl. Acad. Sci. U.S.A.* 116, 8950–8959. doi: 10.1073/PNAS.1900917116

Lavy, O., Gophna, U., Gefen, E., and Ayali, A. (2018). The effect of density-dependent phase on the locust gut bacterial composition. *Front. Microbiol.* 9:3020. doi: 10.3389/fmicb.2018.03020

Li, H., Sosa-Calvo, J., Horn, H. A., Pupo, M. T., Clardy, J., Rabeling, C., et al. (2018). Convergent evolution of complex structures for ant-bacterial defensive symbiosis in fungus-farming ants. *Proc. Natl. Acad. Sci. U.S.A.* 115, 10720–10725. doi: 10.1073/pnas.1809332115

Lizé, A., McKay, R., and Lewis, Z. (2014). Kin recognition in *Drosophila*: the importance of ecology and gut microbiota. *ISME J.* 8, 469–477. doi: 10.1038/ismej.2013.157

Maire, J., Vincent-Monégat, C., Balmand, S., Vallier, A., Hervé, M., Masson, F., et al. (2019). Weevil pgrp-lb prevents endosymbiont TCT dissemination and chronic host systemic immune activation. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5623–5632. doi: 10.1073/pnas.1821806116

Marsh, S. E., Poulsen, M., Pinto-Tomás, A., and Currie, C. R. (2014). Interaction between workers during a short time window is required for bacterial symbiont transmission in *Acromyrmex* leaf-cutting ants. *PLoS One* 9:e103269. doi: 10.1371/journal.pone.0103269

Martinson, V. G., Moy, J., and Moran, N. A. (2012). Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl. Environ. Microbiol.* 78, 2830–2840. doi: 10.1128/AEM.07810-11

Matsuura, K. (2001). Nestmate recognition mediated by intestinal bacteria in a termite, *Reticulitermes speratus. Oikos* 92, 20–26. doi: 10.1034/j.1600-0706.2001.920103.x

Meunier, J. (2015). Social immunity and the evolution of group living in insects. *Philos. Trans. R. Soc. B Biol. Sci.* 370, 20140102–20140102. doi: 10.1098/rstb.2014.0102

Midani, F. S., Weil, A. A., Chowdhury, F., Begum, Y. A., Khan, A. I., Debela, M. D., et al. (2018). Human gut microbiota predicts susceptibility to *Vibrio cholerae* infection. *J. Infect. Dis.* 218, 645–653. doi: 10.1093/infdis/jiy192

Mockler, B. K., Kwong, W. K., Moran, N. A., and Koch, H. (2018). Microbiome Structure Influences Infection by the Parasite *Crithidia bombi* in Bumble Bees. *Appl. Environ. Microbiol.* 84, e2335–e2317. doi: 10.1128/AEM.02335-17

Moeller, A. H., Suzuki, T. A., Phifer-Rixey, M., and Nachman, M. W. (2018). Transmission modes of the mammalian gut microbiota. *Science* 362, 453–457. doi: 10.1126/science.aat7164

Mueller, U. G., Gerardo, N. M., Aanen, D. K., Six, D. L., and Schultz, T. R. (2005). The evolution of agriculture in insects. *Annu. Rev. Ecol. Evol. Syst.* 36, 563–595. doi: 10.1146/annurev.ecolsys.36.102003.152626

Nakashima, K., Watanabe, H., and Azuma, J.-I. (2002). Cellulase genes from the parabasalian symbiont *Pseudotrichonympha grassii* in the hindgut of the wood-feeding termite *Coptotermes formosanus*. *Cell. Mol. Life Sci.* 59, 1554–1560. doi: 10.1007/s00018-002-8528-1

Näpflin, K., and Schmid-Hempel, P. (2018). High gut microbiota diversity provides lower resistance against infection by an intestinal parasite in bumblebees. *Am. Nat.* 192, 131–141. doi: 10.1086/698013

Norris, D. (1965). The complex of fungi essential to growth and development of *Xyleborus sharpi* in wood. *Mater. Org. Beih* 1, 523–529.

Nyholm, S. V., and Graf, J. (2012). Knowing your friends: invertebrate innate immunity fosters beneficial bacterial symbioses. *Nat. Rev. Microbiol.* 10, 815–827. doi: 10.1038/nrmicro2894

Oh, D.-C., Poulsen, M., Currie, C. R., and Clardy, J. (2009). Dentigerumycin: a bacterial mediator of an ant-fungus symbiosis. *Nat. Chem. Biol.* 5, 391–393. doi: 10.1038/nchembio.159

Orenstein, W. A., and Ahmed, R. (2017). Simply put: vaccination saves lives. *Proc. Natl. Acad. Sci. U.S.A.* 114, 4031–4033. doi: 10.1073/pnas.1704507114

Otani, S., Bos, N., and Yek, S. H. (2016). Transitional complexity of social insect immunity. *Front. Ecol. Evol.* 4:69. doi: 10.3389/fevo.2016.00069

Otani, S., Zhukova, M., Koné, N. A., da Costa, R. R., Mikaelyan, A., Sapountzis, P., et al. (2019). Gut microbial compositions mirror caste-specific diets in a major lineage of social insects. *Environ. Microbiol. Rep.* 11, 196–205. doi: 10.1111/1758-2229.12728

Poulsen, M., Cafaro, M. J., Erhardt, D. P., Little, A. E. F., Gerardo, N. M., Tebbets, B., et al. (2010). Variation in *Pseudonocardia* antibiotic defence helps govern parasite-induced morbidity in *Acromyrmex* leaf-cutting ants. *Environ. Microbiol. Rep.* 2, 534–540. doi: 10.1111/j.1758-2229.2009.00098.x

Principi, N., and Esposito, S. (2016). Antibiotic administration and the development of obesity in children. *Int. J. Antimicrob. Agents* 47, 171–177. doi: 10.1016/J.IJANTIMICAG.2015.12.017

Pringle, H. (1998). The slow birth of agriculture. *Science* 282, 1446–1446. doi: 10.1126/SCIENCE.282.5393.1446

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821

Ramadhar, T. R., Beemelmanns, C., Currie, C. R., and Clardy, J. (2014). Bacterial symbionts in agricultural systems provide a strategic source for antibiotic discovery. *J. Antibiot.* 67, 53–58. doi: 10.1038/ja.2013.77

Raymann, K., Shaffer, Z., and Moran, N. A. (2017). Antibiotic exposure perturbs the gut microbiota and elevates mortality in honeybees. *PLoS Biol.* 15:e2001861. doi: 10.1371/journal.pbio.2001861

Sawe, B. E. (2018). *The 10 Largest Cities in the World.* Available at: https://www.worldatlas.com/articles/the-10-largest-cities-in-the-world.html (accessed August 26, 2019).

Schmid-Hempel, P. (2017). Parasites and their social hosts. *Trends Parasitol.* 33, 453–462. doi: 10.1016/J.PT.2017.01.003

Schultz, T. R., and Brady, S. G. (2008). Major evolutionary transitions in ant agriculture. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5435–5440. doi: 10.1073/pnas.0711024105

Schwarz, R. S., Moran, N. A., and Evans, J. D. (2016). Early gut colonizers shape parasite susceptibility and microbiota composition in honey bee workers. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9345–9350. doi: 10.1073/pnas.1606631113

Sharma, A., and Gilbert, J. A. (2018). Microbial exposure and human health. *Curr. Opin. Microbiol.* 44, 79–87. doi: 10.1016/J.MIB.2018.08.003

Sherwin, E., Bordenstein, S. R., Quinn, J. L., Dinan, T. G., and Cryan, J. F. (2019). Microbiota and the social brain. *Science* 366:eaar2016. doi: 10.1126/science.aar2016

Sonnenburg, J. L., and Sonnenburg, E. D. (2019). Vulnerability of the industrialized microbiota. *Science* 366:eaaw9255. doi: 10.1126/science.aaw9255

Stow, A., and Beattie, A. (2008). Chemical and genetic defenses against disease in insect societies. *Brain. Behav. Immun.* 22, 1009–1013. doi: 10.1016/j.bbi.2008.03.008

Syed Ab Rahman, S. F., Singh, E., Pieterse, C. M. J., and Schenk, P. M. (2018). Emerging microbial biocontrol strategies for plant pathogens. *Plant Sci.* 267, 102–111. doi: 10.1016/J.PLANTSCI.2017.11.012

Trappeniers, K., Matetovici, I., Van Den Abbeele, J., and De Vooght, L. (2019). The tsetse fly displays an attenuated immune response to its secondary symbiont, *Sodalis glossinidius*. *Front. Microbiol.* 10:1650. doi: 10.3389/fmicb.2019.01650

Turnbaugh, P. J., and Gordon, J. I. (2009). The core gut microbiome, energy balance and obesity. *J. f Physiol.* 587(Pt 17), 4153–4158. doi: 10.1113/jphysiol.2009.174136

Um, S., Fraimout, A., Sapountzis, P., Oh, D.-C., and Poulsen, M. (2013). The fungus-growing termite *Macrotermes natalensis* harbors bacillaene-producing *Bacillus* sp. that inhibit potentially antagonistic fungi. *Sci. Rep.* 3:3250. doi: 10.1038/srep03250

Van Arnam, E. B., Ruzzini, A. C., Sit, C. S., Horn, H., Pinto-Tomás, A. A., Currie, C. R., et al. (2016). Selvamicin, an atypical antifungal polyene from two alternative genomic contexts. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12940–12945. doi: 10.1073/pnas.1613285113

Vangay, P., Johnson, A. J., Ward, T. L., Al-Ghalith, G. A., Shields-Cutler, R. R., Hillmann, B. M., et al. (2018). US Immigration westernizes the human gut microbiome. *Cell* 175, 962.e10–972.e10. doi: 10.1016/j.cell.2018.10.029

Vuong, H. E., Yano, J. M., Fung, T. C., and Hsiao, E. Y. (2017). The microbiome and host behavior. *Annu. Rev. Neurosci.* 40, 21–49. doi: 10.1146/annurev-neuro-072116-031347

White, J. F., and Torres, M. S. (2009). *Defensive Mutualism in Microbial Symbiosis.* Boca Raton, FL: CRC.

Wilson, E. O. (1987). Causes of ecological success: the case of the ants. *J. Anim. Ecol.* 56:1. doi: 10.2307/4795

Youngster, I., Mahabamunuge, J., Systrom, H. K., Sauk, J., Khalili, H., Levin, J., et al. (2016). Oral, frozen fecal microbiota transplant (FMT) capsules for recurrent *Clostridium difficile* infection. *BMC Med.* 14:134. doi: 10.1186/s12916-016-0680-9

Zhang, Y., Brady, A., Jones, C., Song, Y., Darton, T. C., Jones, C., et al. (2018). Compositional and functional differences in the human gut microbiome correlate with clinical outcome following infection with wild-type *Salmonella enterica* serovar Typhi. *MBio* 9, e686–e618. doi: 10.1128/mBio.00686-18

Zhukovskaya, M., Yanagawa, A., and Forschler, B. (2013). Grooming behavior as a mechanism of insect disease defense. *Insects* 4, 609–630. doi: 10.3390/insects4040609

Zipperer, A., Konnerth, M. C., Laux, C., Berscheid, A., Janek, D., Weidenmaier, C., et al. (2016). Human commensals producing a novel antibiotic impair pathogen colonization. *Nature* 535, 511–516. doi: 10.1038/nature18634

Chapter 2: Colonization Dynamics and Genomic Adaptations in a Defensive Symbiosis

Jennifer R. Bratburd, Joseph A. Sardina, Weilan Gomes da Paixão Melo, Ethan B. Van Arnam, Caitlin M. Carlson, Monica T. Pupo, Adrián A. Pinto-Tomas, Cameron R. Currie

2.1 Abstract

Background: Many species of fungus-growing ants engage in a defensive symbiosis with antibiotic-producing *Pseudonocardia* bacteria. The Actinobacterial genus *Pseudonocardia* is phylogenetically and physiological diverse, occurring across environments such as soil, sediment, plants, and industrial wastes. While fungus-growing ants are known to have structures that support the bacteria's growth, the specificity and adaptations by the bacteria to the host are less clear. Here we investigated specificity of ant-associated *Pseudonocardia* to the ants versus non-ant associated strains.

Results: Non-ant associated *Pseudonocardia* were capable of colonizing ants, but ant-associated strains and related strains were more consistently able to colonize ants. Across a larger set of genomes we had 35 distinct species based on ANI cutoffs, and ant-associated *Pseudonocardia* tended to clade together in two major groups. Ant-associated genomes are slightly reduced compared to genomes from other sources with possible loss of certain functions like nitrate reduction and increase in transposases. More biosynthetic gene clusters were found in ant-associated genomes, which also tended to share families of clusters.

Conclusions: We show that the defensive symbiont *Pseudonocardia* exhibits both specificity in colonization and some corresponding genomic differences, as compared to strains from non-ant associated lineages of the genus. Further colonization experiments competing strains of

Pseudonocardia would help elucidate the factors shaping the specificity in this defensive symbiosis.

2.2 Introduction

Symbiotic associations shape ecological functioning and drives host and symbiont evolution (1–3). Symbioses can range from mutually beneficial associations (mutualism) to exploitative relationships (parasitism) (4), as well as from obligate dependencies to opportunistic associations (5,6). Host-symbiont partner fidelity and specificity is, at least in part, shaped by mechanisms that mediate formation of stable associations (7). The colonization of a host by a potential symbiont is a critical step in form these relationships (8). For some symbioses, a complex set of signals between both partners allows them to recognize each other, as in the squid-*Aliivibrio* (9,10) or legume-rhizobia (11). In other symbioses, hosts may tolerate a wider range of symbionts leading to more stochastic associations, dependent on what is readily available in the environment (12,13). To identify genetic components underlying mechanisms of colonization, genomic comparisons combined with experimental testing of symbiont strains colonization patterns is a useful strategy (14). The degree of adaptation to symbiosis is often evinced in the maintenance of various functions in symbiont genomes, where high host dependency on a microbial mutualistic symbiont correlates with the microbial symbiont's smaller genome size (15). Likewise, genomic signatures can also help reveal functionality of the symbionts (16). Here we use combine colonization experiments with comparative genomics to investigate the defensive mutualism of fungus-growing ants and *Pseudonocardia*.

Fungus-growing ants engage in an ancient defensive symbiosis with bacteria from the genus *Pseudonocardia.* The ants form a monophyletic lineage of more than 230 species that obligately depend on beneficial fungi they cultivate for food (17). The *Pseudonocardia* defensive

symbionts help protect the ants' fungal gardens from specialized fungal pathogens (18–20), through the production of diverse antimicrobial compounds (21–24). In exchange, the bacteria grow on the exoskeleton of the ants and is directly provided nutrients from the ants through glandular secretions (18,25,26). Fungus-farming ants and Actinobacteria are estimated to have been symbiotically associated for approximately 50 million years (20,27).

Fungus-growing ants have specialized behaviors, exocrine glands, and morphological structures for maintaining their relationship with *Pseudonocardia*. For instance, many fungus-growing ants have specialized cuticular structures, including crypts and tubercles, connected to internal exocrine glands that provide protection and nutrients to support *Pseudonocardia* growth (25,27). The metabolic costs to individual ants for maintaining *Pseudonocardia* can be substantial costing workers approximately ~10-20% of their basic metabolic rate (19). *Pseudonocardia* is transmitted vertically from parent to offspring colony by reproductive females (gynes) (18), and within colony transmission occurs from worker to worker. To help ensure partner fidelity, newly eclosed ants acquire *Pseudonocardia* during a narrow time window, primarily within two hours post-eclosion (28). Colonies typically maintain a single strain of *Pseudonocardia*, and in the laboratory-kept colonies these strains remain the same over years, in one case over a decade (29,30). However, individual fungus-growing ant workers are capable of being inoculated with novel strains if exposed during the acquisition window (28,31). Over evolutionary time, switches of *Pseudonocardia* between different ant hosts has occurred (20).

While there is substantial evidence of ant adaptations for hosting a defensive symbiont, the extent to which *Pseudonocardia* has adapted to living on the ants is unclear. Beyond the ants, *Pseudonocardia* strains can be found worldwide in a variety of habitats, including soil, wastewater, and plants (32–34). The bacteria are known for their ability to degrade pollutants

(35) and for their potential to produce novel natural products, such as antibiotics (21–24,36–38). Genomic comparisons of ant-associated *Pseudonocardia* strains have revealed several lineages associated with fungus-growing ants, evidence of codiversification, and variation in biosynthetic gene cluster potential on fine geographical scale (20,39–41), but comparisons of ant-associated *Pseudonocardia* to strains isolated from other sources is lacking. Likewise, previous research with cross fostering ants, in which are ants reared with different colonies to allow them to acquire different *Pseudonocardia* strains, indicates that different ant-associated phylotypes can colonize and grow on the ants to different abundances (42). It is unclear the extent to which non-ant associated strains may colonize the ants.

Here we investigate the specificity of the symbiosis by testing the ability of 16 *Pseudonocardia* strains and 2 *Streptomyces* from ant and non-ant sources to colonize fungus-growing ants. Furthermore, we conduct comparative genomic analyses of 71 *Pseudonocardia* strains isolated from a variety of sources to identify genomic signatures associated with being a defensive exosymbiont of fungus-growing ants. Specifically, we examine genome size, differential gene content, and predict biosynthetic gene clusters for ant and non-ant strains of *Pseudonocardia.*

2.3 Methods

2.3.1 Host-Symbiont Switching

Acromyrmex fungus-growing ant colonies were collected at La Selva Biological Station in Costa Rica and Universidad de Costa Rica campus in March and April 2019. Four colonies of *A. octospinosus* and one colony of *A. volcanus* were collected, with the majority of the pupae used in our switching experiments coming from two *A. octospinosus* colonies with n=87 and n=92 pupae, while the remaining others n=7, n=8, and for *A. volcanus* n=8 pupae were used. For

each colony we prioritized having at least one negative control, one ant-associated strain (*Pseudonocardia* sp. AL050505-11) and *P. spinosispora*. To determine the ability of different strains of *Pseudonocardia* to grow on cuticle of *Acromyrmex* ants we used pupae of female alates (gynes). The basis for using gynes is that: i) most pupae from these colonies were alates being produced for the impending nuptial flights, ii) males alates do not associate with the defensive symbiont and do not have specialized structures for *Pseudonocardia* (27), and iii) gynes engage in the vertical transmission of *Pseudonocardia* from parent to offspring nest. The ant-associated *Pseudonocardia* strains used in our experiments were previously obtained from *A. echinator* and *A. octospinosus.* Non-ant associated strains of *Pseudonocardia* were obtained from the Agricultural Research Services Culture Collection (NRRL) or isolated from other insects (Table 1). We also used *Streptomyces coelicolor* from NRRL and a *Streptomyces* strain isolated from a bee (43). Strains were originally isolated on chitin medium, following Cafaro and Currie (2005) and subsequently maintained on yeast malt extract agar (YMEA). Chitin medium was made with 4g chitin, 0.77g K_2HPO_4, 0.5 $MgSO_4 \cdot 7H_2O$, 0.37g KH_2PO_4, 0.01g $FeSO_4 \cdot 7H_2O$, 0.001 g $MnCl_2 \cdot 4H_2O$, and 0.001 $ZnSO_4 \cdot 7H_2O$ and 15g agar, while YMEA was made using 4 g yeast extract, 10 g malt extract, 4 g dextrose, 15 g agar, and 1 L water, both with 10,000 units/mL of nystatin and 0.05g/L cyclohexamide. We were able to isolate 3 strains from the 5 colonies used in this experiment, including *A. volcanus* (Supplemental Table 1).

Sub-colonies and *Pseudonocardia* inoculating followed methods from Marsh et al (2014). Briefly, sub-colonies were set up in plastic Petri dishes with moist cotton to provide humidity. To rear aposymbiotic *Acromyrmex* gynes for use in our experiments we prevent ants from acquiring *Pseudonocardia* from their own nests by removing pupae from their parent *Acromyrmex* colonies and rearing them to eclosion using *Atta cephalotes* workers. *Atta cephalotes* do not have

Pseudonocardia as external symbionts, and in previous work we have shown they can be used to rear pupae to aposymbiotic callow workers (i.e., *Pseudonocardia* free) (31). Approximately 0.1 g of *A. cephalotes* fungus garden was placed in a small weigh boat, along with a *Acromyrmex* pupa, and 4–6 minor and 1–3 medium caste *A. cephalotes* workers to both maintain the fungus garden and help remove the pupal casing when the pupa undergoes eclosion. Fungus garden was replaced as needed, such as when garden fragments became overgrown by a pathogen.

Ant pupae were monitored every 2–3 hours for signs of pending eclosion, as *Pseudonocardia* colonization is much less successful if it does not occur within the narrow 2 hour post-eclosion inoculation window (Marsh et al 2014). Ants undergoing eclosion overnight with undetermined number of hours post-eclosion were included in the negative control group. Upon eclosion, the aposymbiotic ants were moved to a new sub-colony set up as previously described but without the *Atta* workers. The aposymbiotic ants were then inoculated with 4 uL of the cell slurry directed towards the propleural plate, the location where in *Acromyrmex* spp. *Pseudonocardia* forms the most stable association with individual ant workers (i.e., the bacterium is typically just found on this location in more mature workers that no longer exhibit the symbiont over most of their exoskeleton) (18,44). For each of the 16 *Pseudonocardia* and 2 *Streptomyces* strains used to test colonization, a 1 cm diameter plug was taken from agar plates and diluted into 500 uL of phosphate buffered saline, following previously established methods (31). Negative control gynes were maintained under the same conditions, just without being inoculated with any bacterial strain. Ants were kept alive in sub-colonies for up to 14 days post-eclosion, and then stored in 90% ethanol.

2.3.2 Electron Microscopy

For visualizing growth with microscopy, we used the same procedure for inoculating ants described above with worker pupae from lab-reared *Acromyrmex echinator* colonies. We used environmental scanning electron microscopy (eSEM) to visualize filamentous bacteria on the ant exoskeleton. Ant specimens were stored in 90% ethanol at -20°C until imaging. Prior to examination, samples were air-dried at room temperature and subsequently mounted on eSEM stubs with carbon bi-adhesive tabs. Images were taken at 3.0 torr, 5.0 spot size, and 5°C using a FEI QUANTA 200 eSEM (FEI Company).

2.3.3 DNA Extraction and Genome Sequencing

DNA extractions were performed on the thorax region of individual ants, aseptically dissected from the corresponding gyne, or from pure cultures of bacteria grown on YMEA plates for whole genome sequencing. DNA extractions as follows: Buffer (200 mM Tris HCl pH 8.0, 200 mM NaCl and 20 mM EDTA) plus 20% SDS in water, and 500 uL phenol/chloroform was added to samples. Samples were beat with 1 stainless steel bead in solvent resistant screw cap tubes for 3 minutes, and then spun at 7200 x g at 4°C for 3 minutes. The aqueous layer was transferred to a new tube, where 60 uL NaAcetate and 600 uL of isopropanol were added. Samples were stored at -20°C for 1 hour up to 24 hours. After spinning samples at 18,000 x g for 20 minutes at 4°C, samples were decanted, rinsed with 100% ethanol, dissolved in nuclease free water and stored at -20°C. Library prep and genome sequencing was performed at the University of Wisconsin Madison Biotech Center and the Microbial Genome Sequencing Center in Pittsburgh, Pennsylvania.

2.3.4 PCR for Verifying *Pseudonocardia* Colonization

Successful bacterial colonization of gynes was assessed with PCR on individuals surviving greater than 7 days to ensure sufficient time for *Pseudonocardia* to establish on the ants (Poulsen 2003). We used elongation factor Tu gene (tuf) primers EFTuf 5'-GGCTTCGGCGTTCGACAT-3' and 5'-GCCGCCCTCATCCTTGCCC-3' (29) for PCR colonization assessment for all *Pseudonocardia* strains examined. For *Streptomyces coelicolor,* which does not readily amplify with those primers, we used specific primers for 16S: Coelf3 5'-CGCAGGCATCTGCGAGGTTCG-3' and Strep 261r 5'-GTCTGGGCCGTGTCTCAGTC-3. For PCR reactions, we used 12 uL EconoTaq Green Master Mix (Lucigen Corporation, Madison WI), 1 uL forward primer, 1 uL of reverse primer, 1 uL of template DNA, and 5 uL nuclease free water. For the EF Tu primers, PCR was performed with the following parameters: 94°C for 2.5 minutes; 35 cycles of 94°C for 45 seconds, 55°C for 50 seconds, 72°C for 2 minutes; 72°C for 10 minutes; then stored at 4 C. For the *S. coelicolor* primer, PCR was performed with the following parameters: 95C for 5 minutes; 35 cycles of 95°C for 10 seconds, 55°C for 10 seconds, 72°C for 5 seconds; 72°C for 10 minutes; then stored at 4°C. All the bands showed the expected size except for two, which we sequenced to determine if they could be off-target hits. Of the two, one from the *P. cypriaca* treatment matched the sequence of *Pantoea* sp. and the other from the negative control matched *Stentotrophomonas maltophila*. We did not include these off-target hits as positive results for *Pseudonocardia* hits in subsequent analysis. Sequencing the EF Tu amplified gene from 8 randomly selected ants revealed the closest match to the treatment *Pseudonocardia*.

2.3.5 Genomic Analyses

To investigate genetic differences between strains of *Pseudonocardia* isolated from ants and non-ant sources, as well as those between consistent and inconsistent colonizers, we obtained and/or sequenced 177 strains from a variety of sources (Supplemental Table 1). In addition to isolating *Pseudonocardia* from various taxa of fungus-growing ants (including the ant genera *Trachymyrmex*, *Apterostigma*, *Acromyrmex*, *Atta*, *Cyphomyrmex*, and *Mycetarotes*), we included *Pseudonocardia* isolated from soil, plant roots, marine sediment, bioreactors, and wastewater. After pruning the dataset with Treemmer to reduce closely-related strains (Menardo et al. 2018), we included a total of 75 genomes in our dataset: 35 ant-associated *Pseudonocardia*, 37 non-ant associated *Pseudonocardia*, and 3 outgroups (*Gordonia* sp. SID5947, *Streptomyces coelicolor,* and *Streptomyce*s sp. SID10815). For genome assemblies, reads were trimmed using fastp version 0.19.5 (45) and assembled using SPAdes v3.11.1 (46). Otherwise, assemblies were obtained from NCBI as listed in supplemental table 1. To prune for lower quality genomes, we ran Anvi'o 5.1 over all samples and eliminated genomes that were less than 80% complete or greater than 50% redundancy as predicted based by Anvi'o based on an HMM model of conserved genes (47). We then used Anvi'o with conserved HMM profiles of bacterial single copy gene sequences (48), aligned these sequences using Mafft v7.310 (49) and created a tree with RAxML version 8.2.11 (50) and made a consensus tree after running 100 bootstraps. We used *Gordonia* sp. SID5947 as an outgroup to root the tree. We ran Treemmer (51) to reduce tree nodes while maintaining maximal diversity, complete genomes, and non-ant associated strains. We used fastANI to calculate pairwise ANI values (52).

After pruning the dataset, we annotated genomes with Prokka version 1.12 (53) and ran PyParanoid (54) over the predicted proteins to identify homologous gene families. We analyzed the relationship between homologs and ant colonization phenotype using treeWAS (55). We also

used Fisher's Exact Test to take into account the abundance of the homologs themselves using an in-house script with Bonferroni multiple testing corrections (https://github.com/bratburd/comparative-genomics). Finally, we used Anvi'o's enrichment on the Anvi'o annotated genomes to find enriched COG functional categories.

2.3.6 Biosynthetic Gene Cluster detection

To examine the biosynthetic gene clusters (BGCs) encoding the production of secondary metabolites, we used genomes with N50 values over 100,000 bp, since genome quality can impact predictions of BGCs (56). We used Antismash version 4.2.0 to identify BGCs. To help validate our use of Illumina short reads in our analyses, we confirmed predictions were concordant with the Illumina short read as with PacBio long read sequencing for those same strains where we had generated data using both sequencing platforms. We clustered Antismash-identified BGCs using BiG-SCAPE version 20181005 (57), along with known BGCs from MIBiG database version 1.4 (58). Scripts for organizing and analyzing output are available on Github (https://github.com/bratburd/comparative-genomics).

2.4 Results

2.4.1 *Pseudonocardia* Colonization of *Acromyrmex* Ants

We experimentally examined the ability of 16 strains of *Pseudonocardia* spp. to colonize the exoskeletons of *Acromyrmex* spp. callow gynes, using PCR-based approaches to test for colonization. We did not observe significant differences in survival of gynes in the different treatment groups, with the exception of the negative controls experience a reduction in mortality (Supplemental Figure 1). Relative to normal *Pseudonocardia* growth on ants from natural

transmission (Figure 1a), we did not observe any occurrence of what would be considered normal colonization of the exoskeleton of the gyne by any *Pseudonocardia,* including for ant-associated strains. We photographed a subset of gynes to illustrate the patchy growth on the exoskeleton of the ant (Figure 1B-1D). In addition, using laboratory worker ants reared with a small subset of different strains, we were able to examine these ants with environmental scanning electron microscopy (eSEM), where we observed patchy growth relative to ants naturally colonized with their native strain (Supplemental Figure 2). Of the strains examined in the colonization experiments, we did not detect any presence of bacteria in 3 treatments: *P. petroleophila* (n=12 gynes), *P. cypriaca* (n=7 gynes), or *P. zijingensis* (n=9 gynes) (Figure 1D). For the strains *P. kujensis* and *P. alaniphila* only one gyne out of 9 and 10, respectively, had positive PCR detection of the strain on the ant, while two gynes in each of *P. chloroethenivorans* (n=5 gynes) and *P. compacta* (n=6 gynes) had positive PCR detection support for some growth. For the remaining 9 strains, at least 3 gynes had positive PCR detection, but for the strains *P. spinosispora* and *P. endophytica* positive detection still remained below 50%, with just 3 of 14 and 3 of 8 having positive PCR detection, respectively. Strain detection for the remaining 7 strains occurred in more than half of the gynes tested, with the highest percentage in *P. kongjuensis* (6/7 ants). Overall, strains more closely related to ant-associated clades had more consistent detection, on average 73% ± 8 (mean ± SD), while the more phylogenetically basal *Pseudonocardia* strains were detected much less frequently, averaging 15% ± 16. Of the 7 consistently colonizing strains, 3 (*P. alni*, *P. nitrificans*, and *P. antarctica*) group together as a single species based on a 95% average nucleotide identity (ANI) cutoff (Figure 2).

To explore the ability of other genera of Actinobacteria to grow on the surface of ants, we also tested two strains of *Streptomyces.* We selected this genus in part based on previous reports

of *Streptomyces* strains being isolated from fungus-growing ants (59,60). As with *Pseudonocardia,* we did not observe signs of normal colonization (Supplemental Figure 3B). Nevertheless, we did detect growth of the two strains tested. *Streptomyces* sp. SID10815 (isolated from a bee) showed positive PCR amplification in 4/6 of ants tested, and *S. coelicolor* showed positive PCR amplification on 7/7 of ants tested (but required a primer set more specific to *S. coelicolor* to detect).

2.4.2 Genomic Diversity of *Pseudonocardia*

We calculated ANI and percentage of shared gene content to determine possible distinct species groupings and to compare with previously identified ant-associated groups (Figure 2). Both ANI and shared percentage gene content showed similar patterns, although gene content appeared more influenced by genome quality than ANI. By grouping strains at 95% ANI, a proposed cutoff for species-level taxonomy (52), we found 35 distinct groupings of *Pseudonocardia* in our trimmed dataset, most with a single representative, and only 4 “species groups” with more than 2 strain representatives (Supplemental Figure 4). Of the groupings we identified with greater than 95% ANI, only one had a mixture of ant-associated strains (derived from a variety of ant genera) and non-ant associated strains. All other “species groups” contained only ant-associated or only non-ant associated strains.

Consistently colonizing strains (i.e. strains able to colonize greater than 50% of individual ants tested) all fell within a broad clade that had above 83% ANI, within a previously identified genetic discontinuity between interspecies and intraspecies cutoffs (52). This group contained the majority of the ant strains, plus 9 non-ant strains (Supplemental Figure 4). Three

ant-associated strains fell outside of the consistent colonizing clade (*Pseudonocardia* sp. AL050513-04, *Pseudonocardia* sp. ICBG618, and *Pseudonocardia s*p. CC030328-06).

Previously identified phylotypes associated with ants were generally well-represented "species groups": Group IV/Ps1 (20,41) had 23 genomes sequences in the total untrimmed database, while GroupVI/Ps2 had 10 genomes sequenced. Ps1 included ant-strains mainly sampled from *Apterostigma dentigerum* ants in Panama, while Ps2 was composed primarily of strains isolated from *Acromyrmex* ants. *Pseudonocardia autotrophica* strains (DSM43083, NRRLB16064, DSM535) all grouped separately.

2.4.3 Genomic Comparisons of *Pseudonocardia*

To begin investigating potential genomic signatures associated with the life history of ant-associated *Pseudonocardia,* we compared genome sizes looking for evidence of genomic reduction. Genome size varied significantly across all *Pseudonocardia* strains analyzed, ranging from ~5,056,835 to ~10,179,404 bp (Figure 3). *Pseudonocardia* strains associated with fungus-growing ants were significantly smaller, with an average size of 6,473,749 ± 827,351 (SD) compared to the non-ant average of 7,400,603 ± 1,377,939 (Wilcoxon rank sum test, $p < 0.05$). When limiting the set to just the high quality draft genomes (32 strains with N50 > 100,000), we still found significant reduction in genome size ant strains relative to non-ant strains. Each of the five *Pseudonocardia* strains not associated with ants that clade within one of the two ant-associated lineages, *P. alni*, *P. antarctica*, *P. nitrificans*, *Pseudonocardia* sp. 10165 and *Pseudonocardia* sp. 10385, have reduced genomes, (5,994,807; 6,289,920; 5,070,148; 6,229,613; 6,142,889 respectively), as compared to the average of non-ant associated genomes. Likewise, the three ant-associated strains outside of the main two ant-associated clades, *Pseudonocardia* sp. AL050513-04, *Pseudonocardia* sp. ICBG618, and *Pseudonocardia* sp. CC030328-06, have

larger genome sizes than the average ant-associated strains (9,079,058; 8,707,602; 6,739,435, respectively).

To investigate gene content specific to the ant colonization by *Pseudonocardia*, we used Pyparanoid to identify 31,014 homologous groups of genes, including a core of 682 homologs present in 95% of the randomly trimmed set of 71 *Pseudonocardia* genomes. We compared enrichment of gene content based on two categories: (1) strains isolated originally from ants plus strains falling within the consistent colonizer clade, informed by the experimental colonization results and phylogeny or (2) isolation originally from ants versus other sources. Results from both are summarized in Figure 3.

Employing several strategies to detect enriched homologs, we found consistent colonizers had 220 homologs enriched versus 189 enriched in inconsistent colonizers. Based on COG category, the majority did not match to COGs with a known annotation. Of homologs with annotations, those that showed some enrichment in consistent colonizers included the replication and repair category (11 homologs enriched annotated for replication out of 134 total annotated homologs in consistent colonizers versus 4 of 170 for inconsistent colonizers). Upon inspection, this COG category was primarily comprised of 6 genes annotated as or with core domains of transposases and 3 integrases. Figure 3B shows examples of distribution of significantly enriched transposable elements from consistent versus inconsistent and ant versus non-ant comparisons. The colonization group tended to be depleted for genes in the transcription COG category (13 homologs enriched out of 134 total annotated homologs in consistent colonizers versus 32 of 170 in inconsistent colonizers). Distinct to the inconsistent colonizers in this category were 6 homologs annotated as sigma 70 factors, whereas the ant-associated *Pseudonocardia* lacked enriched homologs annotated as sigma 70 factors. We also examined individual homologs for

patterns that did not fall into the high level COG categorization for their consistent enrichment or depletion among consistent colonizing strains. Many *Pseudonocardia* in the inconsistent colonizer group retained genes for nitrate reduction (present in 1 of 44 consistent colonizer clade, versus 18 of 28 strains in the inconsistent colonizer lineages).

The analysis of homologs enriched by isolation source of ant versus other sources yielded similar results. We found 35 homologs from ant-associated strains, versus 32 homologs enriched in other non-ant isolates (Supplementary Table 2). 46 enriched homologs from ant-isolates overlapped with consistent colonizers. 9 of out 28 ant-associated enriched homologs fell into the replication and repair COG category, versus 0 of 30 in non-ant associated. Likewise, transcription annotated homologs were not enriched in ant-associated isolates (3 of 35 in ant-associates, 5 of 32 in non-ant associated). Similarly, nitrate reduction found in 19 of 37 non-ant associated strains and 0 of 35 ant associated strains.

2.4.4 Biosynthetic Gene Clusters

Given that this is a defensive symbiosis, with ant-associated *Pseudonocardia* strains providing chemical-mediated defense, we next annotated and compared biosynthetic gene clusters (BGCs) between ant and non-ant associated *Pseudonocardia*. In our non-reduced dataset of all available genomes, we had 9 strains that were sequenced with both short-read Illumina and long read PacBio platforms. In the lower quality short read assembled genomes, we found in general over counting of BGCs, especially NRPS and PKS, likely due to assembly issues for these highly repetitive sequences. In the most extreme case, BGCs more than doubled from 17 to 36 in PacBio versus Illumina sequencing (Supplemental Table 3). Thus, we chose to prioritize genome quality and maximize number of strains over random sampling across the phylogeny and

selected 66 genomes from the original nonreduced set of 177 that had not been randomly pruned with N50 greater than 100,000 bp for the remaining analysis (Supplemental Table 1).

Overall, we identified 940 total biosynthetic gene clusters across these genomes, which averages to 14 ± 5 BGCs detected per genome. The majority of BGCs detected belonged to four categories: Nonribosomal peptide synthetase (NRPS at 18%), terpenes (12%), bacteriocin (12%), and other (18%). On average, strains isolated from ants had slightly higher BGCs detected at 15 ± 3 than from other sources at 13 ± 6 ($p = 0.008$, Wilcoxon rank sum test). Strains in the consistent colonizer clade had on average 14 ± 3 BGCs, and those outside the clade had 13 ± 7. Genomes isolated from other sources ranked as having both the highest numbers of BGCs detected (*Pseudonocardia* sp. 15845 with 31) and lowest BGCs detected (*P. thermophila* with 7).

We used BiG-SCAPE to cluster BGCs into families. Few of these families were shared across all strains. Shared elements may have fallen into different family categories, such as the previously identified osmoprotectant ectoine (41) being found in nearly all the strains, BiG-SCAPE analysis suggests that these fall into at least 4 different families. In general, ant-associated *Pseudonocardia* tended to share more BGCs of the same families with each other than with non-ant associated strains (Figure 4). Clustering by BGC family presence revealed three main groups, *Pseudonocardia* Group IV/Ps1, *Pseudonocardia* Group VI/Ps2, and other non-ant strains. Most Group IV/Ps1 *Pseudonocardia* shared an ectoine family, two terpene families, 2 NRPS, oligosaccharide, bacteriocin and 2 other clusters. Within the Group VI/Ps2 clusters, they shared their own ectoine, 2 terpenes, bacteriocin and oligosaccharide. The other group included all of the 22 inconsistent colonizer strains, along with 4 strains that in our phylogenetic analyses grouped with the consistent colonizer strains. Similarly, this group contained all but two of the non-ant strains and only 3 of 38 ant strains. The BGCs in this group were more variable, as

expected given the much greater phylogenetic diversity represented across this group. Unique among predominantly non-ant associated strains, there was 1 cluster of terpenes shared among 13 of 24 of this category that were not found in consistent colonizer strains.

Using the MiBIG database (58) and BGCs previously identified in *Pseudonocardia*, we found some clusters grouping with previously identified compounds (Figure 3A). These included a mixed family linking NRPS and hybrid T1PKS-NRPS with selvamicin (22); a family of nystatin-like compounds (23); a NRPS family grouping gerumycins A, B, and C together; a T1PKS-NRPS hybrid family containing dentigerumycin (24); and a cluster of two linking 9-methoxyrebeccamycin with another indole (21). One basal insect associated strain isolated from a grasshopper contained a PKS-NRPS cluster that grouped with ristocetin and ristomycin A, an antifungal, and an NRPS cluster that grouped with albachelin. Most of the other BGC families identified in these analyses did not group with known BGCs from the MiBIG database.

2.5 Discussion

In this paper, we use experimental colonization and comparative genomics to investigate specificity of bacterial symbiont colonization of fungus-growing ants. We found variation in the ability of strains of *Pseudonocardia* to successfully grow on the exoskeleton of ants, with the strains able to consistently colonize ants grouping together phylogenetically with the two main clades of *Pseudonocardia* previously characterized. Further, we observed some potential genomic adaptations to the host, with ant-associated *Pseudonocardia* strains appear to have slightly smaller genomes, possibly losing some metabolic functions like nitrate reduction while maintaining high numbers of biosynthetic gene clusters.

In our experiments, we found a strong phylogenetic signal associated with the ability of *Pseudonocardia* to grow on the cuticle of *Acromyrmex* gynes. In addition to the two ant-

associated strains tested, non-ant strains exhibiting the ability to consistently grow were closely related to ant-associated strains. It is important to note that we observed patchy growth of *Pseudonocardia* on gynes that is different from normal natural colonization. This patchiness may be influenced by conducting inoculations by pipetting a cell slurry on the ant. Given this, we used a PCR-based assay to evaluate growth, as a proxy for colonization, rather than a visual inspection approach. Thus, this experiment does not address other elements in colonization, such as the role of the ants in spreading their bacteria, the ability of the strain to outcompete others on the newly eclosed ant, and overall differences in growth from different strains.

We focused on *Acromyrmex* spp. here, however, *Pseudonocardia* are also found on a variety of other fungus-growing ants, some of which have different structures such as crypts and tubercles for supporting and maintaining symbionts (27). Many of our consistent colonizing strains overlapped with clades associated with *Acromyrmex*. Other clades of *Pseudonocardia* may better at colonizing different fungus-growing ant hosts. Even among strains of *Pseudonocardia* isolated from *Acromyrmex*, some strains appear to be better colonizers based on cross-fostering experiments in term of ability to colonize and amount of growth on the ants (61).

Overall, we find that most ant-associated strains and consistently colonizing strains group together phylogenetically. Ant-associated strain grouping has been previously observed in multilocus and genome-based trees (20,41). Further, strains isolated from other sources that are closely related to ant-associated strains are able to colonize the ants consistently. This evidence, together with known barriers to colonization like physical structures, timing, and ant behavioral preference, supports claims of co-diversification of the ant host and symbiotic bacteria (20,27,28,62). Repeated loss of *Pseudonocardia* and structures to support it may indicate the cost of maintaining the symbiont occasionally may outweigh the benefits (19,27).

We also report here that in our experimental colonization setup, we detected growth of *Streptomyces* on *Acromyrmex* sp. gynes inoculated with this bacterium. A previous experiment to colonize workers with *Streptomyces* was unsuccessful (31). The dark color in the patch growth of *S. coelicolor* matches the strain colony pigmentation when grown on nutrient media in Petri plates, but to our knowledge no fungus-growing ant has been observed in nature with anything other than white growth of Actinobacteria. Both *Pseudonocardia* and *Streptomyces* are known to be good sources of potential antibiotics, and both can inhibit pathogens of the fungus garden, although *Pseudonocardia* tends to be more effective against the specific fungal pathogen *Escovopsis* of the fungal cultivar (20,63). *Streptomyces* has been isolated from the ants, and some researchers hypothesize that ants acquire their defensive bacteria as they forage, and these bacteria compete with antimicrobials on the surface of the ant which helps the ant acquire better antimicrobial producers (64). However, given the strong phylogenetic signal associated with *Pseudonocardia* growth on gynes we found here, as well as the narrow window of acquisition and specific method for transmission, it is perhaps unlikely that our result ecologically relevant colonization.

Varying degrees of genome reduction are commonly observed in microbial symbiont genomes as compared to free-living relatives and implicated by several factors including increase in genetic drift or loss of formerly essential genes replaced by reliance on host functions (65). Relative to non-ant associated *Pseudonocardia*, in ant-associated strains we see an average 1 Mb genome reduction in the overall size of the genome, as well as expansion of some transposases which could indicate early steps of genomic degradation. Further, there appears to be some loss of nitrate reductase genes which could potentially be related to the nutrition provided by the ants for the symbionts (26). Potentially, further sampling of difficult to culture strains from other

fungus-growing ant hosts (such as those that grow in association with crypts) may have more specific changes resulting from a more intimate symbiosis.

Several non-ant *Pseudonocardia* are isolated from other hosts, such as plants and other insects. Further, some strains have extremely limited metadata, such as *P. hydroxycarbonoxydans* which was originally isolated as an air contaminant, or strains associated with wastes that may have an original unknown source. The wider diversity of isolation sources combined with more variation in evolutionary history of non-ant strains may be partially why we observe a wider range of many metrics examined including genome size and number of biosynthetic gene clusters. Further sampling with detailed metadata could allow more exploration of the other *Pseudonocardia* clades. In particular, isolation of *Pseudonocardia* from various sources may be more feasible with new techniques like fluorescence *in situ* hybridization and flow cytometry (66).

As a defensive symbiont, the most important function *Pseudonocardia* provides to the host is the production of antibiotics. Here we have found that ant-associated strains have on average slightly higher numbers of BGCs detected in their genomes than non-ant strains. although the spread of BGCs overlaps between the two groups. Further, ant strains tend to share BGCs in the same families as compared to non-ant strains. Non-ant strains did share one family of terpene BGCs. Non-ant strains may also have some selective pressure for having BGCs such as competing in the environment or providing defense for other hosts. The highest BGC outliers from the non-ant strains were isolated from associations with soil, *Acacia auriculiformis*, and a grasshopper. Thus far, one endophytic *Pseudonocardia* has been implicated in promoting plant production of a defensive compound to our knowledge, defensive symbiosis of plant-associated *Pseudonocardia* has not been demonstrated yet (34). Of the BGCs identified, several are known

antifungals that can help target the fungal pathogen of the ant system, such as the nystatin cluster, which was highly prevalent in ant-associated *Pseudonocardia.* Other known compounds like selvamicin appeared more rarely. Rare BGCs, as well as modifications to known BGCs, may be important in keeping up with an evolutionary arms race against a resistant pathogen.

This study explores the role of the bacteria in symbiosis with fungus-growing ants and the impact on the bacterial genomes, and sheds light on how an important defensive symbiont may show a few changes in association with its host. Overall, we find that ant-associated strains group phylogenetically, tend to colonize more consistently than less related non-ant strains, have somewhat smaller genomes where some functions like nitrate reduction may be lost and have variation in BGCs, indicative of the host's behavioral role in maintaining fidelity.

Microbial colonization is key to stable associations between hosts and symbionts. Research exploring the limitations of microbial colonization, either with experimental manipulation or observation within the environment reveals that many dynamic variables can influence colonization. These factors include strain competition (10), environmental variables such as temperature (67), which can ultimately impact fitness of host and symbiont (68). This work provides insight into the colonization dynamics of a non-obligate relationship within a defensive symbiosis. Although this work is focused at a genus-level analysis, we still see evidence of genome reduction and possible loss of function in a system where much of the control of the relationship likely lies with the host. Future work is needed to explore how more fine-scale interactions, dynamic environmental variables, and symbiont competition can alter these relationships.

2.6 Author Contributions

Jennifer Bratburd performed colonization experiments, extracted DNA for sequencing genomes, analyzed genomes and wrote the manuscript. Joseph Sardina performed microscopy. Adrian Pinto organized the field research for the experimental colonizations and provided feedback on the experiments and manuscript. Ethan Van Arnam provided feedback on the manuscript, extracted and sequenced some genomes. Caitlin Carlson identified additional strain and extracted the DNA for sequencing. Weilan Gomes da Paixão Melo and Monica Pupo helped organize field research in Brazil, isolated and extracted strains for sequencing. Cameron Currie assisted with permits for all field collecting trips, provided resources and obtained funding for the research, provided feedback on experimental design and edited the manuscript.

2.7 Acknowledgements

Charlotte Francoeur, Allan Artavia, and Miguel for their field support, especially in regards to digging ant colonies. Other members of the Pinto lab and La Selva Biological Station for laboratory support in Costa Rica. Heidi Horn, Kirsten Gotting, Bradon McDonald, Marc Chevrette, Jonathan Klassen, Garret Suen, and Eric Caldera for information and suggestions on the project. Support for this project was provided through National Institutes of Health (NIH) U19 Al109673 and NIH U19 TW009872. Funding for JB provided through the University of Wisconsin-Madison Department of Bacteriology Michael and Winona Foster fellowship and NIH T32 AI55397.

2.8 Tables and Figures

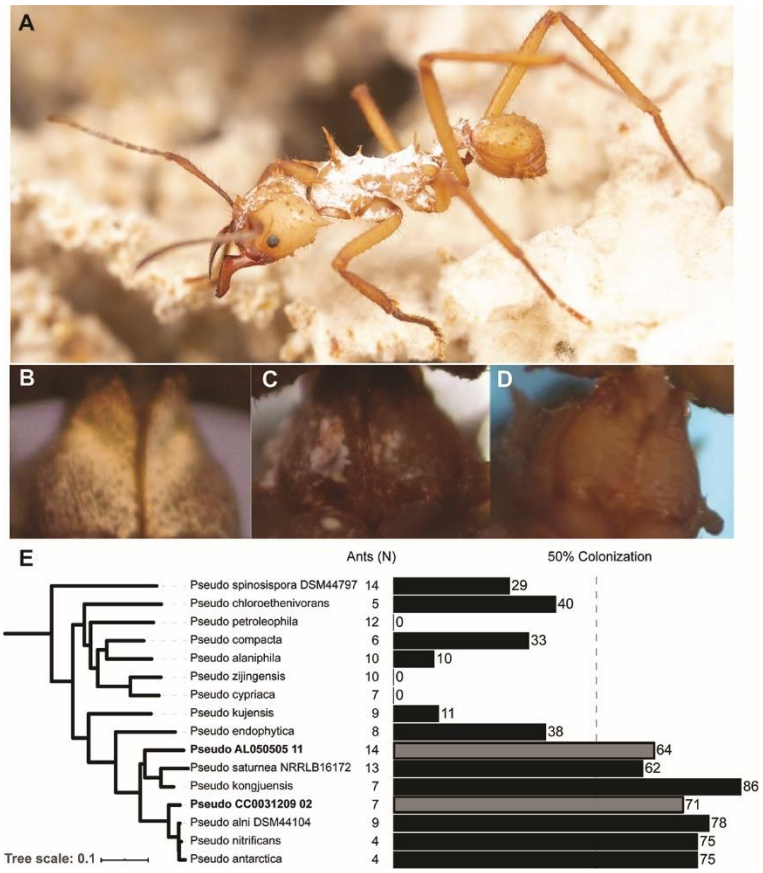


Figure 1. A. Normal *Pseudonocardia* colonization on *Acromyrmex echinator* worker. B. *Acromyrmex octospinosus* gyne colonized with *P. alni*. C. *Acromyrmex octospinosus* gyne unsuccessfully colonized with *P. alaniphila*. D. Percent of ants colonized with *Pseudonocardia* strains as detected with PCR. Number of ants used in each treatment in parenthesis next to strain. Percent of ants colonized at end of bar graph. Dashed line indicates 50% colonization detected. Strains in bold with gray bars indicate ant-associated strains.

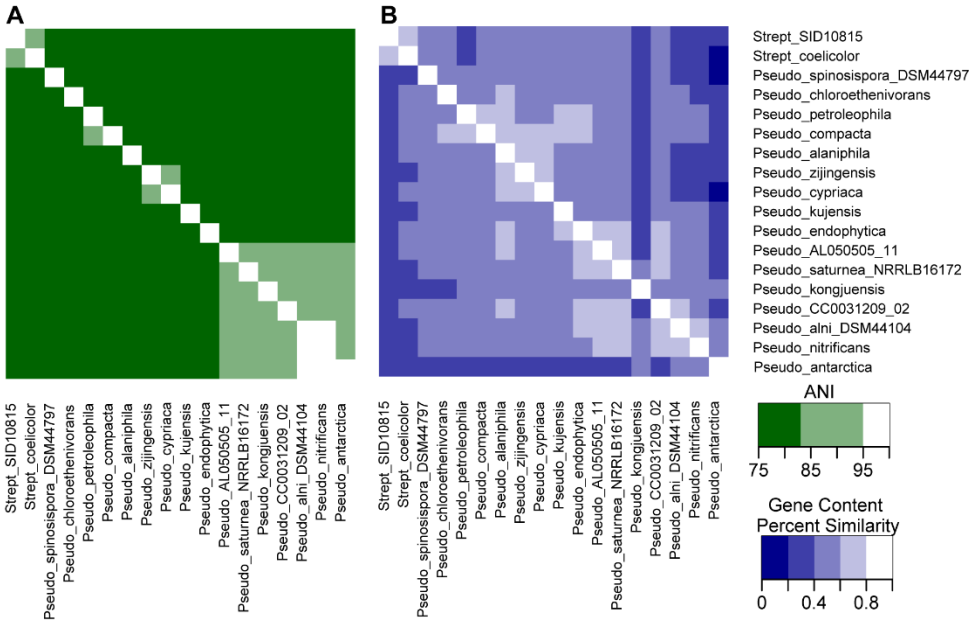


Figure 2. Overall genome similarity for strains used in colonization experiment. A. ANI of strains B. Percentage of shared gene content.

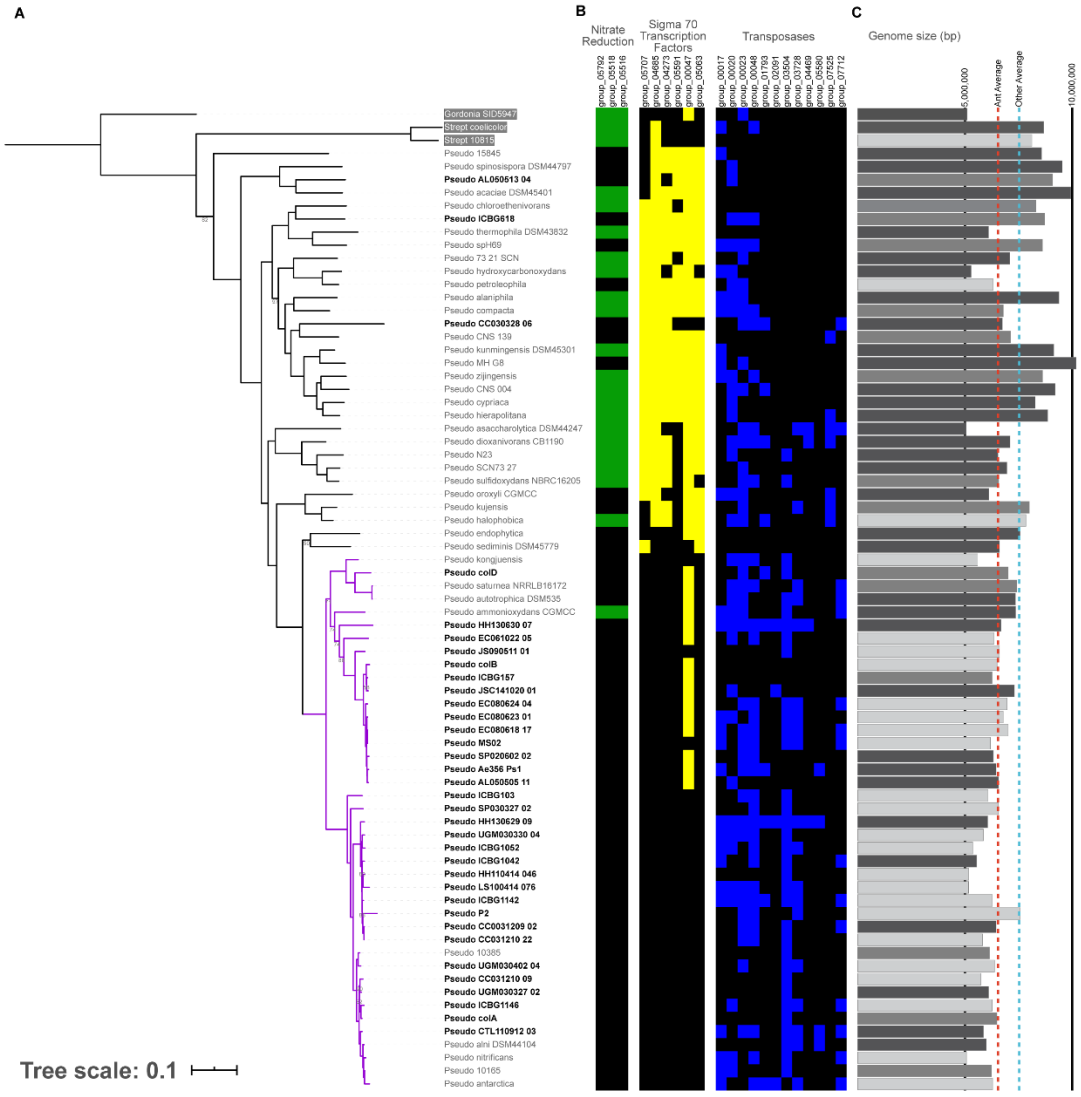


Figure 3. A. Concatenated core gene tree of selected *Pseudonocardia*. Bootstrap support for nodes with less than 97/100 displayed on tree. Taxa in bold black font indicate ant-associated strains, while non-ant associated strains represented by gray font. Purple branches on tree indicate consistently colonizing clades inferred from colonization experiment. B. Heatmap of homologs from PyParanoid selected as enriched based on Bonferroni corrected Fisher's Exact test, column 1 is nitrate reduction genes, column 2 is sigma 70 transcription factors, column 3 is transposable elements. Black indicates no detection. Darker colors indicate multiple copies detected. C. Genome lengths. Red dashed line indicates average ant-associated genome length.

Blue dashed line indicates average non-ant associated genome length. Genome quality indicated by color of bar, where dark gray represents N50 >= 100,000, medium gray represents N50 between 100,000 and 15,000, and the lightest gray represents n50 <= 15,000.

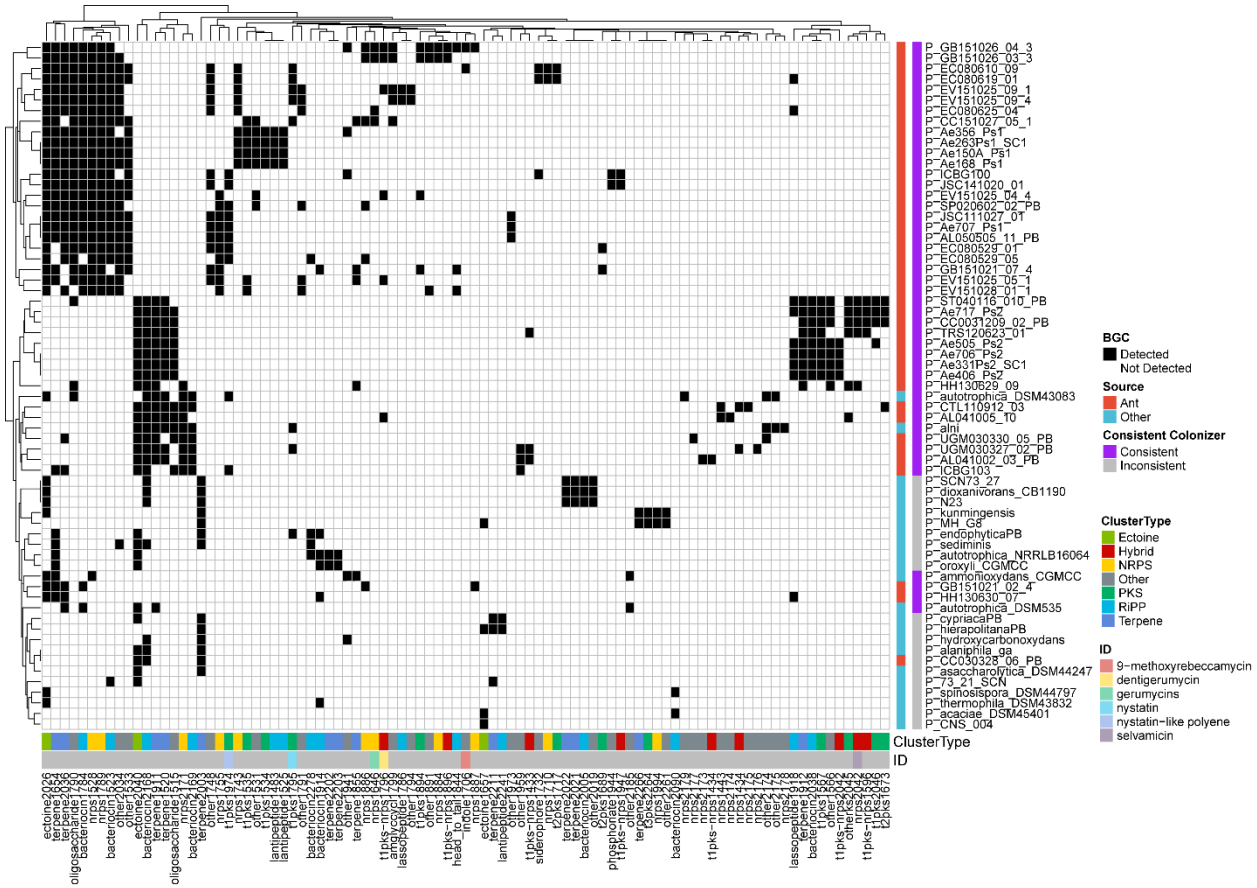


Figure 4. Biosynthetic gene clusters (BGCs) in *Pseudonocardia*. A. Heatmap of BGC family detection. Row and columns are clustered based on similarity. Shaded cell indicate presence of BGC family. Vertical bars represent isolation source and inclusion in consistently colonizing clade. Horizontal bars below heatmap indicate type of BGC detected based on the majority annotation for that family. The second bar below indicates BGC families that include matches to known compounds.

2.9 References

1. Klepzig KD, Adams AS, Handelsman J, Raffa KF. Symbioses: A Key Driver of Insect Physiological Processes, Ecological Interactions, Evolutionary Diversification, and Impacts on Humans. Environ Entomol. 2009 Feb 1;38(1):67–77.

2. Gilbert SF, Sapp J, Tauber AI. A Symbiotic View of Life: We Have Never Been Individuals. Q Rev Biol. 2012 Dec 1;87(4):325–41.

3. Henry LM, Peccoud J, Simon J-C, Hadfield JD, Maiden MJC, Ferrari J, et al. Horizontally Transmitted Symbionts and Host Colonization of Ecological Niches. Curr Biol CB. 2013 Sep 9;23(17):1713–7.

4. Flórez LV, Scherlach K, Gaube P, Ross C, Sitte E, Hermes C, et al. Antibiotic-producing symbionts dynamically transition between plant pathogenicity and insect-defensive mutualism. Nat Commun. 2017 Apr 28;8(1):15172.

5. Werner GDA, Cornwell WK, Cornelissen JHC, Kiers ET. Evolutionary signals of symbiotic persistence in the legume–rhizobia mutualism. Proc Natl Acad Sci. 2015 Aug 18;112(33):10262–9.

6. Werner GDA, Cornelissen JHC, Cornwell WK, Soudzilovskaia NA, Kattge J, West SA, et al. Symbiont switching and alternative resource acquisition strategies drive mutualism breakdown. Proc Natl Acad Sci. 2018 May 15;115(20):5229–34.

7. Kaltenpoth M, Roeser-Mueller K, Koehler S, Peterson A, Nechitaylo TY, Stubblefield JW, et al. Partner choice and fidelity stabilize coevolution in a Cretaceous-age defensive symbiosis. Proc Natl Acad Sci. 2014 Apr 29;111(17):6359–64.

8. Mandel MJ. Models and approaches to dissect host–symbiont specificity. Trends Microbiol. 2010 Nov 1;18(11):504–11.

9. Bongrand C, Koch EJ, Moriano-Gutierrez S, Cordero OX, McFall-Ngai M, Polz MF, et al. A genomic comparison of 13 symbiotic Vibrio fischeri isolates from the perspective of their host source and colonization behavior. ISME J. 2016 Dec;10(12):2907–17.

10. Bongrand C, Ruby EG. Achieving a multi-strain symbiosis: strain behavior and infection dynamics. ISME J. 2019 Mar;13(3):698–706.

11. Wang Q, Liu J, Zhu H. Genetic and Molecular Mechanisms Underlying Symbiotic Specificity in Legume-Rhizobium Interactions. Front Plant Sci [Internet]. 2018 [cited 2020 Jul 6];9. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2018.00313/full

12. Lazzaro BP, Fox GM. Host–Microbe Interactions: Winning the Colonization Lottery. Curr Biol. 2017 Jul 10;27(13):R642–4.

13. Tianero MDB, Kwan JC, Wyche TP, Presson AP, Koch M, Barrows LR, et al. Species specificity of symbiosis and secondary metabolism in ascidians. ISME J. 2015 Mar;9(3):615–28.

14. Kirzinger MWB, Stavrinides J. Host specificity determinants as a genetic continuum. Trends Microbiol. 2012 Feb 1;20(2):88–93.

15. Fisher RM, Henry LM, Cornwallis CK, Kiers ET, West SA. The evolution of host-symbiont dependence. Nat Commun. 2017 Jul 4;8(1):1–8.

16. Waterworth SC, Flórez LV, Rees ER, Hertweck C, Kaltenpoth M, Kwan JC. Horizontal Gene Transfer to a Defensive Symbiont with a Reduced Genome in a Multipartite Beetle Microbiome. mBio [Internet]. 2020 Feb 25 [cited 2020 Mar 31];11(1). Available from: https://mbio.asm.org/content/11/1/e02430-19

17. Schultz TR, Brady SG. Major evolutionary transitions in ant agriculture. Proc Natl Acad Sci. 2008 Apr 8;105(14):5435–40.

18. Currie CR, Scott JA, Summerbell RC, Malloch D. Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. Nature. 1999 Apr;398(6729):701–4.

19. Poulsen M, Bot ANM, Currie CR, Nielsen MG, Boomsma JJ. Within-colony transmission and the cost of a mutualistic bacterium in the leaf-cutting ant Acromyrmex octospinosus. Funct Ecol. 2003;17(2):260–9.

20. Cafaro MJ, Poulsen M, Little AEF, Price SL, Gerardo NM, Wong B, et al. Specificity in the symbiotic association between fungus-growing ants and protective *Pseudonocardia* bacteria. Proc R Soc B Biol Sci. 2011 Jun 22;278(1713):1814–22.

21. Van Arnam EB, Ruzzini AC, Sit CS, Currie CR, Clardy J. A Rebeccamycin Analog Provides Plasmid-Encoded Niche Defense. J Am Chem Soc. 2015 Nov 18;137(45):14272–4.

22. Van Arnam EB, Ruzzini AC, Sit CS, Horn H, Pinto-Tomás AA, Currie CR, et al. Selvamicin, an atypical antifungal polyene from two alternative genomic contexts. Proc Natl Acad Sci. 2016 Nov 15;113(46):12940–5.

23. Kim H-J, Han C-Y, Park J-S, Oh S-H, Kang S-H, Choi S-S, et al. Nystatin-like Pseudonocardia polyene B1, a novel disaccharide-containing antifungal heptaene antibiotic. Sci Rep [Internet]. 2018 Sep 11 [cited 2020 Mar 31];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6134108/

24. Sit CS, Ruzzini AC, Arnam EBV, Ramadhar TR, Currie CR, Clardy J. Variable genetic architectures produce virtually identical molecules in bacterial symbionts of fungus-growing ants. Proc Natl Acad Sci. 2015 Oct 27;112(43):13150–4.

25. Currie CR, Poulsen M, Mendenhall J, Boomsma JJ, Billen J. Coevolved Crypts and Exocrine Glands Support Mutualistic Bacteria in Fungus-Growing Ants. Science. 2006 Jan 6;311(5757):81–3.

26. Steffan SA, Chikaraishi Y, Currie CR, Horn H, Gaines-Day HR, Pauli JN, et al. Microbes are trophic analogs of animals. Proc Natl Acad Sci. 2015 Dec 8;112(49):15119–24.

27. Li H, Sosa-Calvo J, Horn HA, Pupo MT, Clardy J, Rabeling C, et al. Convergent evolution of complex structures for ant–bacterial defensive symbiosis in fungus-farming ants. Proc Natl Acad Sci. 2018 Oct 16;115(42):10720–5.

28. Marsh SE, Poulsen M, Pinto-Tomás A, Currie CR. Interaction between Workers during a Short Time Window Is Required for Bacterial Symbiont Transmission in Acromyrmex Leaf-Cutting Ants. PLoS ONE [Internet]. 2014 Jul 24 [cited 2020 Mar 31];9(7). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4110003/

29. Poulsen M, Cafaro M, Boomsma JJ, Currie CR. Specificity of the mutualistic association between actinomycete bacteria and two sympatric species of Acromyrmex leaf-cutting ants. Mol Ecol. 2005;14(11):3597–604.

30. Andersen SB, Hansen LH, Sapountzis P, Sørensen SJ, Boomsma JJ. Specificity and stability of the Acromyrmex–Pseudonocardia symbiosis. Mol Ecol. 2013;22(16):4307–21.

31. Horn HA. Small molecule dynamics and partner fidelity in an ancient host-microbe symbiosis [Internet]. 2018.; 2018 [cited 2020 Apr 6]. Available from: https://search.library.wisc.edu/catalog/9912753003102121

32. Kohlweyer U, Thiemer B, Schräder T, Andreesen JR. Tetrahydrofuran degradation by a newly isolated culture of Pseudonocardia sp. strain K1. FEMS Microbiol Lett. 2000;186(2):301–6.

33. Lee SD, Kim ES, Kang S-O, Hah YC. Pseudonocardia spinosispora sp. nov., isolated from Korean soil. Int J Syst Evol Microbiol. 2002;52(5):1603–8.

34. Li J, Zhao G-Z, Varma A, Qin S, Xiong Z, Huang H-Y, et al. An Endophytic Pseudonocardia Species Induces the Production of Artemisinin in Artemisia annua. PLoS ONE [Internet]. 2012 Dec 12 [cited 2020 Mar 31];7(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3520919/

35. Vainberg S, McClay K, Masuda H, Root D, Condee C, Zylstra GJ, et al. Biodegradation of Ether Pollutants by Pseudonocardia sp. Strain ENV478. Appl Environ Microbiol. 2006 Aug 1;72(8):5218–24.

36. Carr G, Derbyshire ER, Caldera E, Currie CR, Clardy J. Antibiotic and Antimalarial Quinones from Fungus-Growing Ant-Associated Pseudonocardia sp. J Nat Prod. 2012 Oct 26;75(10):1806–9.

37. Oh D-C, Poulsen M, Currie CR, Clardy J. Dentigerumycin: a bacterial mediator of an ant-fungus symbiosis. Nat Chem Biol. 2009 Jun;5(6):391–3.

38. Schorn MA, Alanjary MM, Aguinaldo K, Korobeynikov A, Podell S, Patin N, et al. Sequencing rare marine actinomycete genomes reveals high density of unique natural product biosynthetic gene clusters. Microbiol Read Engl. 2016;162(12):2075–86.

39. McDonald BR, Chevrette MG, Klassen JL, Horn HA, Caldera EJ, Wendt-Pienkowski E, et al. Biogeography and Microscale Diversity Shape the Biosynthetic Potential of Fungus-growing Ant-associated Pseudonocardia. bioRxiv. 2019 Feb 10;545640.

40. Caldera EJ, Currie CR. The Population Structure of Antibiotic-Producing Bacterial Symbionts of Apterostigma dentigerum Ants: Impacts of Coevolution and Multipartite Symbiosis. Am Nat. 2012 Nov;180(5):604–17.

41. Holmes NA, Innocent TM, Heine D, Bassam MA, Worsley SF, Trottmann F, et al. Genome Analysis of Two Pseudonocardia Phylotypes Associated with Acromyrmex Leafcutter Ants Reveals Their Biosynthetic Potential. Front Microbiol [Internet]. 2016 Dec 26 [cited 2020 Mar 31];7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5183585/

42. Andersen SB, Yek SH, Nash DR, Boomsma JJ. Interaction specificity between leaf-cutting ants and vertically transmitted Pseudonocardia bacteria. BMC Evol Biol. 2015 Feb 25;15(1):27.

43. Chevrette MG, Carlson CM, Ortega HE, Thomas C, Ananiev GE, Barns KJ, et al. The antimicrobial potential of Streptomyces from insect microbiomes. Nat Commun. 2019 Jan 31;10(1):516.

44. Currie CR, Bot ANM, Boomsma JJ. Experimental evidence of a tripartite mutualism: bacteria protect ant fungus gardens from specialized parasites. Oikos. 2003;101(1):91–102.

45. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1;34(17):i884–90.

46. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, et al. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. In: Deng M, Jiang R, Sun F, Zhang X, editors. Research in Computational Molecular Biology. Berlin, Heidelberg: Springer; 2013. p. 158–70. (Lecture Notes in Computer Science).

47. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ. 2015 Oct 8;3:e1319.

48. Campbell JH, O'Donoghue P, Campbell AG, Schwientek P, Sczyrba A, Woyke T, et al. UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. Proc Natl Acad Sci U S A. 2013 Apr 2;110(14):5540–5.

49. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 2013 Apr 1;30(4):772–80.

50. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014 May 1;30(9):1312–3.

51. Menardo F, Loiseau C, Brites D, Coscolla M, Gygli SM, Rutaihwa LK, et al. Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity. BMC Bioinformatics. 2018 May 2;19(1):164.

52. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018 Dec;9(1):5114.

53. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinforma Oxf Engl. 2014 Jul 15;30(14):2068–9.

54. Melnyk RA, Hossain SS, Haney CH. Convergent gain and loss of genomic islands drive lifestyle changes in plant-associated Pseudomonas. ISME J. 2019;13(6):1575–88.

55. Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLOS Comput Biol. 2018 Feb 5;14(2):e1005958.

56. Goldstein S, Beka L, Graf J, Klassen JL. Evaluation of strategies for the assembly of diverse bacterial genomes using MinION long-read sequencing. BMC Genomics. 2019 Dec;20(1):23.

57. Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, et al. A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol. 2020 Jan;16(1):60–8.

58. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft JJJ, et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res. 2020 Jan 8;48(D1):D454–8.

59. Barke J, Seipke RF, Grüschow S, Heavens D, Drou N, Bibb MJ, et al. A mixed community of actinomycetes produce multiple antibiotics for the fungus farming ant Acromyrmex octospinosus. BMC Biol. 2010 Aug 26;8(1):109.

60. Zucchi TD, Guidolin AS, Cônsoli FL. Isolation and characterization of actinobacteria ectosymbionts from Acromyrmex subterraneus brunneus (Hymenoptera, Formicidae). Microbiol Res. 2011 Jan 20;166(1):68–76.

61. Armitage SAO, Broch JF, Marín HF, Nash DR, Boomsma JJ. Immune Defense in Leaf-Cutting Ants: A Cross-Fostering Approach. Evolution. 2011;65(6):1791–9.

62. Zhang MM, Poulsen M, Currie CR. Symbiont recognition of mutualistic bacteria by Acromyrmex leaf-cutting ants. ISME J. 2007 Aug;1(4):313–20.

63. Poulsen M, Cafaro MJ, Erhardt DP, Little AEF, Gerardo NM, Tebbets B, et al. Variation in Pseudonocardia antibiotic defence helps govern parasite-induced morbidity in Acromyrmex leaf-cutting ants. Environ Microbiol Rep. 2010 Aug;2(4):534–40.

64. Barke J, Seipke RF, Yu DW, Hutchings MI. A mutualistic microbiome. Commun Integr Biol. 2011;4(1):41–3.

65. Lo W-S, Huang Y-Y, Kuo C-H. Winding paths to simplicity: genome evolution in facultative insect symbionts. FEMS Microbiol Rev. 2016 Nov;40(6):855–74.

66. Li M, Yang Y, He Y, Mathieu J, Yu C, Li Q, et al. Detection and cell sorting of Pseudonocardia species by fluorescence in situ hybridization and flow cytometry using 16S rRNA-targeted oligonucleotide probes. Appl Microbiol Biotechnol. 2018 Apr 1;102(7):3375–86.

67. Cunning R, Silverstein RN, Baker AC. Investigating the causes and consequences of symbiont shuffling in a multi-partner reef coral symbiosis under environmental change. Proc R Soc B Biol Sci. 2015 Jun 22;282(1809):20141725.

68. McMullen JG, Peterson BF, Forst S, Blair HG, Stock SP. Fitness costs of symbiont switching using entomopathogenic nematodes as a model. BMC Evol Biol. 2017 Apr 17;17(1):100.

Chapter 3: Gut Microbial and Metabolic Responses to *Salmonella enterica* Serovar Typhimurium and *Candida albicans*

Jennifer R. Bratburd, Caitlin Keller, Eugenio Vivas, Erin Gemperline, Lingjun Li, Federico E. Rey, Cameron R. Currie

Reproduced from Bratburd JR, Keller C, Vivas E, Gemperline E, Li L, Rey FE, Currie CR. 2018. Gut microbial and metabolic responses to *Salmonella enterica* serovar Typhimurium and *Candida albicans*. mBio 9:e02032-18. https://doi.org/10.1128/mBio.02032-18.

3.1 Abstract

The gut microbiota confers resistance to pathogens of the intestinal ecosystem, yet the dynamics of pathogen-microbiome interactions and the metabolites involved in this process remain largely unknown. Here, we use gnotobiotic mice infected with the virulent pathogen *Salmonella enterica* serovar Typhimurium or the opportunistic pathogen Candida albicans in combination with metagenomics and discovery metabolomics to identify changes in the community and metabolome during infection. To isolate the role of the microbiota in response to pathogens, we compared mice monocolonized with the pathogen, uninfected mice "humanized" with a synthetic human microbiome, or infected humanized mice. In *Salmonella*-infected mice, by 3 days into infection, microbiome community structure and function changed substantially, with a rise in Enterobacteriaceae strains and a reduction in biosynthetic gene cluster potential. In contrast, *Candida*-infected mice had few microbiome changes. The LC-MS metabolomic fingerprint of the cecum differed between mice monocolonized with either pathogen and humanized infected mice. Specifically, we identified an increase in glutathione disulfide, glutathione cysteine disulfide, inosine 5'-monophosphate, and hydroxybutyrylcarnitine in mice

infected with *Salmonella* in contrast to uninfected mice and mice monocolonized with Salmonella. These metabolites potentially play a role in pathogen-induced oxidative stress. These results provide insight into how the microbiota community members interact with each other and with pathogens on a metabolic level.

Importance

The gut microbiota is increasingly recognized for playing a critical role in human health and disease, especially in conferring resistance to both virulent pathogens such as *Salmonella*, which infects 1.2 million people in the United States every year (E. Scallan, R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, et al., Emerg Infect Dis 17:7–15, 2011, https://doi.org/10.3201/eid1701.P11101), and opportunistic pathogens like Candida, which causes an estimated 46,000 cases of invasive candidiasis each year in the United States (Centers for Disease Control and Prevention, Antibiotic Resistance Threats in the United States, 2013, 2013). Using a gnotobiotic mouse model, we investigate potential changes in gut microbial community structure and function during infection using metagenomics and metabolomics. We observe that changes in the community and in biosynthetic gene cluster potential occur within 3 days for the virulent *Salmonella enterica* serovar Typhimurium, but there are minimal changes with a poorly colonizing *Candida albicans*. In addition, the metabolome shifts depending on infection status, including changes in glutathione metabolites in response to *Salmonella* infection, potentially in response to host oxidative stress.

3.2 Introduction

Symbiotic microbes help shape the biology of plants and animals (1). In humans, gut microbes modulate nutrition and immune function and are correlated with an increasing number

of metabolic and neurological health and disease states (2, 3). The human gastrointestinal tract harbors the largest fraction of microbial life in the body, estimated to range from 108 to 1010 bacteria per gram in the ileum and stool, respectively (4). Bacteria are the dominant taxa in the human gut microbiome, with the most abundant lineages belonging to the phyla Bacteroidetes and Firmicutes. Nevertheless, these communities are highly diverse and include viruses, archaea, fungi, and protists (5–8), and all combined contain 150 times as many genes as the human genome (9). In a healthy state, the human gut microbiome is relatively stable over time (10, 11). Major disruption of the gut microbiome is associated with infections by a number of serious human pathogens, such as *Clostridium difficile*, vancomycin-resistant *Enterococcus* (VRE), and *Salmonella enterica* (12–14).

Preventing exogenous microbes from colonizing the human intestine is critical to the host maintaining a stable and healthy gut microbiome. The role of the microbiome in preventing pathogens from invading the gut has been recognized since the 1950s, when pretreatment with antibiotics was shown to drop the infectious dose of *Salmonella enterica* 100,000-fold (15). Gut microbes confer colonization resistance by outcompeting pathogens for nutrients, priming the host immune system, and directly targeting other microbes with metabolites (16). Several examples of metabolites produced or modified by the microbiota that inhibit pathogens include short-chain fatty acids, secondary bile acids, and modified compounds from the diet (17–19). In addition, some members of the microbiota can create compounds to respond selectively to pathogen infection (20). The gut microbiota has the potential to make a wide variety of novel natural products, and many of the large biosynthetic gene clusters encoding natural products are found in relatively small genomes, indicative of an ecological role for the products (21).

Experiments using gnotobiotic mice with and without human microbiota, in combination with metagenomic and metabolomic approaches, can provide insight on the structure and function of the gut microbiota during pathogen invasion. Gnotobiotic mice are a mammalian model system in which defined microbiomes can be used in a controlled environment. Various metabolomics techniques, including nuclear magnetic resonance and chromatography-mass spectrometry, have been used for large-scale characterization of metabolite changes as a result of microbiome colonization, illustrating the impact of the microbiota on not only intestinal metabolism but also global systems (22, 23). Furthermore, liquid chromatography-mass spectrometry (LC-MS) can help to characterize metabolite changes due to disturbances in the microbiome (24, 25) and to screen for novel secondary metabolites and natural products in bacterial systems (26, 27).

Here we examine colonization resistance in the humanized (HUM) mouse model. Specifically, we perform experimental infection with *Salmonella enterica* serovar Typhimurium and *Candida albicans* in HUM mice and in germfree (GF) mice. *Salmonella enterica* Typhimurium is a disruptive pathogen that causes massive inflammation to outcompete the native microbiota in mice and human models (28–30). *Candida albicans* can cause low-grade inflammation, but in contrast to *Salmonella enterica* Typhimurium is considered a commensal and occasional opportunistic pathogen in the GI tract (31–34). Nevertheless, *C. albicans* has been shown to colonize GF and antibiotic-treated adult mice (33, 35, 36), which appear otherwise resistant, suggesting that gut microbiota play a role in preventing *Candida* colonization in mice and humans. In this study, we investigate how these pathogens alter the structure of the human gut microbiome, the biosynthetic gene cluster potential, and the metabolites produced in a healthy or infected state. We cross the presence and absence of the

microbiome with the presence and absence of pathogen infection, using either *S. enterica* Typhimurium or *C. albicans*. To characterize strain-level diversity that is not resolvable with 16S rRNA gene sequencing, we use shotgun metagenomics on fecal samples over 3 days of infection. We also identify the capacity of community members to produce novel antimicrobials through the biosynthetic gene clusters embedded in bacterial genomes. Further, we characterize metabolites using LC-MS for relative quantification and discovery metabolomics in the host cecum during infection and validate the identifications of several specific metabolites with commercial standards.

3.3 Results

3.3.1 Infection severity in mice with and without microbiota.

Germfree mice, 8 to 12 weeks old, were kept germfree or colonized via oral gavage with a synthetic human community for 2 weeks, and then infected with *Salmonella enterica* Typhimurium or *Candida albicans* (Fig. 1A). All infected mice showed presence of pathogens in fecal samples by growth on selective media. Prior to infection, the mice weighed on average 29.8 g ± 2.3 (mean ± SD). GF mice infected with *Salmonella* (n = 6), henceforth referred to as monocolonized *Salmonella* mice, lost an average body mass of 2.0 ± 1.4 g or 6.8% ± 4.7% within 12 h postinfection. Due to severity of symptoms, three monocolonized Salmonella mice were sacrificed 12 h postinfection, and the remaining monocolonized *Salmonella* mice and one HUM mouse infected with *Salmonella* were sacrificed within 24 h of infection. HUM mice infected with *Salmonella* surviving 3 days into infection (n = 5) lost an average of 4.2 ± 0.6 g or 14.3% ± 1.7%, a significant loss in comparison to weight change from both the monocolonized and HUM mice infected with *Candida* (Mann-Whitney U test, Bonferroni corrected, $P < 0.05$). The monocolonized *Candida* mice (n = 6) gained on average 0.2 ± 0.3 g or 0.8% ± 1.1% of their

original weight, and the *Candida*-infected HUM mice (n = 6) gained on average 0.7 ± 0.5 g or 2.0% ± 1.8% of their original weight. There was no statistically significant difference in the change in weight for the monocolonized *Candida* mice compared to the HUM mice infected with *Candida* by the endpoint of the experiment, 3 days of infection.

3.3.2 Microbial community shifts in response to infection.

We conducted Illumina-based metagenomic sequencing on DNA from fecal pellets collected throughout infection. Each sample had on average 407,535 reads (SD = 63,381), ranging from 295,235 to 523,271 reads. The average number of reads with at least one reported alignment was 385,882 ± 96,477, or 95% of reads per sample. Prior to infection, the most abundant strains, making up over half of the relative abundance in the metagenomes from all groups, were *Bacteroides cellulosilyticus* DSM14838, *Subdoligranulum variabile*, *Bacteroides cellulosilyticus* WH2, *Akkermansia muciniphila*, and *Clostridium bolteae* with an average relative abundance of 15.1%, 14.1%, 9.1%, 7.8%, and 6.5%, respectively (Fig. 2A). By day three in the *Salmonella*-infected HUM mice, most of the communities were dominated by *Salmonella* and other various *Enterobacteriaceae* strains from the original inoculum. Furthermore, diversity significantly decreased in Salmonella-infected mice (see Fig. S1 in the supplemental material). Prior to infection, these strains (*C. youngae, P. penneri, E. cancerogenu*s, and *E. fergusonii*) in total represented an average relative abundance of 0.2%. In the metagenomes from two mice, we observed an increase in the reads mapping to *Enterobacter cancerogenu*s, up to 26.4% and 26.6% of the community, along with a smaller increase in *Proteus penneri*. One mouse had an increase in *Escherichia fergusonii* to 22.9% of the metagenome, while it remained below 1% of the metagenome in all the other mice. In another mouse, *Citrobacter youngae* reads increased to 15.2%, while in other mice *C. youngae* reads remained below 7.9%. After excluding *Salmonella*

reads, we continued to observe a large shift in the relative abundance of community members. Using principal component analysis (PCA), we show large separation of the HUM *Salmonella* microbiome communities, 3 days postinfection, from a tight cluster of all other time points and treatments, with the first component explaining 31.4% of the variation (Fig. 2B).

In all *Candida*-infected HUM mice, less than 1% of reads mapped to the *Candida albicans* SC5314 reference genome. The metagenome of this group was not significantly different from uninfected HUM mice. The community structure remained fairly consistent over the infection period, although there was some variation in strain relative abundance over time (Fig. 2). The largest change in any individual strain's relative abundance was an 8.4% increase in *Subdoligranulum variabile* in one mouse from 1 day postinfection to 3 days postinfection.

3.3.3 Prevalence of biosynthetic gene clusters within genomes and metagenomes.

In total, from the genomes of the human microbiome used in this study, using antiSMASH 4.0 (37), we detected 1,081 biosynthetic gene clusters (BGCs). Of these clusters, when grouped together using BiG-SCAPE with a cutoff distance of 30 calculated based on a weighted combination of Jaccard, domain sequence similarity, and adjacency index, we identified 128 cluster nodes in 51 groups. The remaining 953 BGCs did not form any groupings with each other. Based on antiSMASH-predicted classifications, most clusters were classified as other, which included putative clusters (486), fatty acids (117), fatty acid-saccharide combined clusters (22), aryl polyenes (14), siderophores (4), and resorcinol (3). Another large category was saccharides (345), followed by 62 ribosomally synthesized and posttranslationally modified peptides (RiPPs), a group that includes bacteriocins, sactipeptides, lantipeptides, and thiopeptides. We also found 20 nonribosomally synthesized peptide clusters and one hybrid polyketide-NRPS cluster in *Desulfovibrio piger* (Table S1).

We found significant differences in the percentages of total metagenomic reads mapping to BGCs in *Salmonella*-infected HUM mice prior to infection versus 3 days postinfection (Wilcoxon $P < 0.05$, corrected with Benjamini-Hochberg), excluding reads mapping to BGCs from *Salmonella* itself. Saccharides, lantipeptides, aryl polyenes, sactipeptides, fatty acids, fatty acid-saccharides, terpenes, and putative clusters were significantly reduced, while thiopeptides significantly increased 3 days postinfection (Fig. 3). The majority of non-*Salmonella* reads mapping to thiopeptide clusters mapped to *Citrobacter youngae*, *Enterobacter cancerogenus*, *Proteus penneri,* and *Escherichia fergusonii,* consistent with the overall increase relative abundance in *Enterobacteriaceae* described above.

3.3.4 Differential metabolomics during infection and novel metabolite potential.

Analysis of the LC-MS results with Compound Discoverer (Thermo Fisher Scientific) resulted in the grouping of 8,613 merged features (chromatographic peaks) into 8,259 putative compounds. The compounds detected from the cecum samples of one or more mice from each treatment group totaled 3,254 for the monocolonized *Candida* mice, 3,696 compounds for the monocolonized *Salmonella* mice, 3,349 compounds for the uninfected HUM mice, 2,924 compounds for the HUM mice infected with *Candida*, and 2,815 compounds for the HUM mice infected with *Salmonella*.

LC-MS m/z values and relative intensities from cecum contents showed separation of samples with PCA. Two components were able to explain 67.7% of the variance (Fig. S2). Using partial least-squares discriminant analysis (PLS-DA), we observed distinct separation of all groups with two components ($R2 = 0.70799$, $Q2 = 0.66183$ for component 1 and $R2 = 0.85972$ and $Q2 = 0.81188$ for component 2; Fig. 4A). Using permutation testing of the PLS-DA, we obtained statistical significance ($P < 0.001$) for 1,000 permutations. The outliers in the

Salmonella-infected HUM mouse group were from two technical replicates of one sample that had to be sacrificed 24 h into infection. We also found distinct patterns for different groups of metabolites (Fig. 4B), which indicate similar patterns between uninfected HUM and *Candida*-infected HUM mice compared to monocolonized infected mice and HUM *Salmonella* mice. Additionally, we identified numerous features overrepresented in the monocolonized groups compared to the HUM groups (Fig. S3).

To examine metabolites potentially produced by the microbiome in response to infection, we looked for metabolites that were typically not found in pathogen-monocolonized mice (absent in at least 8 of 12 samples, representing 6 biological replicates with 2 technical replicates each) and were at least 1.5-fold higher in abundance in infected HUM mice compared to the highest normalized area of the controls (HUM mice with no infection). Using these guidelines, we narrowed our metabolites of interest to 31 out of 8,085 features detected overall. We detected 22 features in higher abundance in HUM *Salmonella*-infected mice. In HUM *Candida*-infected mice, we found 10 features of interest based on the above criteria. One metabolite (m/z 347.0626, retention time 1.05 min) appeared to be shared between the lists, and also had matching tandem MS fragmentation from both infection groups. This metabolite had similar MS/MS to 3′AMP and 2′AMP standards, but the experimental retention time did not match that of the standards (1.37 min for 3′AMP and 2.22 min for 2′AMP). From the 31 selected compounds of interest, only 6 from HUM *Salmonella* and 4 from *Candida* infection had putative identifications based upon accurate mass matching to KEGG, HMDB, or AntiBase, leaving a remaining total of 21 potentially novel compounds (Table S2). In silico fragmentation with MetFrag (38) was performed using MS/MS spectra obtained on the targets. If the top peaks in the experimental MS/MS were explained by the *in silico* fragmentation, then standards were

obtained to confirm the identification. Using this procedure, we identified glutathione disulfide, glutathione cysteine disulfide, inosine 5′-monophosphate, and hydroxybutyrylcarnitine as compounds upregulated from the HUM *Salmonella* group (Fig. S4). Although the in silico fragmentation approach worked well for the targets with KEGG matches, the increasing number of compounds in the more inclusive databases made it difficult to find putative identifications with MS/MS for targets that did not have matches to the KEGG databases.

3.4 Discussion

Understanding how microbial communities change in response to perturbation is crucial for health, not only because the microbiota can protect the host against pathogenic microbes but also because changes in the gut microbiota have been associated with multiple health conditions (39). Increasingly it has been recognized that pathogenicity and virulence can depend on the context of specific microbe-microbe interactions or the whole community, indicating the importance of studying pathogen-microbiome interactions (40, 41). In this study, we compare how two pathogenic perturbations affect the structure and function of human gut microbiota in a gnotobiotic mouse model. We find that during infection with *Salmonella*, the structure and functional capacity of the microbiota change. Corresponding to these changes, we see significant changes in metabolites before versus during infection that vary with and without the human microbiota.

Our infection experiments revealed significant differences among treatments as measured by weight loss. *Candida*-infected mice had weights that remained around their baseline starting weight. While we did isolate CFUs of *Candida* from mouse feces using media with antibiotics, indicating that viable yeast cells passed through the host, reads mapping to *Candida* from the metagenomic data were at or below the limit of detection, suggesting that *Candida* did not

readily colonize these mice. Alternatively, the lack of fungal DNA may be influenced by our DNA extraction method (42). In contrast, *Salmonella*-infected mice lost significantly more weight than *Candida*-infected mice by 3 days into infection, regardless of microbiome presence or absence. GF mice infected with *Salmonella* were moribund within 24 h, while HUM mice infected with *Salmonella* were able to survive until the end of the 3 days, with the exception of one mouse, indicative of the protective effects of the microbiota against *Salmonella*.

Salmonella infection perturbed the microbiota and led to an increase in the relative abundance of different *Enterobacteriaceae*, whereas Candida did not. Prior to infection, the microbiota contained similar dominant taxa including Bacteroidetes and Firmicutes with relatively few Gammaproteobacteria. During *Salmonella* infection in humanized mice, the metagenomic data indicated an increase in the relative abundance of *Enterobacteriaceae* (including strains besides *Salmonella*). This result is consistent with previous work examining changes in gut microbial communities during *Salmonella* infection (28, 43, 44), and resembles increases in *Enterobacteriaceae* during antibiotic treatment (13), both of which may ultimately be driven by the oxygenation of the gut (45). These changes may represent a bloom of closely related strains or a reduction in the size of the bacterial community overall. Although *Enterobacteriaceae* increased in the samples, which particular strains increased appeared stochastic. Some of the variation may be due to read mapping of conserved genes to closely related strains; however, we saw similar results using different read mapping programs (Bowtie and Burrows-Wheeler Algorithm) and using parameters to exclude non-uniquely mapping reads. Given that these strains may compete with *Salmonella* over electron acceptors and trace elements, further investigation on these dynamic interactions is warranted (46, 47). The stochasticity may also reflect the general instability of the community. While *Salmonella*

dramatically perturbs the community, *Candida* did not seem to readily colonize the mice, and although some changes occurred in the microbial communities, these fluctuations are within the range of natural variation.

The synthetic human microbiome used in this study contained many biosynthetic gene clusters, and the potential functional capacity changed with infection treatment. In our input strains we found potential for unknown biosynthetic gene clusters, including RiPPs, NRPS clusters, and many putative clusters. This fits with previous observations; biosynthetic gene clusters are common in human gut microbiota and anaerobic bacteria (21, 48). Metagenomic analysis indicated a decrease in most cluster types during *Salmonella* infection, which likely reflects a drop in community diversity. One exception was the increase in reads mapping to gene clusters involved in thiopeptide biosynthesis, which was increased even after removing reads mapping to *Salmonella*'s own thiopeptide biosynthesis cluster. Thiopeptides are a class of peptide antibiotics that target Gram-positive bacteria (49). Since *Salmonella* is Gram-negative and has one putative thiopeptide BGC of its own, it seems unlikely that these thiopeptide clusters, if produced, would target *Salmonella.* Other possibilities are that if produced, these secondary metabolites encoded by clusters might add to the community instability, or that these genes are not transcribed or translated. Alternatively, this result may suggest that the pathogen-induced disruption in the microbiome helps diminish members that would have been capable of producing BGC products. Further research will be needed to characterize what role, if any, these BGCs play during infection.

Our discovery metabolomics showed differences in the metabolites present in the mouse cecum based on presence of microbiome as well as infection. For example, the metabolomes of *Salmonella*-infected, *Candida*-infected, and uninfected mouse ceca grouped separately on PLS-

DA analysis, suggesting distinct metabolic responses between a virulent bacterial pathogen and opportunistic fungal pathogen. The changes in overall metabolites based on gut microbiota support previous research comparing germfree and colonized mice and mice with different gut microbiome donors (50). We found more putative metabolites of interest (based on higher abundance in HUM infected mice and generally absent in GF mice) from *Salmonella*-infected mice than *Candida*-infected mice. Previous studies investigating global metabolomics in *Salmonella* infections have focused on the hosts with conventional mouse microbiota, finding disruptions in host hormone pathways (51), changes in common microbial metabolites, including trimethylamine N-oxide (TMAO) and hippurate (52), and changes in sugar moieties (43). Our study differed from these previous studies in that we used gnotobiotic mice to specifically focus on metabolites produced when human-associated gut microbiota strains were exposed to pathogens. While using native microbiota to look for pathogen interactions is valuable especially in an ecological context, the humanized mouse model enables exploration of potentially distinct chemical interactions between human microbiota strains and human pathogens (53). Furthermore, human gut microbiota extracts have been previously shown to inhibit virulence of *Salmonella in vitro* (40). Mice monocolonized with pathogens serve as key controls that allowed us to focus on compounds apparently made by the microbiota during infection rather than overall host changes. Nevertheless, the possibility exists that we may detect metabolites made by *Salmonella* in response to gut microbiota in our experiments or metabolites that differ due to GF mice exhibiting colitis rather than the typical systemic typhoid-like infection (54). In addition, we scanned for molecular features with an m/z greater than 200, to avoid discovery of smaller commonly made microbial metabolites. In our metabolites of interest from humanized infection conditions, we had many molecular features that were not identified with KEGG, HMDB, or

AntiBase, potentially indicating novel metabolites. One drawback in studying these metabolite interactions *in vivo* is the challenge in isolating individual novel molecules from a complex mixture, even in a well-described community with full genomes (55), as we were unable to match known and predicted metabolites to a majority of our target m/z values. Although work is being done to increase MS/MS databases for natural products (56), identifying natural products is still challenging, as many natural product databases, including AntiBase, are not MS compatible.

We were able to identify a few metabolites specific to the humanized *Salmonella*-infected group, including two metabolites in the glutathione pathway. In particular, we identified glutathione disulfide and glutathione cysteine disulfide in higher abundance in humanized *Salmonella*-infected mice. *Salmonella* infection triggers vast amounts of oxidative stress (57), and glutathione metabolism is important for protection against oxidative stress (58). Changes in genes encoding antioxidant proteins have also been identified in humans exposed to *Salmonella enterica* serovar Typhi (59). Further, glutathione cysteine disulfide has been shown to reduce colonic lesions in a mouse model of colitis (60). Previous work indicates that germfree mice have a disrupted glutathione metabolism relative to conventional mice (61). It remains to be seen whether experimentally manipulating glutathione metabolite amounts affects *Salmonella* infections *in vivo*, and to what extent different gut microbes contribute to the glutathione pool. In contrast to the hypothesis that microbes may make specific metabolites that inhibit certain pathogens, this evidence suggests more generalized responses to certain kinds of dysbiosis, such as oxidative stress (62). The possibility of microbial metabolites with specific responses to pathogens cannot be eliminated; however, many metabolites remain unidentified, and the roles of

those identified are unclear. Further characterization of microbial metabolites made during infection is necessary to identify these responses.

Colonization resistance conferred by the microbiota helps the host resist a variety of pathogens, including *Salmonella*. Understanding the complex interactions between the host, microbiota, and pathogens will enable better microbiome based-therapies, from fecal microbiota transplants to microbiota-derived compounds (63, 64). Combining gnotobiotic mice with genomics and metabolomics has allowed us to interrogate changes in community composition and function during infection in an unbiased manner and demonstrates distinct metabolic responses to a virulent or opportunistic pathogen.

3.5 Materials and Methods

3.5.1 Human gut microbiota and pathogens.

For our synthetic human microbiome gut community, we used a collection of previously obtained isolates cultured from human fecal samples and maintained in long-term storage in the Rey lab at the University of Wisconsin-Madison. Bacterial isolates (Table S3) were grown from glycerol stock on Mega Medium (65), which was filter sterilized and held in a Coy anaerobic chamber (5% H2, 20% CO2, and 75% N2). An even mix from each bacterial culture was inoculated into each anaerobic tube. From stock cultures, *Salmonella enterica* Typhimurium ATCC 14028 was grown aerobically overnight in LB broth at 37°C, while *Candida albicans* K1 was grown on Sabouraud dextrose agar (SDA).

Gnotobiotic mice and experimental infections. The University of Wisconsin-Madison Animal Care and Use Committee approved protocols used in mouse experiments. GF male C57BL/6J mice were maintained in gnotobiotic isolators until 8 to 12 weeks of age with 12-h light cycle and sterilized food and water ad libitum. These GF mice were then randomly assigned

to 1 of 5 treatment groups, moved to out-of-the-isolator gnotobiotic cages in autoclaved filter-top cages, and subsequently gavaged in a biosafety cabinet using aseptic technique (66). Mice were housed 3 per cage, with a total of 6 mice per group.

To humanize mice, GF mice were colonized via oral gavage with 0.2 ml mixed bacterial culture as shown in Table S3. All HUM mice were given the same inoculum, where bacteria were mixed with roughly similar proportions. Prior to infection, HUM mice were given 2 weeks to allow stabilization of the community. For mouse infections, mice were inoculated via oral gavage with 0.2 ml of overnight culture of *Salmonella enterica* Typhimurium ATCC 14028 or *Candida albicans* K1. Humanization and infection treatments were performed in a biosafety cabinet using aseptic technique (66). Mice were sacrificed 3 days postinfection or earlier depending on symptom severity and weight loss. Cecal contents were collected, flash frozen and stored at −80°C until processing. We selected cecum contents for LC-MS due to their high microbial loads and proximity to the distal ileum to which *Salmonella* localizes (28, 67).

Salmonella and *Candida* quantification was performed by serial dilutions of fecal samples in phosphate-buffered saline, followed by plating for quantification for *Salmonella* on xylose lysine deoxycholate (XLD) agar, and for *Candida* on SDA with chloramphenicol and gentamicin. Fecal samples from uninfected mice showed no growth on the SDA plates, as well as no growth of black colonies on XLD plates, indicating no colonies capable of metabolizing thiosulfate into hydrogen sulfide as *Salmonella* does.

3.5.2 Metagenomics.

To characterize the gut microbiome of HUM mice, we conducted metagenomics using Illumina MiSeq. Genomic DNA was extracted from fecal pellets following the Turnbaugh et al. protocol (68). Briefly, the protocol is as follows: to each frozen fecal pellet, we added 500 µl of

extraction buffer (200 mM Tris, 200 nM NaCl, 20 mM EDTA), 210 µl 20% SDS, 500 µl phenol-chloroform, 500 µl 0.1-mm zirconia-silica beads, and one 3.2-mm stainless steel bead. Cells were beaten for 3 min at room temperature. To remove contaminants, the Wizard SV Gel and PCR Clean-up kit was used. DNA library preparation and sequencing were done at the University of Wisconsin-Madison Biotechnology Center. Samples were prepared with the TruSeq Nano DNA LT Library Prep kit (Illumina Inc., San Diego, CA, USA) with minor modifications. After shearing samples with a Covaris M220 Ultrasonicator (Covaris Inc., Woburn, MA, USA), samples were size selected for an average insert size of 550 bp using SPRI bead-based size exclusion, and then libraries were standardized to 2 nM. Sequencing was done using single ends on the Illumina MiSeq sequencer with a 50-bp (v2) sequencing cartridge.

Metagenomic data were preprocessed using BBDuk (https://sourceforge.net/projects/bbmap/) to trim adapters, remove phi-X contamination, and quality trim reads to Q10. We analyzed the reads using the COPROseq (Community profiling by sequencing) pipeline (69), which mapped reads to reference genomes using Bowtie version 1.0 (70), and normalized reads based on genome length. We also compared read mapping using the Burrows-Wheeler Alignment tool to verify that reads mapped consistently (71). Reference genomes were obtained from NCBI. Diversity was analyzed using the vegan package in R with a Kruskal-Wallis test. Biosynthetic gene clusters were identified using antiSMASH 4.0 (37). Gene clusters were then grouped by similarity using BiG-SCAPE (J. Navarro-Muñoz et al., unpublished data; https://git.wageningenur.nl/medema-group/BiG-SCAPE). Data were analyzed and figures produced in R. Statistical testing was done using a Wilcoxon rank sum test (Mann-Whitney U test) with a Benjamini-Hochberg correction.

3.5.3 Metabolomics.

All chemicals were obtained from Fisher Scientific unless otherwise noted. Mouse cecum samples were placed in 10-ml PTFE tubes for extraction with a methanol-chloroform/water extraction. Three parts methanol, 1 part chloroform, and 4 parts water (Milli-Q system, Millipore, Billerica, MA) were added, in order, to each sample (total volume, 4 ml) and centrifuged for 20 min at 4,575 × g at 4°C. The aqueous fraction was removed, and 4 parts methanol were then added. After brief vortexing, samples were centrifuged for 5 min at 1,500 × g and 4°C. The organic layer was removed. Samples were dried in a SpeedVac and stored at −80°C. To clean up the sample, the aqueous fraction was further processed with a 3-kDa molecular weight cutoff (MWCO) (Amicon Ultra, Millipore). The MWCO device was rinsed with 0.2 ml 0.1 M NaOH and 0.5 ml 50/50 methanol-water. The sample was loaded in 0.5 ml 50/50 methanol-water and rinsed with 0.1 ml 50/50 methanol-water. All centrifugations occurred at 14,000 × g until the rinse or sample was through the device. The MWCO flowthrough was dried with a SpeedVac and stored at −80°C until analysis.

Aqueous samples were resuspended in optima-grade water at a concentration of 10 mg/ml. A Dionex Ultimate 3000 UHPLC system (Thermo Scientific, Waltham, MA, USA) and a Cortecs C18 column (2.1-mm internal diameter × 100-mm length, 1.6-µm particle size; Waters, Milford, MA, USA), equipped with a corresponding guard column were used to separate the samples. The column temperature was 35°C, and the mobile phases were optima-grade water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). The separation occurred with a 35-min gradient at a flow rate of 0.3 ml/minutes with the following conditions: 0 to 5 min, 1% B; 5 to 10 min, linear gradient from 1% to 3% B; 10 to 18 min, linear gradient from 3% to 40% B; 18 to 22 min, linear gradient from 40% to 80% B; 22 to 27 min, column cleaning at 95% B; and 27 to 35 min, reequilibration at 1% B. The injection volume was 3 µl and the samples

were kept at 10°C during analysis. Metabolite MS data were acquired on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Waltham, MA, USA), which was equipped with an ESI source and operated in positive ion mode with a scan range of m/z 200 to 1,700. The MS parameters were as follows: 70,000 resolution, 1E6 AGC, and 100-ms maximum injection time.

3.5.4 Metabolomics data analysis.

Relative quantification of the metabolomics data for the different sample types was performed with Compound Discoverer software (Thermo Scientific, Waltham, MA, USA). Spectra underwent retention time alignment (adaptive curve 5 ppm, 1-min tolerances), detection of unknown compounds (5 ppm, 30 intensity threshold, 3 S/N threshold, 1,000,000 minimum peak intensity), and grouping of unknown compounds (5 ppm, 0.05 retention time tolerance). The Compound Discoverer workflow also included fill gaps, mark background, predict compositions, ChemSpider search, normalize areas (constant sum), merge features, and differential analysis. To isolate metabolites of interest, m/z values detected in the blanks or in more than 4 of 12 replicates in either of the germfree infected conditions were removed. Additionally, m/z were selected if they showed 1.5-fold upregulation in 8 of 12 replicates of the infected humanized group, with the ratios being calculated from the control with the highest normalized area. MetaboAnalyst (72, 73) was used for further statistical analysis after exporting m/z values, retention time, and normalized areas from Compound Discoverer. Data were filtered with an interquantile range (IQR) estimate and log transformed. Heatmaps were produced using Pearson and Ward clustering.

Compound identification.MS/MS spectra for the compounds on the target lists for both infections were collected on the Dionex UltiMate 3000 UHPLC and Q-Exactive instrument described above. The injection volume was 20 µl. An inclusion list was used for the targets with

a retention time window of ± 0.7 min. All charge states and salt adducts observed in the Compound Discoverer analysis were included in the inclusion list. The MS2 parameters were as follows: 70,000 resolution, 5 E5 AGC, 100-ms maximum injection time, 1.0 m/z isolation window, and 30 NCE. MetFrag in silico fragmentation prediction software was used to aid in metabolite identification (38). Target molecules were searched against KEGG and PubChem databases with a 5-ppm error. Candidate molecules from the databases were then processed against the MS/MS spectra of the target molecule with 5-ppm and 0.01-m/zabs settings. The top results of the in silico fragmentation were analyzed for putative identification. Putative identifications were then verified by comparing the experimental MS/MS to the MS/MS of the commercial standard.

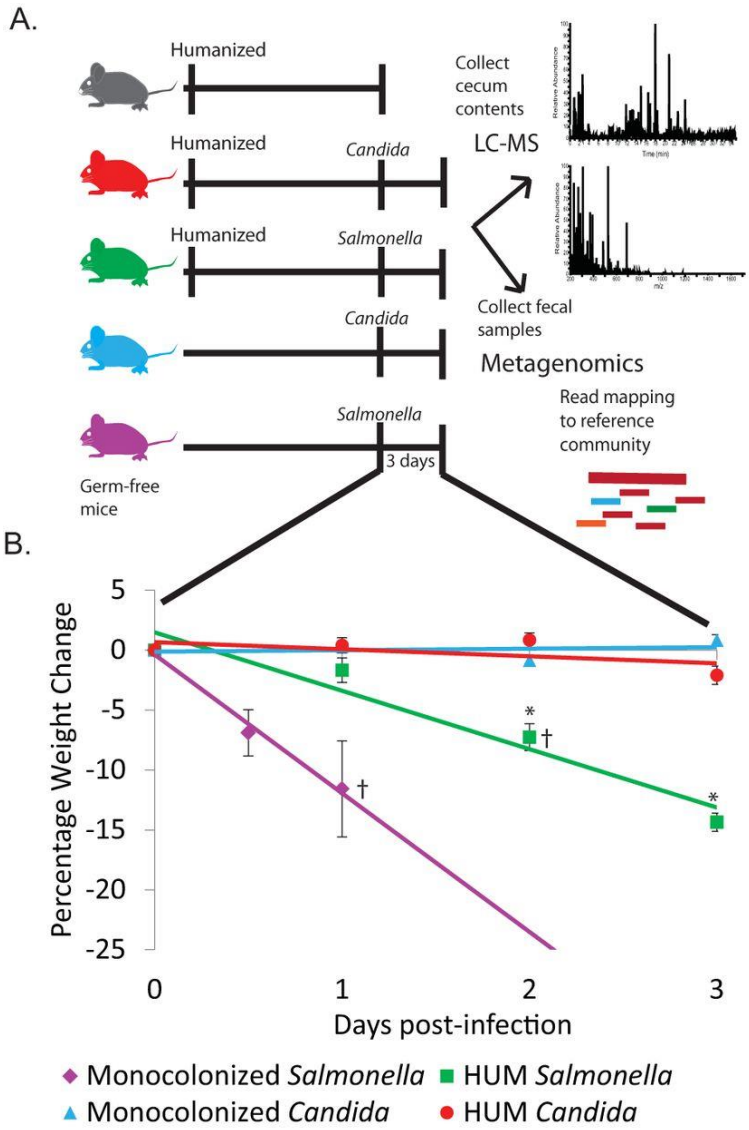
Accession number(s).The metagenome sequences from this study are available under the BioProject identifier PRJNA491522 (https://www.ncbi.nlm.nih.gov/sra/PRJNA491522). The metabolomics data are available from the MetaboLights database under the accession number MTBLS753 (https://www.ebi.ac.uk/metabolights/MTBLS753).

3.6 Acknowledgements

This work was supported by T32AI55397 and U19AI109673 from the National Institutes of Health, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation. Sequencing was provided by the University of Wisconsin Biotech Center. This work was supported in part by grants NIH DK108259 (to F.E.R.). L.L. acknowledges a Vilas Distinguished Achievement Professorship, NIH DK071801, and S10RR029531. The Q-Exactive instrument was purchased through an NIH shared instrument grant (NCRR S10RR029531).

We thank Kimberley Romano for assistance in mouse experimental procedures and Marc Chevrette for advice on secondary metabolite clustering methods. We also thank Camila Carlos, Lily Khadempour, Heidi Horn, and Lindsay Kalan for critical feedback on the manuscript.

3.7 Figures



3.6. 1 Figure 1. (A) Overview of experimental design. (B) Percent body weight loss during 3 days of infection. Errors bars indicate standard error. Significant difference from HUM *Candida* ($P < 0.05$) using Wilcoxon test denoted by * next to relevant group. Mice sacrificed early indicated with † (monocolonized *Salmonella*, 3 at 12 h and 3 at 24 h, HUM *Salmonella* 1 at 24 h).

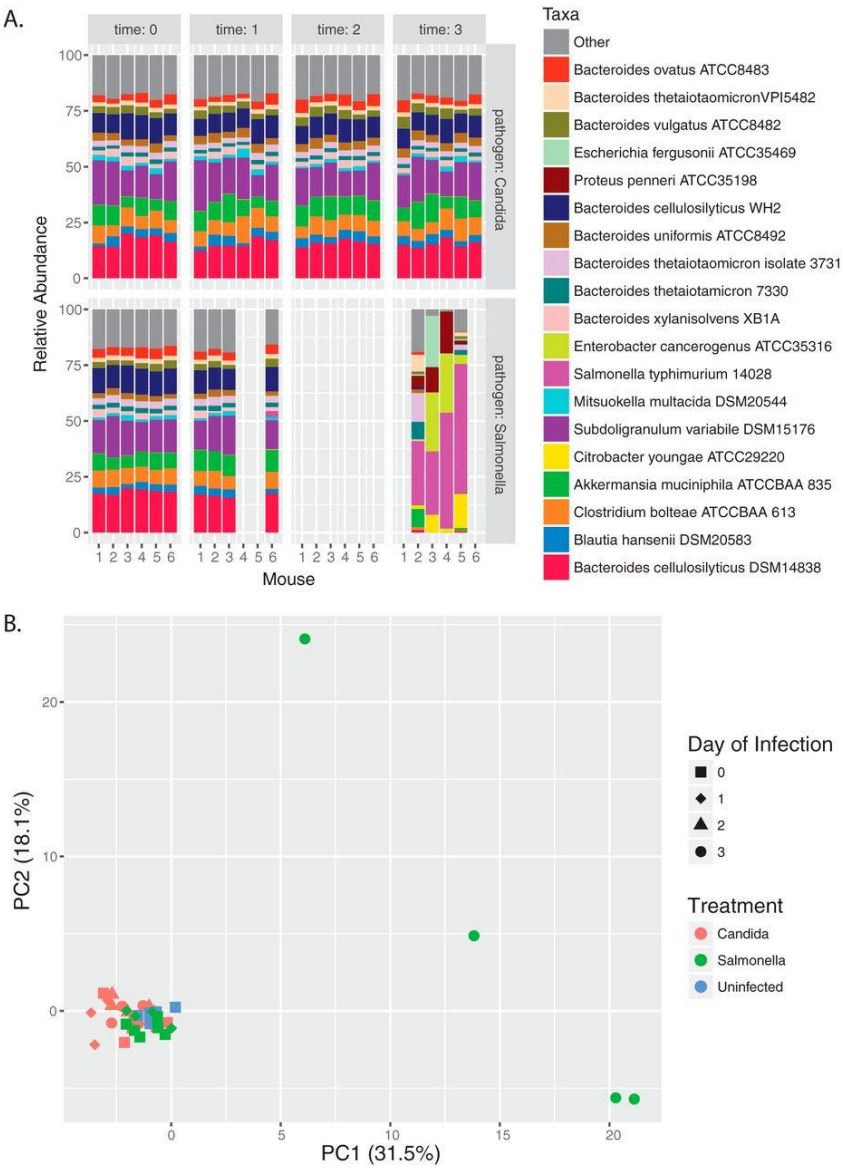


Figure 2. Variation in fecal microbiota metagenomes during infection. (A) Relative abundance of top 19 strains in HUM Candida albicans and Salmonella enterica Typhimurium infection group. (B) PCA of strain relative abundance.

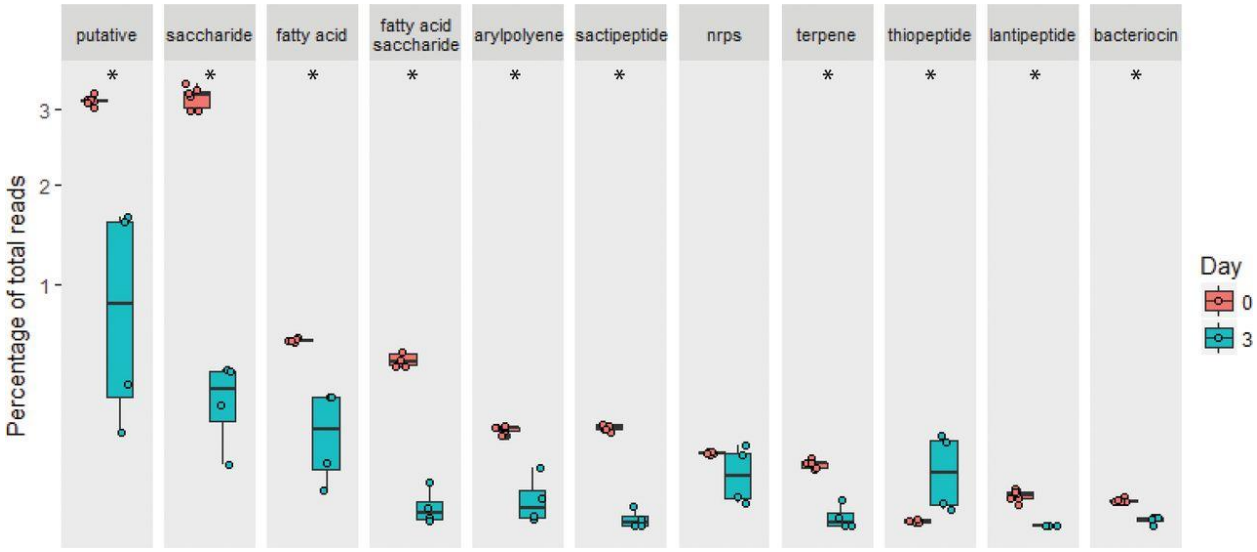


Figure 3. Percent abundance of reads mapping to biosynthetic gene clusters out of total reads that were mapped from the metagenome from HUM *Salmonella*-infected mice prior to infection (n = 6) and 3 days into infection (n = 4), on a square root-adjusted axis. Significance ($P < 0.05$ with Benjamini-Hochberg correction) is indicated with *.

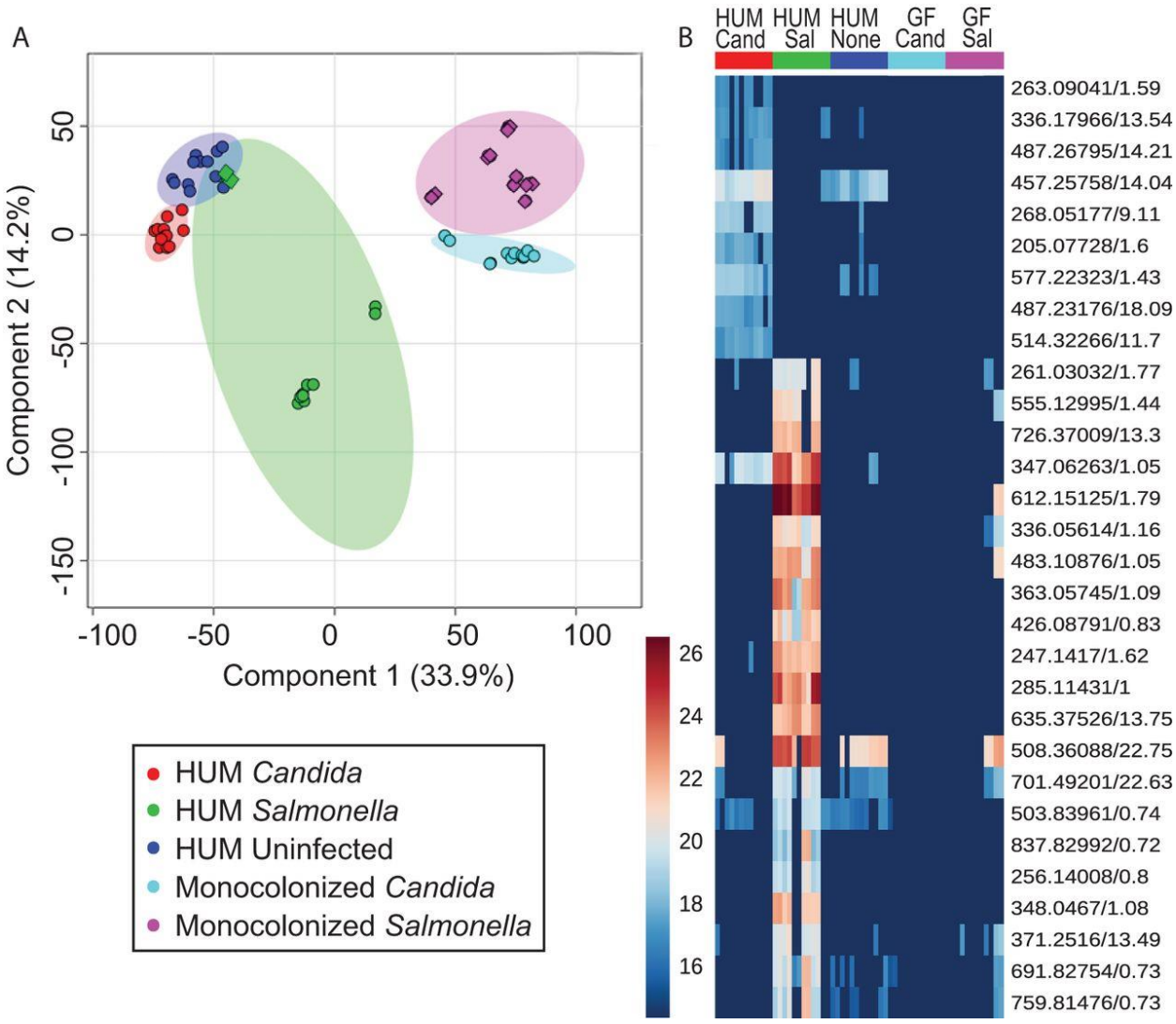


Figure 4. (A) PLS-DA of metabolites from all groups, with 95% confidence intervals. (B) Metabolites of interest 1.5× higher in HUM infected groups than uninfected mice, absent in 8/10 technical replicates for monocolonized mice. Circles are samples collected three days postinfection; diamonds are from animals sacrificed 1 day postinfection.

3.8 References

1. Phane Hacquard S, Garrido-Oter R, Gonzá A, Spaepen S, Ackermann G, Lebeis S, McHardy AC, Dangl JL, Knight R, Ley R, Schulze-Lefert P. 2015. Microbiota and host nutrition across plant and animal kingdoms. Cell Host Microbe 17:603–616. doi:10.1016/j.chom.2015.04.009.

2. Sharon G, Garg N, Debelius J, Knight R, Dorrestein PC, Mazmanian SK. 2014. Specialized metabolites from the microbiome in health and disease. Cell Metab 20:719–730. doi:10.1016/j.cmet.2014.10.016.

3. Lynch SV, Pedersen O. 2016. The human intestinal microbiome in health and disease. N Engl J Med 375:2369–2379. doi:10.1056/NEJMra1600266.

4. Sender R, Fuchs S, Milo R. 2016. Revised estimates for the number of human and bacteria cells in the body. PLoS Biol 14:e1002533. doi:10.1371/journal.pbio.1002533.

5. Walter J, Ley R. 2011. The human gut microbiome: ecology and recent evolutionary changes. Annu Rev Microbiol 65:411–429. doi:10.1146/annurev-micro-090110-102830.

6. Scanlan PD, Marchesi JR. 2008. Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. ISME J 2:1183–1193. doi:10.1038/ismej.2008.76.

7. Hoffmann C, Dollive S, Grunberg S, Chen J, Li H, Wu GD, Lewis JD, Bushman FD. 2013. Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. PLoS One 8:e66019. doi:10.1371/annotation/31412345-fc86-4d67-b37c-93d42f5f0a59.

8. Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, Lewis JD, Bushman FD. 2011. The human gut virome: inter-individual variation and dynamic response to diet. Genome Res 21:1616–1625. doi:10.1101/gr.122705.111.

9. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J-M, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD, Wang J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464:59–65. doi:10.1038/nature08821.

10. Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, Brady A, Creasy HH, McCracken C, Giglio MG, McDonald D, Franzosa EA, Knight R, White O, Huttenhower C. 2017. Strains, functions and dynamics in the expanded Human Microbiome Project. Nature 550:61. doi:10.1038/nature23889.

11. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Heath AC, Leibel RL, Rosenbaum M, Gordon JI. 2013. The long-term stability of the human gut microbiota. Science 341:1237439. doi:10.1126/science.1237439.

12. Buffie CG, Jarchum I, Equinda M, Lipuma L, Gobourne A, Viale A, Ubeda C, Xavier J, Pamer EG. 2012. Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to *Clostridium difficile*-induced colitis. Infect Immun 80:62–73. doi:10.1128/IAI.05496-11.

13. Ubeda C, Taur Y, Jenq RR, Equinda MJ, Son T, Samstein M, Viale A, Socci ND, van den Brink MRM, Kamboj M, Pamer EG. 2010. Vancomycin-resistant *Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. J Clin Invest 120:4332–4341. doi:10.1172/JCI43918.

14. Lawley TD, Bouley DM, Hoy YE, Gerke C, Relman DA, Monack DM. 2008. Host transmission of *Salmonella enterica* serovar Typhimurium is controlled by virulence factors and indigenous intestinal microbiota. Infect Immun 76:403–416. doi:10.1128/IAI.01189-07.

15. Bohnhoff M, Drake BL, Miller CP. 1954. Effect of streptomycin on susceptibility of intestinal tract to experimental *Salmonella* infection. Proc Soc Exp Biol Med 86:132–137. doi:10.3181/00379727-86-21030.

16. Buffie CG, Pamer EG. 2013. Microbiota-mediated colonization resistance against intestinal pathogens. Nat Rev Immunol 13:790–801. doi:10.1038/nri3535.

17. Spees AM, Lopez CA, Kingsbury DD, Winter SE, Bäumler AJ. 2013. Colonization resistance: battle of the bugs or ménage à trois with the host? PLoS Pathog 9:e1003730. doi:10.1371/journal.ppat.1003730.

18. Steed AL, Christophi GP, Kaiko GE, Sun L, Goodwin VM, Jain U, Esaulova E, Artyomov MN, Morales DJ, Holtzman MJ, Boon ACM, Lenschow DJ, Stappenbeck TS. 2017. The microbial metabolite desaminotyrosine protects from influenza through type I interferon. Science 357:498–502. doi:10.1126/science.aam5336.

19. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, No D, Liu H, Kinnebrew M, Viale A, Littmann E, van den Brink MRM, Jenq RR, Taur Y, Sander C, Cross JR, Toussaint NC, Xavier JB, Pamer EG. 2015. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. Nature 517:205–208. doi:10.1038/nature13828.

20. Hsiao A, Ahmed AMS, Subramanian S, Griffin NW, Drewry LL, Petri WA, Haque R, Ahmed T, Gordon JI, Gordon JI. 2014. Members of the human gut microbiota involved in recovery from Vibrio cholerae infection. Nature 515:423–426. doi:10.1038/nature13738.

21. Donia MS, Cimermancic P, Schulze CJ, Wieland Brown LC, Martin J, Mitreva M, Clardy J, Linington RG, Fischbach MA. 2014. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. Cell 158:1402–1414. doi:10.1016/j.cell.2014.08.032.

22. Claus SP, Tsang TM, Wang Y, Cloarec O, Skordi E, Martin F-P, Rezzi S, Ross A, Kochhar S, Holmes E, Nicholson JK. 2008. Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes. Mol Syst Biol 4:219. doi:10.1038/msb.2008.56.

23. Martin F-PJ, Dumas M-E, Wang Y, Legido-Quigley C, Yap IKS, Tang H, Zirah S, Murphy GM, Cloarec O, Lindon JC, Sprenger N, Fay LB, Kochhar S, van Bladeren P, Holmes E, Nicholson JK. 2007. A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. Mol Syst Biol 3:112. doi:10.1038/msb4100153.

24. Zheng X, Xie G, Zhao A, Zhao L, Yao C, Chiu NHL, Zhou Z, Bao Y, Jia W, Nicholson JK, Jia W. 2011. The footprints of gut microbial–mammalian co-metabolism. J Proteome Res 10:5512–5522. doi:10.1021/pr2007945.

25. Lu K, Abo RP, Schlieper KA, Graffam ME, Levine S, Wishnok JS, Swenberg JA, Tannenbaum SR, Fox JG. 2014. Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis. Environ Health Perspect 122:284–291. doi:10.1289/ehp.1307429.

26. Hou Y, Braun DR, Michel CR, Klassen JL, Adnani N, Wyche TP, Bugni TS. 2012. Microbial strain prioritization using metabolomics tools for the discovery of natural products. Anal Chem 84:4277–4283. doi:10.1021/ac202623g.

27. Krug D, Zurek G, Revermann O, Vos M, Velicer GJ, Müller R. 2008. Discovering the hidden secondary metabolome of Myxococcus xanthus: a study of intraspecific diversity. Appl Environ Microbiol 74:3058–3068. doi:10.1128/AEM.02863-07.

28. Barman M, Unold D, Shifley K, Amir E, Hung K, Bos N, Salzman N. 2008. Enteric salmonellosis disrupts the microbial ecology of the murine gastrointestinal tract. Infect Immun 76:907–915. doi:10.1128/IAI.01432-07.

29. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, Chaffron S, Macpherson AJ, Buer J, Parkhill J, Dougan G, von Mering C, Hardt W-D. 2007. Salmonella enterica serovar Typhimurium exploits inflammation to compete with the intestinal microbiota. PLoS Biol 5:e244. doi:10.1371/journal.pbio.0050244.

30. Rivera-Chávez F, Bäumler AJ. 2015. The pyromaniac inside you: *Salmonella* metabolism in the host gut. Annu Rev Microbiol 69:31–48. doi:10.1146/annurev-micro-091014-104108.

31. Kumamoto CA. 2011. Inflammation and gastrointestinal *Candida* colonization. Curr Opin Microbiol 14:386–391. doi:10.1016/j.mib.2011.07.015.

32. Mason KL, Erb Downward JR, Falkowski NR, Young VB, Kao JY, Huffnagle GB. 2012. Interplay between the gastric bacterial microbiota and *Candida albicans* during postantibiotic recolonization and gastritis. Infect Immun 80:150–158. doi:10.1128/IAI.05162-11.

33. Koh AY. 2013. Murine models of *Candida* gastrointestinal colonization and dissemination. Eukaryot Cell 12:1416–1422. doi:10.1128/EC.00196-13.

34. Auchtung TA, Fofanova TY, Stewart CJ, Nash AK, Wong MC, Gesell JR, Auchtung JM, Ajami NJ, Petrosino JF. 2018. Investigating colonization of the healthy adult gastrointestinal tract by fungi. mSphere 3:e00092-18. doi:10.1128/mSphere.00092-18.

35. Schofield DA, Westwater C, Balish E. 2005. Divergent chemokine, cytokine and -defensin responses to gastric candidiasis in immunocompetent C57BL/6 and BALB/c mice. J Med Microbiol 54:87–92. doi:10.1099/jmm.0.45755-0.

36. Kennedy MJ, Volz PA. 1985. Effect of various antibiotics on gastrointestinal colonization and dissemination by *Candida albicans*. Med Mycol 23:265–273. doi:10.1080/00362178585380391.

37. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, de los Santos ELC, Kim HU, Nave M, Dickschat JS, Mitchell DA, Shelest E, Breitling R, Takano E, Lee SY, Weber T, Medema MH. 2017. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res 45:W36–W41. doi:10.1093/nar/gkx319.

38. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. 2016. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. J Cheminform 8:3. doi:10.1186/s13321-016-0115-9.

39. Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA. 2012. The application of ecological theory toward an understanding of the human microbiome. Science 336:1255–1262. doi:10.1126/science.1224203.

40. Antunes LCM, McDonald JAK, Schroeter K, Carlucci C, Ferreira RBR, Wang M, Yurist-Doutsch S, Hira G, Jacobson K, Davies J, Allen-Vercoe E, Finlay BB. 2014. Antivirulence activity of the human gut metabolome. mBio 5:e01183-14. doi:10.1128/mBio.01183-14.

41. Byrd AL, Segre JA. 2016. Criteria for disease causation must take microbial interactions into account. Science 351:224–226. doi:10.1126/science.aad6753.

42. Wesolowska-Andersen A, Bahl MI, Carvalho V, Kristiansen K, Sicheritz-Pontén T, Gupta R, Licht TR. 2014. Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. Microbiome 2:19. doi:10.1186/2049-2618-2-19.

43. Deatherage Kaiser BL, Li J, Sanford JA, Kim Y-M, Kronewitter SR, Jones MB, Peterson CT, Peterson SN, Frank BC, Purvine SO, Brown JN, Metz TO, Smith RD, Heffron F, Adkins JN. 2013. A multi-omic view of host-pathogen-commensal interplay in *Salmonella*-mediated intestinal infection. PLoS One 8:e67155. doi:10.1371/journal.pone.0067155.

44. Stecher B, Chaffron S, Käppeli R, Hapfelmeier S, Freedrich S, Weber TC, Kirundi J, Suar M, McCoy KD, von Mering C, Macpherson AJ, Hardt W-D. 2010. Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria. PLoS Pathog 6:e1000711. doi:10.1371/journal.ppat.1000711.

45. Rivera-Chávez F, Lopez CA, Bäumler AJ. 2017. Oxygen as a driver of gut dysbiosis. Free Radic Biol Med 105:93–101. doi:10.1016/j.freeradbiomed.2016.09.022.

46. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, Sterzenbach T, Tsolis RM, Roth JR, Bäumler AJ. 2011. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proc Natl Acad Sci U S A 108:17480–17485. doi:10.1073/pnas.1107857108.

47. Deriu E, Liu JZ, Pezeshki M, Edwards RA, Ochoa RJ, Contreras H, Libby SJ, Fang FC, Raffatellu M. 2013. Probiotic bacteria reduce *Salmonella Typhimurium* intestinal colonization by competing for iron. Cell Host Microbe 14:26–37. doi:10.1016/j.chom.2013.06.007.

48. Letzel A-C, Pidot SJ, Hertweck C. 2013. A genomic approach to the cryptic secondary metabolome of the anaerobic world. Nat Prod Rep 30:392–428. doi:10.1039/C2NP20103H.

49. Just-Baringo X, Albericio F, Álvarez M. 2014. Thiopeptide antibiotics: retrospective and recent advances. Mar Drugs 12:317–351. doi:10.3390/md12010317.

50. Marcobal A, Kashyap PC, Nelson TA, Aronov PA, Donia MS, Spormann A, Fischbach MA, Sonnenburg JL. 2013. A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. ISME J 7:1933–1943. doi:10.1038/ismej.2013.89.

51. Antunes LCM, Arena ET, Menendez A, Han J, Ferreira RBR, Buckner MMC, Lolic P, Madilao LL, Bohlmann J, Borchers CH, Finlay BB. 2011. Impact of salmonella infection on host hormone metabolism revealed by metabolomics. Infect Immun 79:1759–1769. doi:10.1128/IAI.01373-10.

52. Zhu X, Lei H, Wu J, Li JV, Tang H, Wang Y. 2014. Systemic responses of BALB/c mice to *Salmonella typhimurium* Infection. J Proteome Res 13:4436–4445. doi:10.1021/pr500770x.

53. Arrieta M-C, Walter J, Finlay BB. 2016. Human microbiota-associated mice: a model with challenges. Cell Host Microbe 19:575–578. doi:10.1016/j.chom.2016.04.014.

54. Stecher B, Macpherson AJ, Hapfelmeier S, Kremer M, Stallmach T, Hardt W-D. 2005. Comparison of *Salmonella enterica* serovar Typhimurium colitis in germfree mice and mice pretreated with streptomycin. Infect Immun 73:3228–3241. doi:10.1128/IAI.73.6.3228-3241.2005.

55. Li JW-H, Vederas JC. 2009. Drug discovery and natural products: end of an era or an endless frontier? Science 325:161–165. doi:10.1126/science.1168243.

56. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T, Porto C, Bouslimani A, Melnik AV, Meehan MJ, Liu W-T, Crüsemann M, Boudreau PD, Esquenazi E, Sandoval-Calderón M, Kersten RD, Pace LA, Quinn RA, Duncan KR, Hsu C-C, Floros DJ, Gavilan RG, Kleigrewe K, Northen T, Dutton RJ, Parrot D, Carlson EE, Aigle B, Michelsen CF, Jelsbak L, Sohlenkamp C, Pevzner P, Edlund A, McLean J, Piel J, Murphy BT, Gerwick L, Liaw C-C, Yang Y-L, Humpf H-U, Maansson M, Keyzers RA, Sims AC, Johnson AR, Sidebottom AM, Sedio BE, Klitgaard A, Larson CB, Boya P CA, Torres-Mendoza D, Gonzalez DJ, Silva DB, Marques LM, Demarque DP, Pociute E, O'Neill EC, Briand E, Helfrich EJN, Granatosky EA, Glukhov E, Ryffel F, Houson H, Mohimani H, Kharbush JJ, Zeng Y, Vorholt JA, Kurita KL, Charusanti P, McPhail KL, Nielsen KF, Vuong L, Elfeki M, Traxler MF, Engene N, Koyama N, Vining OB, Baric R, Silva RR, Mascuch SJ, Tomasi S, Jenkins S, Macherla V, Hoffman T, Agarwal V, Williams PG, Dai J, Neupane R, Gurr

J, Rodríguez AMC, Lamsa A, Zhang C, Dorrestein K, Duggan BM, Almaliti J, Allard P-M, Phapale P, Nothias L-F, Alexandrov T, Litaudon M, Wolfender J-L, Kyle JE, Metz TO, Peryea T, Nguyen D-T, VanLeer D, Shinn P, Jadhav A, Müller R, Waters KM, Shi W, Liu X, Zhang L, Knight R, Jensen PR, Palsson BO, Pogliano K, Linington RG, Gutiérrez M, Lopes NP, Gerwick WH, Moore BS, Dorrestein PC, Bandeira N. 2016. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol 34:828–837. doi:10.1038/nbt.3597.

57. Farr SB, Kogoma T. 1991. Oxidative stress responses in Escherichia coli and Salmonella typhimurium. Microbiol Rev 55:561–585.

58. Hayes JD, McLellan LI. 1999. Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. Free Radic Res 31:273–300. doi:10.1080/10715769900300851.

59. Zhang Y, Brady A, Jones C, Song Y, Darton TC, Jones C, Blohmke CJ, Pollard AJ, Magder LS, Fasano A, Sztein MB, Fraser CM. 2018. Compositional and functional differences in the human gut microbiome correlate with clinical outcome following infection with wild-type Salmonella enterica serovar Typhi. mBio 9:e00686-18.

60. Oz HS, Chen TS, Nagasawa H. 2007. Comparative efficacies of 2 cysteine prodrugs and a glutathione delivery agent in a colitis model. Transl Res 150:122–129. doi:10.1016/j.trsl.2006.12.010.

61. Mardinoglu A, Shoaie S, Bergentall M, Ghaffari P, Zhang C, Larsson E, Bäckhed F, Nielsen J. 2015. The gut microbiota modulates host amino acid and glutathione metabolism in mice. Mol Syst Biol 11:834. doi:10.15252/msb.20156487.

62. Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. 2017. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. Nat Commun 8:1784. doi:10.1038/s41467-017-01973-8.

63. Kassam Z, Lee CH, Yuan Y, Hunt RH. 2013. Fecal microbiota transplantation for Clostridium difficile infection: systematic review and meta-analysis. Am J Gastroenterol 108:500–508. doi:10.1038/ajg.2013.59.

64. Suez J, Elinav E. 2017. The path towards microbiome-based metabolite treatment. Nat Microbiol 2:17075. doi:10.1038/nmicrobiol.2017.75.

65. Romano KA, Vivas EI, Amador-Noguez D, Rey FE. 2015. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. mBio 6:e02481-14. doi:10.1128/mBio.02481-14.

66. Faith JJ, Ahern PP, Ridaura VK, Cheng J, Gordon JI. 2014. Identifying gut microbe-host phenotype relationships using combinatorial communities in gnotobiotic mice. Sci Transl Med 6:220ra11. doi:10.1126/scitranslmed.3008051.

67. Dunne C. 2001. Adaptation of bacteria to the intestinal niche: probiotics and gut disorder. Inflamm Bowel Dis 7:136–145. doi:10.1097/00054725-200105000-00010.

68. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. 2009. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 1:6ra14. doi:10.1126/scitranslmed.3000322.

69. McNulty NP, Yatsunenko T, Hsiao A, Faith JJ, Muegge BD, Goodman AL, Henrissat B, Oozeer R, Cools-Portier S, Gobert G, Chervaux C, Knights D, Lozupone CA, Knight R, Duncan AE, Bain JR, Muehlbauer MJ, Newgard CB, Heath AC, Gordon JI. 2011. The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. Sci Transl Med 3:106ra106. doi:10.1126/scitranslmed.3002701.

70. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. doi:10.1186/gb-2009-10-3-r25.

71. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. doi:10.1093/bioinformatics/btp324.

72. Xia J, Sinelnikov IV, Han B, Wishart DS. 2015. MetaboAnalyst 3.0—making metabolomics more meaningful. Nucleic Acids Res 43:W251–W257. doi:10.1093/nar/gkv380.

73. Xia J, Wishart DS, Xia J, Wishart DS. 2016. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. Curr Protoc Bioinformatics 55:14.10.1–14.10.91. doi:10.1002/cpbi.11.

Chapter 4: Variation in Human Gut Microbiome and Resistance to *Salmonella enterica* Typhimurium Infection

Jennifer R. Bratburd, Caitlin Keller, Jericha Mill, Eugenio Vivas, Darin Wiesner, Bruce Klein, Lexis Wedell, Lingjun Li, Federico E. Rey, Cameron R. Currie

4.1 Abstract

Colonization resistance is an important feature of a healthy microbiome that protects the host from infection. Gut microbiomes have been shown to vary in the amount of resistance they provide against invading pathogens such as *Salmonella*. Here, we used gnotobiotic mice colonized with human microbiomes to assess the degrees of variation across communities in their capacity to confer more resistance to or protection from *Salmonella*. Of the ten communities we tested, we found variation in susceptibility. We sequenced metagenomes of gut microbial communities prior to infection to explore commonalities among resistant versus susceptible communities. We found one strain (Clostridiales sp. 1 7 47FAA) shared in resistant metagenomes. We also found several functional categories enriched in susceptible communities including pathways related to purine and pyrimidine biosynthesis, the methylerythritol phosphate pathway, rhamnose degradation, glycolysis, chorismate biosynthesis, glycogen degradation, and CMP-3-deoxy-D-manno-octulosonate. Follow-up metabolomics of one highly susceptible and one highly resistant community also suggested communities had many different metabolites prior to infection. Further exploration is needed to determine underlying shared mechanisms.

4.2 Introduction

The human gut microbiome is linked to many aspects of health, including metabolic and neurological disorders as well as protection from infectious disease (Kim, Covington, and Pamer 2017; Boulangé et al. 2016; Tremlett et al. 2017). Despite the increasing recognition of the

microbiome's role in health, there is no agreement on what constitutes a healthy microbiome or an unhealthy, dysbiotic state (Hooks and O'Malley 2017; Lloyd-Price, Abu-Ali, and Huttenhower 2016). Efforts to identify a core healthy microbiome in humans have instead found a wide diversity of strains and some broadly defined core metabolic functions ("Structure, Function and Diversity of the Healthy Human Microbiome" 2012). Some definitions of a healthy microbiome consider overall ecological features such as diversity, where higher diversity often is associated with gut microbiomes of healthy subjects. Other definitions of a healthy community center on resistance and resilience to perturbations, rather than on particular strains and functions (Lloyd-Price, Abu-Ali, and Huttenhower 2016). Of all the proposed definitions, resistance can most directly contribute to host health by preventing infectious pathogens from establishing; however, understanding the underlying contributors to resistance remains a challenge.

The idea of microbiome resistance has been widely explored in the context of how the existing community prevents the establishment of infectious pathogens, particularly species in the genus *Salmonella*, gut pathogens responsible for over 1 million illnesses per year in the United States (Bohnhoff, Drake, and Miller 1954; Rivera-Chávez and Bäumler 2015, Scallan et al. n.d.). Several taxa and their functions within gut communities have been associated with resistance or susceptibility to *Salmonella*. For example, some *Clostridia* produce butyrate that feeds epithelial cells and keeps the intestines anaerobic. When antibiotics disrupt the butyrate-producers, epithelial cells, without a supply of butyrate switch to lactate fermentation with less oxygen use, which *Salmonella* can then use (Gillis et al. 2018). Likewise, *Salmonella* can trigger inflammation in order to disturb the microbiome and reduce competition for other carbon sources or nutrients that are diet derived like fructose-asparagine (Wu et al. 2018), microbiota-derived like succinate (Spiga et al. 2017), or host and diet derived like 1,2 propanediol (Staib and Fuchs

2015; Faber et al. 2017). *Salmonella* can also use the byproducts of inflammation like tetrathionate or oxygen as electron acceptors (Winter et al. 2010; Rivera-Chávez et al. 2016). Closely-related Enterobacteriaceae species have been proposed to compete with *Salmonella* in disturbed environments, as well as produce metabolites that may inhibit *Salmonella* (Velazquez et al. 2019; Rivera-Chávez and Bäumler 2015). In addition to competition, microbes can also modulate the immune system and impact the outcome of *Salmonella* infection (Thiemann et al. 2017). Some strains could potentially make the host more susceptible, as proposed with a mouse model monocolonized with the mucus-utilising bacteria *Akkermansia muciniphila*, which is suggested thin the mucus layer and allow *Salmonella* to better infect the host (Ganesh et al. 2013).

Variation in the microbiome can impact the effectiveness of colonization resistance (Thiemann et al. 2017; Velazquez et al. 2019), and strains thought to increase or decrease susceptibility may exist simultaneously. In this study, we explore the variability of colonization resistance in the human gut microbiome using a gnotobiotic mouse model. We leveraged use of samples from the Wisconsin Longitudinal Study (Herd, Carr, and Roan 2014), a cohort with more than 60 years of metadata in order to test how different human microbiomes affect resistance or susceptibility to *Salmonella enterica* Typhimurium. We used metagenomic analysis to identify common features among resistant communities, and performed more in-depth metagenomic and immunological analysis on two donor communities. We found variability in resistance based on donor metagenomes, and a few commonalities between these microbiomes, including an enriched strain and several putative functions in higher abundance in resistant metagenomes.

4.3 Methods

4.3.1 Microbiota and Pathogen Growth

Ten human stool samples were obtained from the Wisconsin Longitudinal Study, a cohort of Wisconsin high school graduates from the class of 1957, along with siblings and spouses. The Institutional Review Board (IRB) at the University of Wisconsin-Madison approved WLS data collection with informed consent (2014-1066, 2015-0955). Samples were stored at -80°C until use. We grew *Salmonella enterica* Typhimurium ATCC 14028 aerobically overnight in lysogeny broth (LB) at 37°C.

4.3.2 Mouse Microbiome Colonization and Experimental Infections

Protocols used for mouse experiments were approved by the University of Wisconsin-Madison Animal Care and Use Committee. Germ-free female C57BL/6J mice were maintained in gnotobiotic isolators until 8 to 12 weeks of age with 12-h light cycle and sterilized food and water *ad libitum*. Mice were then moved to out-of-isolator sterile cages, with 2 mice per cage. Once out of the isolator, all procedures including inoculation of mice with human feces and infection with mice used aseptic technique in a biological safety cabinet.

Mice were orally gavaged with a 0.2 mL fecal slurry from frozen fecal samples from a human donor following previously established protocol (Romano et al. 2018). Mice were given 2 weeks for microbiomes to stabilize. After this stabilization, mice were infected with *Salmonella enterica* Typhimurium ATCC 14028 via oral gavage with 0.2 mL overnight culture. In the first set of experiments, 3 cages of mice with 2 mice per cage were used for each donor microbiome. Mice were sacrificed when weight loss exceeded 10% of their original weight. In the second set of experiments, mice were sacrificed at intervals: prior to infection, 2 days post-infection and 5 days post-infection. Cecal contents, sections of small and large intestines were collected for all

mice at time of death and stored at -80°C . Fecal samples were collected daily when possible and used to determine colony-forming units of *Salmonella* by growing on selective xylose lysine deoxycholate (XLD) agar to confirm *Salmonella* colonization. Survival analysis performed in R using packages survminer and survival, using a Kaplan-Meier estimator. Pairwise differences used Benjamini-Hochberg adjusted p-values.

4.3.3 Metagenomic Sequencing and Analysis

To characterize microbiomes, fecal samples taken from mice immediately prior to infection and human samples were sequenced using Illumina HiSeq. Genomic DNA was extracted from fecal pellets following the protocol described in Bratburd et al 2018. DNA library preparation and sequencing were done at the University of Wisconsin-Madison Biotechnology Center. Samples were prepared with the TruSeq Nano DNA LT Library Prep kit (Illumina Inc., San Diego, CA, USA) with minor modifications. After shearing samples with a Covaris M220 Ultrasonicator (Covaris Inc., Woburn, MA, USA), samples were size selected for an average insert size of 550 bp using SPRI bead-based size exclusion, and then libraries were standardized to 2 nM.

Metagenomic data was processed using fastp to trim adapters, sequencing reagent contamination and low quality reads (Chen et al. 2018). Humann2 was used to analyze and annotate reads for taxonomic identification and functional characterization (Franzosa et al. 2018). Diversity was analyzed using the vegan package in R with a Kruskal-Wallis test. To identify taxa and functional categories enriched in susceptible or resistance metagenome, I split microbiomes into resistant and susceptible categories based on average days survived, with susceptible microbiomes average survival less than 7 days (n=3), and resistant greater than 8 days (n=5). Microbiomes with average survival in between these cutoffs (n=2) were not used in

this analysis. Linear discrimination analysis effect size was perfomed using Lefse (Segata et al. 2011). Parameters were set to all-against-all (more strict) multi-class analysis, with default thresholds. Subsequent data visualization and analysis was performed in R, version 1.1.456. Code is available at https://github.com/bratburd/wgs_metagenome.

4.3.4 Metabolomics

Extraction of metabolites from mouse cecum samples was performed using methods as previously described in Bratburd et al (2018), with some modifications described below. Aqueous samples were resuspended in optima grade water at a concentration of 10 mg/mL. A Dionex Ultimate 3000 UHPLC system (Thermo Scientific, Waltham, MA, USA) and a Cortecs C18 column (2.1 mm internal diameter x 100 mm length, 1.6 μm particle size; Waters, Milford, MA, USA), equipped with a corresponding guard column were used to separate the samples. The column temperature was 35°C, and the mobile phases were optima grade water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). The separation occurred with a 35 minute gradient at a flow rate of 0.3 mL/minutes with the following conditions: 0–5 min, 1% B; 5–10 min, linear gradient from 1–3% B; 10–18 min, linear gradient from 3–40% B; 18–22 min, linear gradient from 40–80% B; 22–27 min, column cleaning at 95% B; and 27–35 min, re-equilibration at 1% B. The injection volume was 3 μL and the samples were kept at 10°C during analysis. Metabolite MS data was acquired on a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific, Waltham, MA, USA), which was equipped with an ESI source and operated in positive ion mode with a scan range of *m/z* 100–1500. The MS parameters were as follows: 70,000 resolution, 1 E6 AGC, and 100 ms maximum injection time. MS/MS was collected with a top 3 DDA with the following parameters: 35,000 resolution, 1 E5 AGC, 100 ms maximum injection time, 1.0 *m/z* isolation window, and 30 NCE.

4.3.5 Metabolomics Data Analysis

Relative quantification of the metabolomics data for the different sample types was performed with Compound Discoverer software (Thermo Scientific, Waltham, MA, USA). Spectra underwent retention time alignment, detection of unknown compounds, and grouping of unknown compounds. The Compound Discoverer workflow also included fill gaps, mark background, predict compositions, ChemSpider search, normalize areas, merge features, and differential analysis.

4.3.6 Compound Identification

MS/MS spectra for the compounds on the target lists for both infections were collected on the Dionex UltiMate 3000 UHPLC and Q-Exactive instrument described above. The injection volume was 3 μL. An inclusion list was used for the targets with a retention time window of +/- 0.7 min. All charge states and salt adducts observed in the Compound Discoverer analysis were included in the inclusion list. The MS^2 parameters were as follows: 70,000 resolution, 5 E5 AGC, 100 ms maximum injection time, 1.0 *m/z* isolation window, and 30 NCE.

4.3.7 Immunology

Mesenteric lymph nodes and spleens were collected at time of sacrifice. Mesenteric lymph nodes (MesLN) and spleens were treated similarly, unless otherwise noted. Organs were harvested, homogenized with frosted glass slides, suspended in PBS+1%BSA, and passed through a 70um filter. Splenocytes were treated with red blood cell lysis for 15 minutes at RT and washed with PBS + 1% BSA. ¼ of MesLN cells and 1/20 spleen cells were treated with fluorescent coupled antibodies (see below) for 30 minutes at 4C. Cells were washed, fixed with 10% PFA, suspended in 200uL PBS + 1% BSA, and analyzed by flow cytometry (BD Fortessa). Data were analyzed with Flowjo 10 (Treestar). Gating strategy (Supplemental Figure 1). The

following reagents were used: RBC Lysis Biolegend 420301, Live/Dead Near-IR, Thermo Fisher L34976. The following antigens were used (Antigen, Fluorophore, Clone, Vendor, Cat #**):** CD90.2, BV785, 30-H12, Biolegend, 105331; CD19, PE-Dazzle, 6D5, Biolegend, 115553; Ly6C, BV510, HK1.4, Biolegend, 128033; Ly6G, BUV395, 1A8, BD Biosciences, 563978; CD64, PE, X54-5/7.1, Biolegend, 139303; CD11b, BV650, M1/70, Biolegend, 101239; CD11c, PE-Cy7, N418, Biolegend, 117317; MHCII, AF700, M5/114.15.2, Biolegend, 107621

4.4 Results

4.4.1 Human Microbiome Engraftment in Mice

We colonized germ-free mice with 10 different randomly selected human donor samples in order to test which communities enhanced resistance to *Salmonella* infection (Figure 1A). Donors ranged in age from 68-77 and 50% were female. Only one donor (WLS5) had antibiotics within the last 6 months prior to sample collection. Based on a Principal Coordinates Analysis (PCoA) of Bray-Curtis dissimilarities, communities separated first based on whether the source material came from the human donor or mouse colonized with that donor's communities. In mice, communities from the same donor also grouped together (Permanova on Bray-Curtis Dissimilarity, p=0.049)(Figure 1B). Shannon and inverse Simpson indices of diversity significantly differed based on host animal, with human donor samples (n=10) on average having a Shannon index of 2.93 and an inverse Simpson index of 12.75. The values from human samples were significantly higher than mice colonized with human donors (n=59), which had an average Shannon index of 2.50, inverse Simpson of 7.96 (Inverse simpson index: Kruskal-Wallis, p-value=0.0002826; Shannon index, Kruskal-Wallis p=8.836e-05)(Figure 1C). Donor communities in mice also showed variability in community composition and diversity (Figure 1D).

4.4.2 Microbiome Resistance to *Salmonella*

On average, all mice with any donor communities succumbed to infection at 7.8 days post infection, standard deviation (s.d.) 2.4. In contrast, germ-free infected mice succumbed within 1-2 days post infection (n=3). When analyzed based on cage, we found an overall effect of donor on survival (p=0.00021, log-rank test) (Figure 2A). Pairwise testing with each individual mouse as the unit of replication (n=6 mice/microbiome donor) revealed significant differences of survival between different donor microbiomes. Using cages as units of replication (n=3 cages/microbiome donor), Benjamini-Hochberg corrected pairwise testing did not reveal significant differences based on donor (Figure 3A). Note that while cages are a better unit for the microbiome due to shared taxa in the same environment, measuring death for an individual is more straightforward than for cages.

4.4.3 Metagenomic Analysis of Microbiomes Resistant to *Salmonella*

We sequenced metagenomes from fecal samples from at least three mice per donor collected immediately prior to infection. These communities did not show grouping based on time to death after plotting the Bray-Curtis dissimilarities on NMDS (Figure 2C). Diversity, as measured by the Shannon and inverse Simpson index had a positive correlation but insignificant with survival measured in days to death (inverse Simpson index: linear model, p=0.167, R2=0.02317, Shannon index: p=0.215, R2=0.0145, Figure 2B). When splitting by donor, the strongest positive correlation was for community WLS39, which also had the largest number of samples. Based on linear models, this correlation was significant for inverse Simpson index (p=0.0385, R2=0.6217) and insignificant for Shannon diversity (p=0.05907, R2=0.5388). When we averaged the Shannon or inverse Simpson's index diversities of the communities in mice, we did not see a clear or significant relationship between the donor community diversity and average

days to death in mice colonized with those communities (inverse Simpson Mice p=0.52 R2=0.06; Shannon p=0.598, R2=0.036, Figure 2D). Likewise, when we compared diversity of the original human donors versus the average days to death in mice, we did not find a significant relationship (inverse Simpson Human donors p=0.77, R2=0.01, Shannon p=0.99, R2=8.9e-6, Figure 2D).

We searched for taxa associated with resistance (on average dying after 8 days) or susceptibility (dying prior to 7 days of infection) using Lefse to identify features differing between resistant and susceptible microbiomes (Segata et al. 2011) (Figure 3A). We identified Clostridiales sp. 1 7 47FAA GCF000155435 (log LDA score of 3.41), but this appeared mainly enriched in 2 metagenomes from one donor (Figure 3B). We used the same method with metagenomic data mapped to metacyc functional annotations. The general pathways associated with susceptible metagenomes include: methylerythritol phosphate pathway, inosine monophosphate biosynthesis III, guanosine ribonucleotide de novo biosynthesis, and rhamnose degradation I, glycogen degradation II, CMP-3-deoxy-D-manno-octulosonate biosynthesis I, superpathway of histidine, purine and pyrimidine biosynthesis, glycolysis I, and chorismate biosynthesis I (log LDA scores > 2). No general pathways (not specific to a particular strain) were associated with resistance.

4.4.4 Characterizing a Susceptible and a Resistant Microbiome

We selected a consistently susceptible community that survived on average 5-6 days (WLS39), and one community that survived on average 8-10 days (WLS28) for follow-up experiments. Initially, we saw distinct separation in weight loss within the first 5 days post-infection; however, in the follow-up experiments we did not see a clear distinction (Figure 4A). For subsequent immunology and metabolomics, we used samples from mice prior to infection.

We performed LC-MS on aqueous extracts collected from mice cecum from this follow-up experiment. We found that mice colonized with different communities grouped distinctly (Figure 4B). Of the putative metabolites we were able to match to database with MS/MS, we found that community WLS28 was enriched for acetyl-B-methylcholine, and 5'-S-Methyl-5'-thioadenosine while community WLS39 was enriched for glycitein, trans-3-indole, and equol.

Using flow cytometry, we measured immune cells (B cells, lymphocytes, myeloid cells, macrophages, neutrophils, monocytes and dendritic cells) in the mesenteric lymph nodes and spleen. Prior to infection, we found no significant differences in these cells from mice with either community (Supplemental Figure 2).

4.5 Discussion

Here we used human donor microbiomes to explore variation in the human gut microbiota's ability to confer resistance to *Salmonella*. As with previous studies (Thiemann et al. 2017; Velazquez et al. 2019), we find that the microbiota seems to impact resistance or susceptibility to *Salmonella*, with some donor configurations more effective than others. In contrast to those studies which relied on mice microbiota from different vendors or facilities, we used a collection from human donors. These donor communities colonized somewhat consistently within mice, although lost members and overall diversity. This is consistent with previous studies assessing microbiota changes from host to host (Rawls et al. 2006; Romano et al. 2018), and lost members may be human specific taxa (Hugenholtz and de Vos 2018; Li et al. 2019). As other studies have identified, the microbiota confers some protection, and resistance seems to depend on a complex community covering many metabolic functions (Brugiroux et al. 2016; Stecher et al. 2010).

In contrast to previous papers linking gut microbiome diversity with resistance to *Salmonella* (Thiemann et al. 2017; Stecher et al. 2010), we did not find a significant link between diversity and resistance as measured in days to death, with donors and their average days to death in mice or the communities in mice and their average day to death. We are limited by the number of donors used (n=10). Further, the donors themselves come from a similar Western cohort and may not represent a large enough range of diversity overall.

We identified one taxa that was associated with resistant communities, Clostridiales sp. 1 7 47FAA GCF000155435. Surveying more human microbiomes and assessing their resistance to *Salmonella* infection could shed light on the significance of the presence of this strain. The fact that we only found one strain consistently in resistant communities may indicate that specific strains are not as important as the functions those strains provide. As efforts to match specific strains to determinants of health have been met with limited success, the lack of common strains associated with protection may indicate other methods are needed to understand microbiome resistance and resilience.

Functionally we identified several categories associated with susceptible metagenomes, and none associated with resistant metagenomes. This analysis is limited by the lack of annotations for many genes in the microbiome. While the pathways may be associated with resistance, further studies and more sampling could help identify which pathways may be the most consistent or relevant. Two categories enriched in the susceptible metagenomes could be related: superpathway of histidine, purine, and pyrimidine biosynthesis, inosine 5 phosphate biosynthesis III (the first step in de novo purine biosynthesis) and guanosine ribonucleotide de novo biosynthesis. Purine biosynthesis is used in a variety of processes including in DNA and RNA, generating energy, and signaling. IMP biosynthesis III pathway differs from the

biosynthesis I pathway in that it uses two enzymes for the final step as opposed to one multifunctional enzyme, and is traditionally associated with Archaea. In previous discovery metabolomics research we identified IMP enriched in mice with a microbiome during *Salmonella* infection (Bratburd et al. 2018). There are many possibilities for the role purines could play during infection, and purines are central metabolites with wide-ranging impacts on the immune system (Hasko et al. 2004).

Other pathways enriched in metagenomes of mice more susceptible to *Salmonella* include rhamnose degradation pathway I, methylerythritol phosphate (MEP) pathway I, CMP-3-deoxy-D-manno-octulosonate biosynthesis, glycolysis I, chorismate biosynthesis pathway I, superpathway of aromatic amino acids, glycogen degradation II. Rhamnose is a hexose sugar that can be found in bacterial cell walls as part of polysaccharides in plants, as well as in mucin oligosaccharides, and is known to be used by *Salmonella* and other Enterobacteriaceae as a carbon source (Staib and Fuchs 2015; Akhy, Brown, and Old 1984). One hypothesis here is that abundance of genes may reflect the metabolite's availability in the gut. The other pathways' relationship to *Salmonella* colonization is more difficult to speculate on. The MEP pathway is known to form precursors to isoprenoids, and can be found in pathogens like *Salmonella* as well as in many gut microbiome members. MEP can affect the immune system, having antioxidant intermediates that can accumulate during oxidative stress, and with intermediates that elicit production of Vγ9/Vδ2 T cells (Heuston et al. 2012). CMP-3-deoxy-D-manno-octulosonate is part of bacterial lipopolysaccharide (Strohmaier et al. 1995). Glycolysis (Embden-Meyerhof-Parnas pathway) can be used both to generate energy and precursors (Wolfe 2015). Glycogen degradation can be used to liberate glucose from this storage molecule during cell stress (Wilson et al. 2010). The superpathway of aromatic amino acids, responsible for synthesis of tryptophan,

phenylalanine, and tyrosine, also includes the chorismate pathway which was also enriched. The chorismate biosynthesis pathway (also referred to as the shikimate pathway) leads to synthesis of some vitamins, ubiquinone, and siderophores (Dosselaere and Vanderleyden 2001).

From the two communities we pulled for having more and less resistance, we found a lack of distinction in the immune systems prior to infection. The inconsistency in the second experiment may have resulted from a lowered infectious dose used inadvertently in these experiments.This is indicative that two-week colonization with the human donor community in the adult mice did not substantially change the priming of the immune system. Colonizing germ-free mice with human communities does not restore their immune systems entirely (Chung et al. 2012), so while it is possible that the gut microbiota may impact pathogen resistance via the immune system, we are unlikely to distinguish those effects here.

We saw distinctions between the cecum metabolites of the two communities. Previous studies indicate that human donor microbiota in mice can confer unique metabolomic fingerprints (Marcobal et al. 2013). We previously assessed what metabolites changed during infection with or without a microbiome (Bratburd et al. 2018); however, here we were interested in metabolites that could be associated with preventing an infection. While these communities show distinct metabolites, it is unclear if these metabolites play any role in infection. Differences in metabolites may reflect differences in microbiomes that do not impact resistance to *Salmonella*, other stochastic differences in amount of diet consumed or sensitivity of metabolites to be identified by our metabolomic techniques. For example, two of the metabolites from the susceptible communities, glycitein and equol may derive from soy products or gut microbes (Mayo, Vázquez, and Flórez 2019), while methylthioadenosine identified in the resistant communities is a metabolite found in all mammalian tissue. Indole, found in higher abundance in

the susceptible mice, is produced by many bacteria and may be used for signaling (Lee and Lee 2010). In mice studies with *Citrobacter rodentium*, decreased indole levels were associated with higher virulence (Kumar and Sperandio 2019), and suggest that the pathogens sense the concentration of indole levels throughout the GI tract. In studies with human cell lines, indole inhibits *Salmonella enterica* Typhimurium (Kohli et al. 2018). It is difficult to reconcile why we would find this metabolite more abundant in susceptible mice.

There are many possible options for future research. Developing a more standardized and reproducible approach to measuring effectiveness of different microbiomes would be critical for continuation of experiments described above. Coordinating efforts between different labs at different institutions would be an insightful and higher reaching effort for studying how microbiome variability impacts *Salmonella* colonization. Given that human microbiomes are only partially recapitulated in gnotobiotic mice, continued efforts may be limited to finding species or mechanisms that are also relevant in mice hosts. Synthetic communities in gnotobiotic hosts may better interrogate specific mechanisms that have been suggested by this and other research. For example, follow-up experiments could compare the ratio of *Salmonella* colonization with wild-type and knock-outs for various functions, such as rhamnose degradation, to support these explorations experimentally. As mice may bias human microbiome representation, further high throughput *in vitro* studies could be another alternative.

While dysbiosis is often a vaguely defined concept, understanding how specific mechanisms of health, like prevention of colonization by infections agents may help fill in the details. These approaches could potentially help identify which features, if any, predict risk for *Salmonella* infection. Furthermore, this may lead to finding microbiome attributes that affect general resilience to other perturbations, and ultimately what factors are associated with health.

4.6 Author Contributions

JRB designed and executed the experiment, performed analysis and wrote the manuscript. Caitlin Keller and Jericha Mill performed LC-MS and analyzed the LC-MS data and wrote methods pertaining to these experiments. Eugenio Vivas maintained the germ-free mouse colony and animal husbandry. Darin Wiesner performed the cell sorting for immune cell detection and assisted in mice sacrifices. Lexis Wedell performed cell plating. Lingjun Li, Bruce Klein, Federico Rey and Cameron Currie advised on experimental design and edited the manuscript.

4.7 Figures

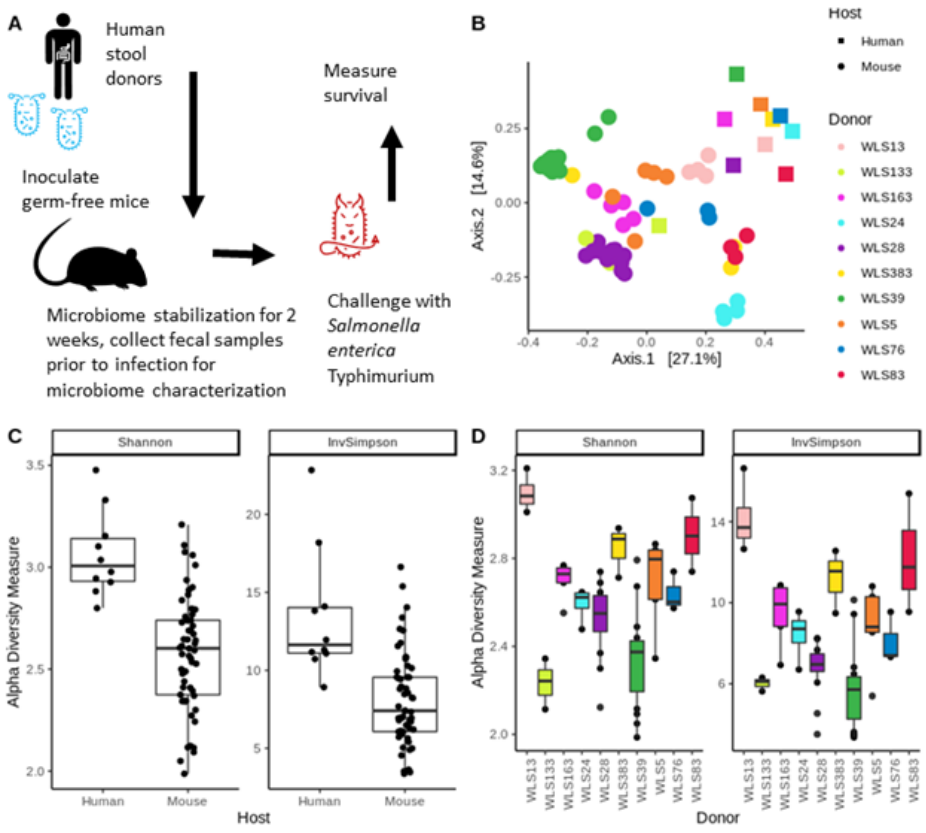


Figure 1. A. Diagram of initial experimental set up. B. PCoA of Bray-Curtis dissimilarity of microbiome taxa. Host indicated by shape (square=human, circle=mouse) and donor source indicated by color (WLS13=peach, WLS133=lime, WLS163=pink, WLS24=aqua,

WLS28=purple, WLS383=yellow, WLS39=green, WLS5=orange, WLS76=blue, WLS83=red). C. Alpha diversity measures (Shannon=right panel, inverse Simpson=left panel) split by host source. D. Alpha diversity measures from mice samples split by donor sources (Shannon=right panel, inverse Simpson=left panel). NounProject Project Symbols: bacteria by Anthony Ledoux, human by Jakob Vogel, mice by designer468

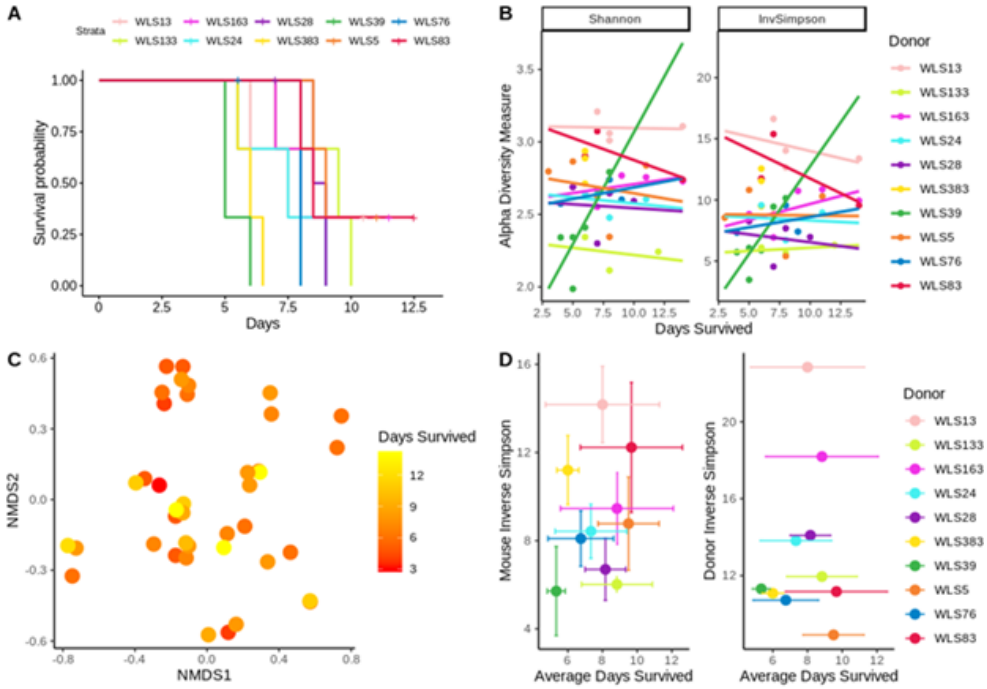


Figure 2. Mice survival with relation to metagenomic composition. A. Percent survival over days of infection, colored by donor source. B. Alpha diversity measures over days survived in mice, split and colored by donor. Linear regression estimated for each donor. Only the regression from the WLS39 donor on the inverse Simpson index was significant. C. NMDS plot of Bray-Curtis dissimilarity of mouse microbiomes, colored by days survived with infection (red=lower survival, yellow=longer survival). D. Inverse Simpson index over average days mice survived per donor. Right = Mouse inverse Simpson index. Left = Donor inverse Simpson index, note that

only one sample used per human donor.

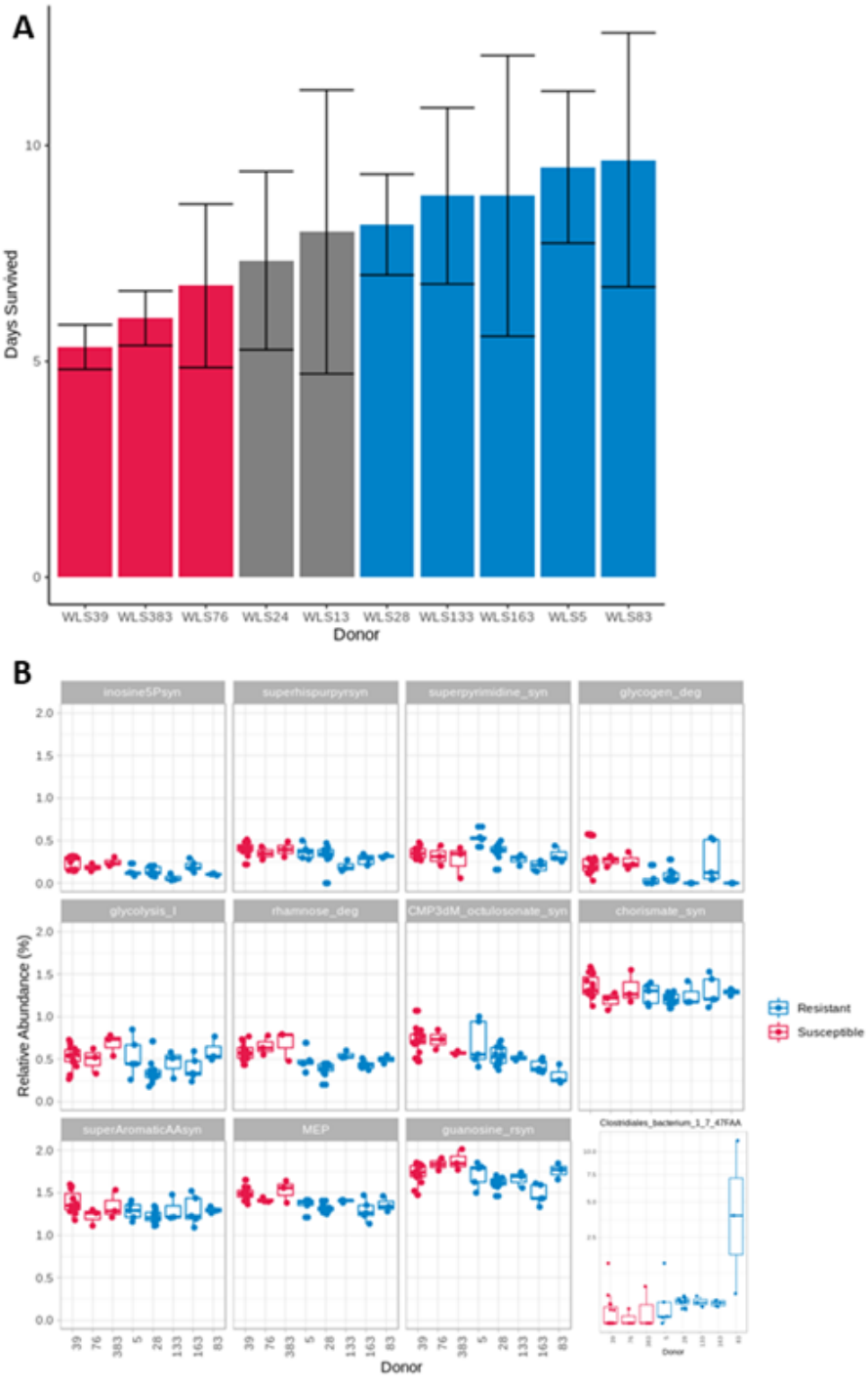


Figure 3. Differences detected in enriched versus susceptible communities. A. Days survived per donor. Bars colored in red used as susceptible donors, bars in blue used for resistant donors. Middle category not used in downstream analysis. B. Function and taxa enrichment in microbiomes in susceptible versus resistant microbiomes as identified by Lefse. Plotted here with relative abundance (%) for each mouse in each donor group. Taxon (Clostridales bacteria) plotted on square root adjusted y-axis.

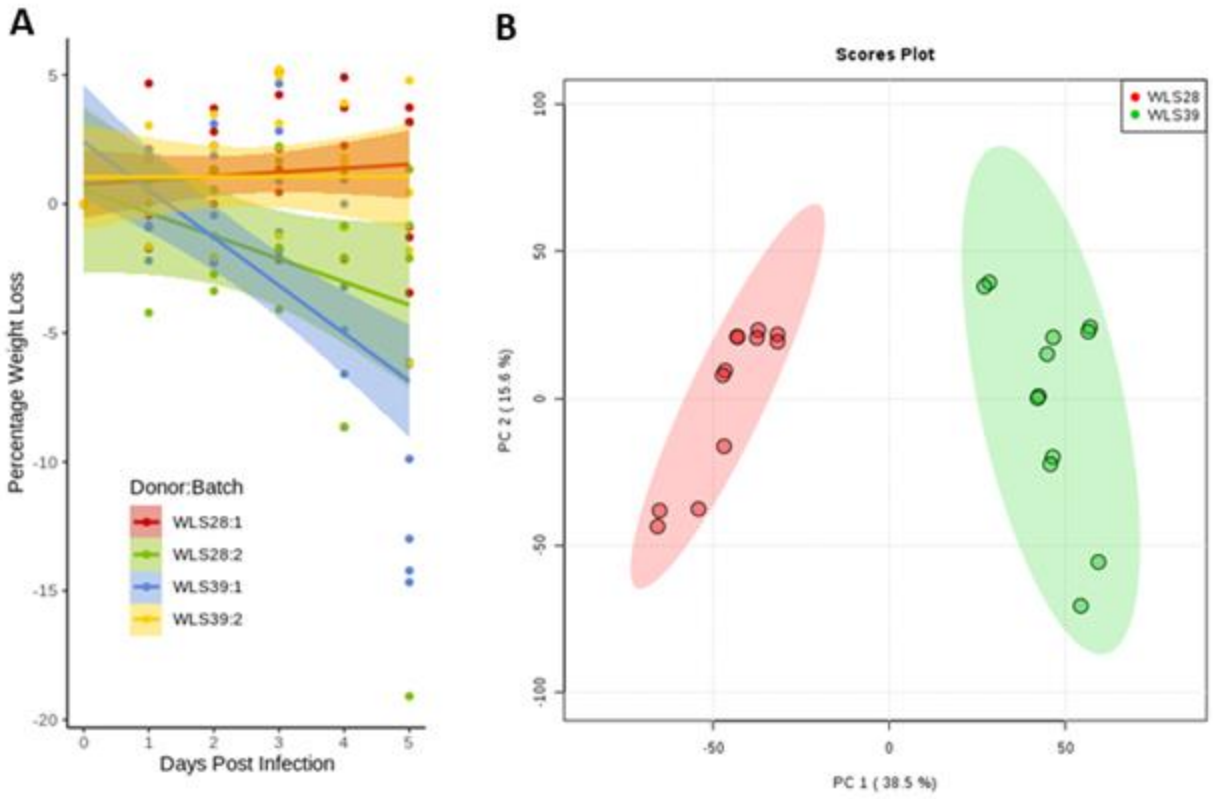


Figure 4. A. Percentage weight loss for mice in WLS28 and WLS39 group split by batch. B. Principal component analysis (PCA) for cecum metabolites prior to infection for mice from these two donors. Each point represents one technical replicate, each biological sample had two technical replicates.

4.8 References

Akhy, M. T., C. M. Brown, and D. C. Old. 1984. “L-Rhamnose Utilisation in Salmonella Typhimurium.” *The Journal of Applied Bacteriology* 56 (2): 269–74. https://doi.org/10.1111/j.1365-2672.1984.tb01347.x.

Bohnhoff, M., B. L. Drake, and C. P. Miller. 1954. “Effect of Streptomycin on Susceptibility of Intestinal Tract to Experimental Salmonella Infection.” *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)* 86 (1): 132–37. https://doi.org/10.3181/00379727-86-21030.

Boulangé, Claire L., Ana Luisa Neves, Julien Chilloux, Jeremy K. Nicholson, and Marc-Emmanuel Dumas. 2016. “Impact of the Gut Microbiota on Inflammation, Obesity, and Metabolic Disease.” *Genome Medicine* 8 (1): 42. https://doi.org/10.1186/s13073-016-0303-2.

Bratburd, Jennifer R., Caitlin Keller, Eugenio Vivas, Erin Gemperline, Lingjun Li, Federico E. Rey, and Cameron R. Currie. 2018. “Gut Microbial and Metabolic Responses to Salmonella Enterica Serovar Typhimurium and Candida Albicans.” *MBio* 9 (6). https://doi.org/10.1128/mBio.02032-18.

Brugiroux, Sandrine, Markus Beutler, Carina Pfann, Debora Garzetti, Hans-Joachim Ruscheweyh, Diana Ring, Manuel Diehl, et al. 2016. “Genome-Guided Design of a Defined Mouse Microbiota That Confers Colonization Resistance against Salmonella Enterica Serovar Typhimurium.” *Nature Microbiology* 2 (2): 1–12. https://doi.org/10.1038/nmicrobiol.2016.215.

Chung, Hachung, Sünje J. Pamp, Jonathan A. Hill, Neeraj K. Surana, Sanna M. Edelman, Erin B. Troy, Nicola C. Reading, et al. 2012. “Gut Immune Maturation Depends on Colonization with a Host-Specific Microbiota.” *Cell* 149 (7): 1578–93. https://doi.org/10.1016/j.cell.2012.04.037.

Dosselaere, F., and J. Vanderleyden. 2001. “A Metabolic Node in Action: Chorismate-Utilizing Enzymes in Microorganisms.” *Critical Reviews in Microbiology* 27 (2): 75–131. https://doi.org/10.1080/20014091096710.

Faber, Franziska, Parameth Thiennimitr, Luisella Spiga, Mariana X. Byndloss, Yael Litvak, Sara Lawhon, Helene L. Andrews-Polymenis, Sebastian E. Winter, and Andreas J. Bäumler. 2017. “Respiration of Microbiota-Derived 1,2-Propanediol Drives Salmonella Expansion during Colitis.” *PLoS Pathogens* 13 (1). https://doi.org/10.1371/journal.ppat.1006129.

Franzosa, Eric A., Lauren J. McIver, Gholamali Rahnavard, Luke R. Thompson, Melanie Schirmer, George Weingart, Karen Schwarzberg Lipson, et al. 2018. “Species-Level Functional Profiling of Metagenomes and Metatranscriptomes.” *Nature Methods* 15 (11): 962–68. https://doi.org/10.1038/s41592-018-0176-y.

Ganesh, Bhanu Priya, Robert Klopfleisch, Gunnar Loh, and Michael Blaut. 2013. “Commensal Akkermansia Muciniphila Exacerbates Gut Inflammation in Salmonella Typhimurium-

Infected Gnotobiotic Mice." *PLoS ONE* 8 (9). https://doi.org/10.1371/journal.pone.0074963.

Gillis, Caroline C., Elizabeth R. Hughes, Luisella Spiga, Maria G. Winter, Wenhan Zhu, Tatiane Furtado de Carvalho, Rachael B. Chanin, et al. 2018. "Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth." *Cell Host & Microbe* 23 (1): 54-64.e6. https://doi.org/10.1016/j.chom.2017.11.006.

Herd, Pamela, Deborah Carr, and Carol Roan. 2014. "Cohort Profile: Wisconsin Longitudinal Study (WLS)." *International Journal of Epidemiology* 43 (1): 34–41. https://doi.org/10.1093/ije/dys194.

Heuston, Sinéad, Máire Begley, Cormac G. M. Gahan, and Colin Hill. 2012. "Isoprenoid Biosynthesis in Bacterial Pathogens." *Microbiology,* 158 (6): 1389–1401. https://doi.org/10.1099/mic.0.051599-0.

Hooks, Katarzyna B., and Maureen A. O'Malley. 2017. "Dysbiosis and Its Discontents." *MBio* 8 (5). https://doi.org/10.1128/mBio.01492-17.

Hugenholtz, Floor, and Willem M. de Vos. 2018. "Mouse Models for Human Intestinal Microbiota Research: A Critical Evaluation." *Cellular and Molecular Life Sciences* 75 (1): 149–60. https://doi.org/10.1007/s00018-017-2693-8.

Kim, Sohn, April Covington, and Eric G. Pamer. 2017. "The Intestinal Microbiota: Antibiotics, Colonization Resistance, and Enteric Pathogens." *Immunological Reviews* 279 (1): 90–105. https://doi.org/10.1111/imr.12563.

Kohli, Nandita, Zeni Crisp, Rebekah Riordan, Michael Li, Robert C. Alaniz, and Arul Jayaraman. 2018. "The Microbiota Metabolite Indole Inhibits Salmonella Virulence: Involvement of the PhoPQ Two-Component System." *PLOS ONE* 13 (1): e0190613. https://doi.org/10.1371/journal.pone.0190613.

Kumar, Aman, and Vanessa Sperandio. 2019. "Indole Signaling at the Host-Microbiota-Pathogen Interface." *MBio* 10 (3). https://doi.org/10.1128/mBio.01031-19.

Lee, Jin-Hyung, and Jintae Lee. 2010. "Indole as an Intercellular Signal in Microbial Communities." *FEMS Microbiology Reviews* 34 (4): 426–44. https://doi.org/10.1111/j.1574-6976.2009.00204.x.

Li, Yanze, Wenming Cao, Na L. Gao, Xing-Ming Zhao, and Wei-Hua Chen. 2019. "Consistent Alterations of Human Fecal Microbes after Transplanted to Germ-Free Mice." *BioRxiv*, April, 495663. https://doi.org/10.1101/495663.

Lloyd-Price, Jason, Galeb Abu-Ali, and Curtis Huttenhower. 2016. "The Healthy Human Microbiome." *Genome Medicine* 8 (1): 51. https://doi.org/10.1186/s13073-016-0307-y.

Marcobal, A, P C Kashyap, T A Nelson, P A Aronov, M S Donia, A Spormann, M A Fischbach, and J L Sonnenburg. 2013. "A Metabolomic View of How the Human Gut Microbiota

Impacts the Host Metabolome Using Humanized and Gnotobiotic Mice." *The ISME Journal* 7 (10): 1933–43. https://doi.org/10.1038/ismej.2013.89.

Mayo, Baltasar, Lucía Vázquez, and Ana Belén Flórez. 2019. "Equol: A Bacterial Metabolite from The Daidzein Isoflavone and Its Presumed Beneficial Health Effects." *Nutrients* 11 (9). https://doi.org/10.3390/nu11092231.

Rawls, John F., Michael A. Mahowald, Ruth E. Ley, and Jeffrey I. Gordon. 2006. "Reciprocal Gut Microbiota Transplants from Zebrafish and Mice to Germ-Free Recipients Reveal Host Habitat Selection." *Cell* 127 (2): 423–33. https://doi.org/10.1016/j.cell.2006.08.043.

Rivera-Chávez, Fabian, and Andreas J. Bäumler. 2015. "The Pyromaniac Inside You: Salmonella Metabolism in the Host Gut." *Annual Review of Microbiology* 69 (1): 31–48. https://doi.org/10.1146/annurev-micro-091014-104108.

Rivera-Chávez, Fabian, Lillian F. Zhang, Franziska Faber, Christopher A. Lopez, Mariana X. Byndloss, Erin E. Olsan, Gege Xu, et al. 2016. "Depletion of Butyrate-Producing Clostridia from the Gut Microbiota Drives an Aerobic Luminal Expansion of Salmonella." *Cell Host & Microbe* 19 (4): 443–54. https://doi.org/10.1016/j.chom.2016.03.004.

Romano, Kymberleigh A., Kimberly A. Dill-McFarland, Kazuyuki Kasahara, Robert L. Kerby, Eugenio I. Vivas, Daniel Amador-Noguez, Pamela Herd, and Federico E. Rey. 2018. "Fecal Aliquot Straw Technique (FAST) Allows for Easy and Reproducible Subsampling: Assessing Interpersonal Variation in Trimethylamine-N-Oxide (TMAO) Accumulation." *Microbiome* 6 (May). https://doi.org/10.1186/s40168-018-0458-8.

Scallan, Elaine, Robert M. Hoekstra, Frederick J. Angulo, Robert V. Tauxe, Marc-Alain Widdowson, Sharon L. Roy, Jeffery L. Jones, and Patricia M. Griffin. n.d. "Foodborne Illness Acquired in the United States—Major Pathogens - Volume 17, Number 1—January 2011 - Emerging Infectious Diseases Journal - CDC." Accessed April 8, 2020. https://doi.org/10.3201/eid1701.p11101.

Segata, Nicola, Jacques Izard, Levi Waldron, Dirk Gevers, Larisa Miropolsky, Wendy S Garrett, and Curtis Huttenhower. 2011. "Metagenomic Biomarker Discovery and Explanation." *Genome Biology* 12 (6): R60. https://doi.org/10.1186/gb-2011-12-6-r60.

Spiga, Luisella, Maria G. Winter, Tatiane Furtado de Carvalho, Wenhan Zhu, Elizabeth R. Hughes, Caroline C. Gillis, Cassie L. Behrendt, et al. 2017. "An Oxidative Central Metabolism Enables Salmonella to Utilize Microbiota-Derived Succinate." *Cell Host & Microbe* 22 (3): 291-301.e6. https://doi.org/10.1016/j.chom.2017.07.018.

Staib, Lena, and Thilo M. Fuchs. 2015. "Regulation of Fucose and 1,2-Propanediol Utilization by Salmonella Enterica Serovar Typhimurium." *Frontiers in Microbiology* 6. https://doi.org/10.3389/fmicb.2015.01116.

Stecher, Bärbel, Samuel Chaffron, Rina Käppeli, Siegfried Hapfelmeier, Susanne Freedrich, Thomas C. Weber, Jorum Kirundi, et al. 2010. "Like Will to Like: Abundances of

Closely Related Species Can Predict Susceptibility to Intestinal Colonization by Pathogenic and Commensal Bacteria." *PLoS Pathogens* 6 (1). https://doi.org/10.1371/journal.ppat.1000711.

Strohmaier, H, P Remler, W Renner, and G Högenauer. 1995. "Expression of Genes KdsA and KdsB Involved in 3-Deoxy-D-Manno-Octulosonic Acid Metabolism and Biosynthesis of Enterobacterial Lipopolysaccharide Is Growth Phase Regulated Primarily at the Transcriptional Level in Escherichia Coli K-12." *Journal of Bacteriology* 177 (15): 4488–4500.

"Structure, Function and Diversity of the Healthy Human Microbiome." 2012. *Nature* 486 (7402): 207–14. https://doi.org/10.1038/nature11234.

Thiemann, Sophie, Nathiana Smit, Urmi Roy, Till Robin Lesker, Eric J. C. Gálvez, Julia Helmecke, Marijana Basic, et al. 2017. "Enhancement of IFNγ Production by Distinct Commensals Ameliorates Salmonella-Induced Disease." *Cell Host & Microbe* 21 (6): 682-694.e5. https://doi.org/10.1016/j.chom.2017.05.005.

Tremlett, Helen, Kylynda C. Bauer, Silke Appel-Cresswell, Brett B. Finlay, and Emmanuelle Waubant. 2017. "The Gut Microbiome in Human Neurological Disease: A Review." *Annals of Neurology* 81 (3): 369–82. https://doi.org/10.1002/ana.24901.

Velazquez, Eric M., Henry Nguyen, Keaton T. Heasley, Cheng H. Saechao, Lindsey M. Gil, Andrew W. L. Rogers, Brittany M. Miller, et al. 2019. "Endogenous Enterobacteriaceae Underlie Variation in Susceptibility to Salmonella Infection." *Nature Microbiology* 4 (6): 1057–64. https://doi.org/10.1038/s41564-019-0407-8.

Wilson, Wayne A., Peter J. Roach, Manuel Montero, Edurne Baroja-Fernández, Francisco José Muñoz, Gustavo Eydallin, Alejandro M. Viale, and Javier Pozueta-Romero. 2010. "Regulation of Glycogen Metabolism in Yeast and Bacteria." *FEMS Microbiology Reviews* 34 (6): 952–85. https://doi.org/10.1111/j.1574-6976.2010.00220.x.

Winter, Sebastian E., Parameth Thiennimitr, Maria G. Winter, Brian P. Butler, Douglas L. Huseby, Robert W. Crawford, Joseph M. Russell, et al. 2010. "Gut Inflammation Provides a Respiratory Electron Acceptor for Salmonella." *Nature* 467 (7314): 426–29. https://doi.org/10.1038/nature09415.

Wolfe, Alan J. 2015. "Glycolysis for the Microbiome Generation." *Microbiology Spectrum* 3 (3). https://doi.org/10.1128/microbiolspec.MBP-0014-2014.

Wu, Jikang, Anice Sabag-Daigle, Mikayla A. Borton, Linnea F. M. Kop, Blake E. Szkoda, Brooke L. Deatherage Kaiser, Stephen R. Lindemann, et al. 2018. "Salmonella-Mediated Inflammation Eliminates Competitors for Fructose-Asparagine in the Gut." *Infection and Immunity* 86 (5). https://doi.org/10.1128/IAI.00945-17.

Chapter 5: Gut Microbes: Good Versus Illness

The Wisconsin Initiative for Science Literacy invites doctoral candidates in science and engineering to include a chapter in their Ph.D. thesis that describes their scholarly research to non-science audiences. The goal is to explain the candidate's scholarly research and its significance to a wider audience that includes family members, friends, civic groups, newspaper reporters, program officers at appropriate funding agencies, state legislators, and members of the U.S. Congress. WISL encourages the inclusion of such chapters in all Ph.D. theses everywhere through the cooperation of Ph.D. candidates and their mentors.

Symbiosis is defined by relationships. Symbiosis refers to any unlike organisms living together, and the relationships between these organisms can vary widely. If two partners benefit each other, like bees getting nutrients while pollinating flowers, they are called mutualists. If one partner exploits the other, like ticks biting hosts to feed on blood and harming the host, the relationship is parasitic. The bulk of our relationships with microbes is beneficial or at least not particularly damaging. Although most microbes are not harmful to us, those that are (often known as pathogens) may have a terrible impact on our health: *Salmonella*, norovirus, influenza, etc. In the current coronavirus pandemic of 2020, these impacts are not limited to our individual health, but even our collective societal functions.

Though often more attention is paid to our microbial nemeses, microbes can also be our best defenders against pathogens. If you count up all the cells of our bodies, approximately half of those cells are microbial, not human (Sender, Fuchs, and Milo 2016). Most of those microbes reside in the gut and are collectively known as the gut microbiome or microbiota. With large numbers, and large diversity (hundreds of gut microbial species may be found in one person)

(Qin et al. 2010), come many interactions: microbes interacting with our bodies, and microbes interacting with other microbes, including pathogens.

In my work, I have explored how these microbes respond to infection with *Salmonella*. *Salmonella* is a group of pathogens that cause illnesses including food poisoning and typhoid fever. Even as early as the 1950s researchers found that beneficial microbes had an effect on *Salmonella*. Early studies in mice showed that mice had much less resistance to *Salmonella* when treated in advance with antibiotics, which disrupt the existing microbes in the gut (Bohnhoff, Drake, and Miller 1954).

Today, we know a great deal more about *Salmonella'*s interactions in the gut environment. *Salmonella*, a rabble-rouser in the gut, first triggers the immune system, causing inflammation. The body releases reactive chemicals containing oxygen which disturbs the normally low-oxygen environment of the gut. Oxygen is highly reactive and can kill cells by damaging cell walls, which people rely on when using hydrogen peroxide (H_2O_2) to treat a cut for bacteria. The wily *Salmonella* bacteria conveniently take advantage of the newly released chemicals and the disrupted gut environment, growing to large numbers in the gut so *Salmonella* can then be shed and transmitted via the fecal-oral route (for example, preparing food after not washing hands in the bathroom) to the next unlucky host.

From the defensive microbes' perspective, this situation is less than ideal. A better outcome for us and our beneficial microbes is if the microbes prevent *Salmonella* from gaining a foothold in the gut. The microbes have many options--they can try to change the immune system's response, take up space and food, and make compounds that stop the growth or control the pathogen. To picture this on a macroscopic scale, you can imagine the efforts to maintain a garden against weeds. Some of your plants may naturally outcompete weeds, perhaps by shading

them or using up the nutrients in the plot fastest. Garden plots also benefit from hand-weeding, which we can imagine as the equivalent of the immune system role. Plants also have their own chemical warfare from chemicals they produce to prevent growth of other species (also called allelopathy) akin to herbicides. On the microscopic scale, chemical battles are especially potent in bacterial competition since microbes are excellent chemical engineers, with incredibly unique and diverse enzymes for making different compounds. For this reason, microbes are also a major source of antibiotics and other drugs (Chevrette and Currie 2019).

Figure 1. Housing of germ-free mice

With the help of Dr. Federico Rey's lab, I used germ-free mice--laboratory mice kept in sterile bubbles or cages, without any outside contact to any microbes (Figure 1). Using these mice allowed me to colonize them with whatever microbes I wanted. In my first set of experiments, I gave them specific strains of bacteria that had been previously isolated from humans in order to "humanize" the mice. A few mice I left germ-free.

After 2 weeks waiting for these communities to stabilize in the mice, I infected some of the mice with microbiota and the still germ-free mice with *Salmonella enterica* Typhimurium (a strain that infects both mice and humans, although it causes somewhat different symptoms). By comparing these two groups, I could find compounds made during infection only when the microbiota was present. In addition, I had a third group of mice with a microbiota that were not infected so I could eliminate compounds made normally by the microbes and focus on those made during infection (Figure 2).

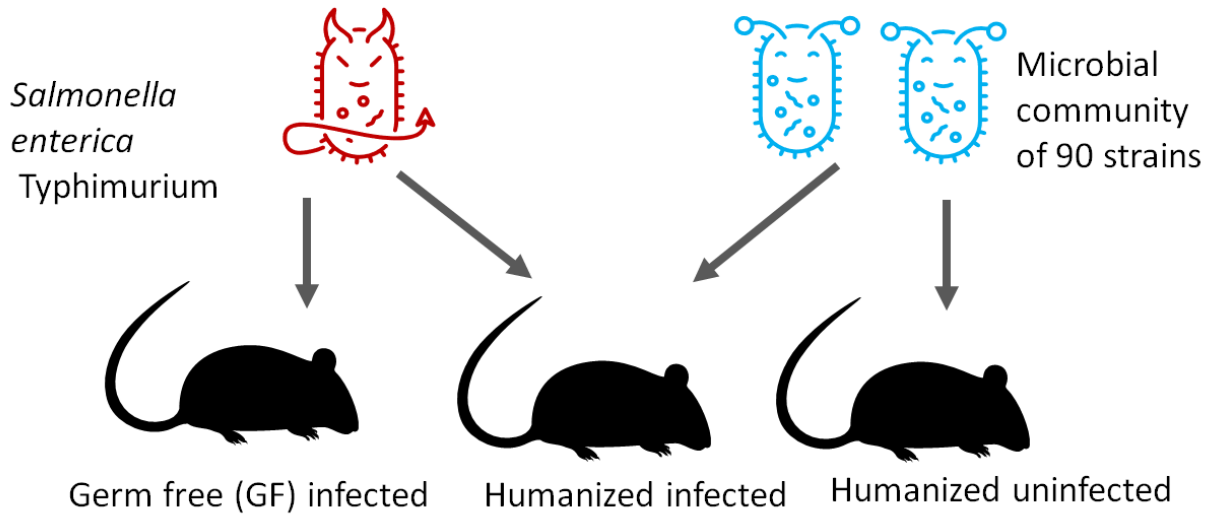


Figure 2. Diagram of experiment setup

By collaborating with Dr. Lingjun Li's lab, I was also able to assess compounds in the guts that are found when both the microbiome and *Salmonella* are present. We used liquid-chromatography mass spectrometry, which can be thought of as splitting up all the compounds in a sample and then measuring their weight (more accurately mass/charge ratios). Many of these compounds are difficult to identify as their "weight" does not match anything in databases of known compounds. Fortunately, we found matches for a few compounds, and could identify them by comparing each to a reference. Of these, two were from the glutathione pathway. Glutathione is an antioxidant, which can help protect from the immune system's reactive chemicals with oxygen. Potentially, the gut microbes may regulate and produce these metabolites that may impact the infection.

By sequencing DNA from the feces collected over three days of infection, I could get a sense of which microbes were most abundant. I found that without an infection, the microbial communities stayed fairly consistent, but with *Salmonella* they rapidly changed. As had been seen by other researchers, microbes that are more related to *Salmonella* were enriched in the samples after infection. These microbes have similarities in their metabolism to *Salmonella*, perhaps most importantly their ability to tolerate an environment with oxygen (as most of the other gut bacteria live strictly without oxygen).

In the work I have just described, we used one representative microbiome with lab grown strains mixed together. However, each of us has our own individual communities of microbes. This variation might help explain how, with the help of their microbiomes, some people are better able to resist infection than others. How might these different microbiomes with their different strains of bacteria affect which metabolites are produced and our ability to resist disease?

To explore differences among people, I used human microbiome samples (poop) and colonized the mice with these different samples. Then I infected the mice with *Salmonella* and measured how long the mice survived. I found that some people's microbiomes protected the mice better against infection. In addition, I collected samples prior to infection to gain insight into how the microbiome plays a role in preventing *Salmonella* from colonizing, rather than how microbial communities changed after colonization.

From DNA sequencing, I could compare several of the protective microbiomes to see what they shared. I only found a single microbial species that was shared by each of the protective microbiomes, but was not present in the susceptible microbiomes. In the susceptible microbiomes, I also found that several gene pathways were enriched, including those responsible for degradation of the sugar rhamnose and for creating basic components for cell growth, such as purines, which are compounds used for many things including building DNA.

I compared in detail one of the best communities against one of the worst and found that several different metabolites were enriched in one over the other, although these compounds were different from the kinds I had seen in my previous work. Some of them may have derived from microbial breakdown of soy products. At this point it is unclear if these metabolites play a role in resistance to infection or just happen to be produced in different abundances by the different microbiomes.

Overall, these projects helped us identify microbes and metabolites that may play a role in defending us from *Salmonella* infection. In the future, more experiments could study if compounds identified here play a role in infection and whether those compounds have any potential therapeutic use. The study in which I examined different human microbiomes suggests that there are many compounds and microbial functions that may play a role during infection.

References

Bohnhoff, M., B. L. Drake, and C. P. Miller. 1954. “Effect of Streptomycin on Susceptibility of Intestinal Tract to Experimental Salmonella Infection.” *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)* 86 (1): 132–37. https://doi.org/10.3181/00379727-86-21030.

Chevrette, Marc G., and Cameron R. Currie. 2019. “Emerging Evolutionary Paradigms in Antibiotic Discovery.” *Journal of Industrial Microbiology & Biotechnology* 46 (3): 257–71. https://doi.org/10.1007/s10295-018-2085-6.

Qin, Junjie, Ruiqiang Li, Jeroen Raes, Manimozhiyan Arumugam, Kristoffer Solvsten Burgdorf, Chaysavanh Manichanh, Trine Nielsen, et al. 2010. “A Human Gut Microbial Gene Catalogue Established by Metagenomic Sequencing.” *Nature* 464 (7285): 59–65. https://doi.org/10.1038/nature08821.

Sender, Ron, Shai Fuchs, and Ron Milo. 2016. “Revised Estimates for the Number of Human and Bacteria Cells in the Body.” *PLOS Biology* 14 (8): e1002533. https://doi.org/10.1371/journal.pbio.1002533.

Chapter 6: Conclusions and Future Directions

Microbes have substantial impacts on their hosts, with both positive and negative impacts on host fitness. In this dissertation, I explored how host-microbe interactions impact colonization across a spectrum of mutualistic and pathogenic relationships, with systems of varying complexity. Metagenomic and metabolomic approaches enable an opportunity to discover new mediators at the host-microbe interface and explore the capacity of microbes to make unique compounds and engage in unusual metabolic functions. My work suggests ways in which microbes may influence host defense against pathogens via specific metabolites and gene pathways. Examining these microbe-microbe interactions and microbe-host interactions can help with understanding how microbes contribute to our health and developing microbiome-based therapeutics.

In Chapter 1, I discussed the broad implications of defensive microbes, with a focus on social insects and humans and how different species promote colonization of beneficial microbes that limit colonization of pathogens. This chapter discusses various strategies of social insects and humans to fight pathogens with a focus on the role of defensive microbes. I point out several key open areas of investigation in the field: how beneficial microbes are transmitted and maintained while minimizing the spread of pathogens (also discussed in Chapter 2), and how perturbations change the host and microbial response to pathogens (also discussed in Chapters 3 and 4). Upon reflection during a time of social distancing due to SARS-CoV-2, I find these questions to be even more pressing and fascinating. In the review, we argue that insects are good models to use to address these important questions. In Appendix 4, I elaborate further on

different animal models for microbiome research. In particular, I highlight the flexibility using microbial communities of defined complexity in gnotobiotic animals.

In Chapter 2, I explore colonization of a defensive mutualist *Pseudonocardia* using the fungus-growing ant system. In this ancient symbiosis, we know from previous research that the ant appears to have developed structural modifications to support *Pseudonocardia* and behaviors to transmit the bacteria, yet less is known about the bacteria's adaptations to the host. By comparing genomes of strains isolated from ants to those isolated from a variety of different environmental sources, I find some indication of genome reduction in the ant-associated strains which have reduced genome lengths and are missing some gene clusters. I also find a variety of biosynthetic gene cluster potential. This analysis is somewhat limited by how many ant-strains fall within two related clades. Further efforts to isolate *Pseudonocardia* from non-ant sources may help illuminate the full extent of diversity in the genus. By experimentally colonizing the ants during the brief period that they acquire *Pseudonocardia* with a variety of ant-associated and non-ant associated strains, I identified that strains outside of these ant-associated clades can colonize ants but do so somewhat erratically. My colonization experiments together with the genomic evidence indicates that while ant likely control the relationship, ant-associated *Pseudonocardia* may have concordantly developed adaptations to their hosts. As with other defensive symbionts, typically genome reduction is limited in comparison to nutritional symbionts. This may suggest a lack of strong selective pressure on *Pseudonocardia* from the ants, potentially because those kinds of pressures could diminish *Pseudonocardia*'s ability to produce antibiotics, or because ant-associated *Pseudonocardia* have lifestages where they end up surviving in the environment. The ant's ability to maintain *Pseudonocardia* in their colonies is perhaps more interesting in light of the lack of adaptations from *Pseudonocardia.*

In Chapters 3 and 4, I investigated interactions between the human gut microbiome and *Salmonella enterica* Typhimurium in a germ-free mouse model. As an early indication of the complexities of the microbiome, germ-free mammalian models were first developed in 1895, but maintaining these animal colonies only became possible by the 1950's with better housing systems and an understanding how to supplement the diet with nutrients normally derived from the gut microbiota (Wostman 1996). In Chapter 3, I use a defined consortium of bacteria, and employ metagenomic sequencing to determine that bacterial communities in infected mice change with a higher representation of non-*Salmonella* Enterobacteriaceae. In addition, we used liquid chromatography-mass spectrometry on ceca collected three days post-infection to identify metabolites, including glutathione metabolites that may be related to oxidative stress occurring during infection. As *Salmonella* induces a large amount of oxidative stress by triggering inflammation, the microbes in the gut might be responding to this general perturbation rather than a specific reaction to one microbe. On the other hand, we were unable to identify a majority of the metabolites made during infection, so other many metabolite-mediated interactions may be occurring as well and worth exploring in future studies.

In Chapter 4, instead studying one synthetic community, I used donor human stool samples to colonize germ-free mice and interrogate how variability in the microbiome impacts resistance to infection. With this approach, I identified several communities more resistant to *Salmonella* and used metagenomic sequencing to find similarities in gene content of the *Salmonella*-resistant versus susceptible communities, revealing one shared strain among resistant communities and several shared functional categories among susceptible communities. In contrast to Chapter 3 where I investigated changes across infection and focused on changes near the end of infection, in this study I focused on the comparing the communities prior to infection.

By looking at the communities prior to perturbation, I would have a more acute perspective on resistance to infection rather than resilience. I found one strain enriched in resistant communities versus susceptible communities, which may indicate more redundancy and variation of human gut microbiota strains rather than the importance of this one strain. For gene pathways, I identified several overrepresented in the susceptible microbiomes versus the resistant microbiomes, including both anabolic pathways (including purines, chorismite, and CMP-3-deoxy-D-manno-octulosonate) and catabolic pathways (including glycogen and rhamnose, a sugar found bacterial, plants cells and mucin oligosaccharides). I did not find a significant relationship of survival with community diversity. To explore different communities in more depth, I chose one resistant and one susceptible community and found that prior to infection they produce a variety of different metabolites, but the proportion of immunological cells present is similar. Notably, many of the pathways and metabolites identified in this chapter differed from the findings in Chapter 3 (although inosine monophosphate appeared in both chapters), which may indicate the high variation in in the gut microbiomes as well as different potential mechanisms for preventing disease rather than responding to perturbation from infection. Overall, these findings suggest a great deal of complexity in the metabolic potential of the gut microbiome and its impact on pathogen resistance.

Future directions could include testing specific metabolites or pathways of interest. With the fungus-farming ant model, follow-ups could explore increasingly complexity by colonizing with two different strains at the size time and use qPCR to distinguish which colonizes better, to add the microbe-microbe dimension of colonization dynamics. Developing genetic manipulation of *Pseudonocardia* would enable knocking out specific genes or pathways to test their impact on colonization. With follow-up in mice, multiple opportunities are available, including creating

synthetic communities to test specific impact of presence or absence of specific functions. Likewise, the treatment with particular metabolites to see the impact on host and microbiota.

Many challenging questions remain in elucidating microbe-microbe interactions and microbe-host interactions. Increasingly high-throughput and low-cost sequencing has enabled new tools for characterizing microbiomes, as described in Appendix 4. This work helps identify possible avenues of microbial interactions and the variability of different strains in their ability to colonize hosts and with different microbiomes and their ability to prevent pathogen colonization. While much work has focused on interactions between individual hosts under relatively static conditions, species interactions often stretches beyond these narrow laboratory conditions. As suggested in Chapter 1, to better understand microbial lifestyles, the interactions between microbes and multiple hosts is increasing worth consideration. For example, acquisition and transmission dynamics of a whole ant colony rather than focusing on the single ant. In another example of increasing complexity, Chapter 3 relies on a simplified synthetic community, while Chapter 4 uses complex donor communities. Yet both of these projects still maintain laboratory-reared mice in precise, defined environments and examine only the perturbation caused by infection with *Salmonella.* In addition, hosts and microbes exist in fluctuating and sometimes stressful environments with a variety of perturbations. While studying additional variables in systems that already contain a vast amount of complexity is challenging, examining variation, perturbation, and further interactions would help find the extremes of microbes ability to colonize and impact their environment under various conditions. Ultimately, understanding how microbiomes respond to perturbations will help identify key features of microbiomes with important implications for their use in medicine and technology.

Appendix 1: Supplemental Material for Chapter 2

Strain	length (bp)	GC content	Completion (%)	Redundancy (%)	num_genes	avg_gene_length	num_genes_per_kb	Contigs	N50
Pseudo_hierapolitana	8856958	0.726	100	7.194	8391	957.4	0.95	1	8856958
Strept_coelicolor	8667207	0.721	98.6	5.755	7753	992.1	0.89	1	8667507
Pseudo_dioxanivorans_CB1190	7096571	0.733	99.3	12.95	6667	957.3	0.94	1	7096571
Pseudo_ICBG100	6776068	0.739	99.3	7.194	6212	995.3	0.92	1	6776068
Pseudo_JSC141020_01	7303443	0.736	99.3	6.475	6642	990.7	0.91	4	6658632
Pseudo_CC030328_06	6739435	0.719	99.3	7.914	6249	978.9	0.93	2	6654308
Pseudo_GB151026_04_3	7703877	0.731	100	7.194	7039	978	0.91	6	6527214
Pseudo_EC080529_01	7343270	0.733	99.3	5.036	6839	962.7	0.93	2	6524406
Pseudo_sediminis_DSM45779	6612042	0.729	100	4.317	6215	958.5	0.94	2	6501404
Pseudo_CC151027_05_1	6981162	0.734	99.3	7.914	6484	966.8	0.93	4	6484431
Pseudo_GB151021_07_4	7131515	0.734	99.3	7.194	6655	962.3	0.93	4	6483020
Pseudo_GB151026_03_3	7557021	0.732	100	6.475	6948	973.6	0.92	5	6427668
Pseudo_AL050505_11	6566921	0.738	99.3	7.914	6057	984.3	0.92	3	6389648
Pseudo_EV151028_01_1	6738030	0.735	99.3	6.475	6293	969	0.93	5	6372068
Pseudo_Ae707_Ps1	6832450	0.736	99.3	7.194	6274	983.4	0.92	4	6361983
Pseudo_JSC111027_01	6362052	0.739	99.3	5.036	5840	992.2	0.92	2	6342421
Pseudo_GB151021_02_4	6325603	0.747	99.3	3.597	5744	1012.5	0.91	1	6325603
Pseudo_SP020602_02	6322523	0.74	98.6	6.475	5777	995.7	0.91	1	6322523
Pseudo_Ae717_Ps2	7431781	0.732	100	5.755	7046	932.8	0.95	26	6288042
Pseudo_EV151025_05_1	6972056	0.734	99.3	6.475	6501	961.1	0.93	3	6255433
Pseudo_EV151025_04_4	6656636	0.737	99.3	5.036	6127	982.2	0.92	3	6248279
Pseudo_CC0031209_02	6452094	0.738	100	4.317	5948	982.1	0.92	3	6222459
Pseudo_HH130630_07	6680451	0.737	100	7.194	6241	959.1	0.93	3	6186048
Pseudo_TRS120623_01	6318161	0.738	100	5.755	5796	988.7	0.92	2	6184572
Pseudo_EV151025_09_4	6841670	0.735	99.3	6.475	6371	965	0.93	4	6182302
Pseudo_EV151025_09_1	7068612	0.734	99.3	6.475	6573	964.2	0.93	5	6182243
Pseudo_Ae356_Ps1	6451889	0.736	99.3	5.036	6162	940.9	0.96	13	6162815
Pseudo_Ae263Ps1_SC1	6512444	0.736	99.3	5.036	6189	945.8	0.95	14	6157670
Pseudo_EC080619_01	7205934	0.733	99.3	6.475	6736	960.4	0.93	3	6149338
Pseudo_EC080529_05	6236915	0.74	98.6	7.914	5960	940.1	0.96	2	6146021

Pseudo_AL041005_10	6143341	0.744	99.3	7.914	6100	898.3	0.99	1	6143341
Pseudo_EC080610_09	7131853	0.733	99.3	5.036	6656	962.8	0.93	3	6138223
Pseudo_EC080625_04	6554452	0.735	99.3	6.475	6170	957.7	0.94	3	6135769
Pseudo_Ae150A_Ps1	6388870	0.737	99.3	5.036	6069	947.8	0.95	8	6135547
Pseudo_Ae706_Ps2	6762097	0.734	100	4.317	6336	957.9	0.94	21	6111893
Pseudo_UGM030327_02	6104372	0.744	100	6.475	5691	982	0.93	1	6104372
Pseudo_Ae406_Ps2	6502255	0.737	100	4.317	6108	958.9	0.94	27	6080519
Pseudo_Ae331Ps2_SC1	6433125	0.737	100	4.317	6015	967.9	0.94	23	6075932
Pseudo_ICBG1034	5883796	0.734	97.1	9.353	6459	820.4	1.1	1	6072940
Pseudo_HH130629_09	6058802	0.736	99.3	2.878	5717	949.5	0.94	1	6058802
Pseudo_Ae505_Ps2	6389892	0.736	100	4.317	5989	956.1	0.94	13	6031156
Pseudo_AL041002_03	5978138	0.744	100	5.036	5607	975.1	0.94	1	5978138
Pseudo_UGM030330_05	5798895	0.741	100	4.317	5522	954.1	0.95	2	5736366
Pseudo_alni_PB	5994807	0.742	100	5.755	5567	977.8	0.93	3	5686562
Pseudo_CTL110912_03	5862849	0.742	100	5.036	5633	939.8	0.96	3	5531428
Pseudo_Ae168_Ps1	6509772	0.736	99.3	5.036	6206	942.9	0.95	14	5473146
Gordonia_SID5947	5099185	0.666	99.3	2.158	4760	976.2	0.93	2	5011469
Pseudo_ST040116_010	6613040	0.737	100	5.755	6114	977.8	0.92	7	4256075
Pseudo_endophytica	7567994	0.728	100	10.791	6994	982.9	0.92	2	4021098
Pseudo_cypriaca	8279222	0.727	100	6.475	7844	969.3	0.95	3	2964811
Pseudo_kunmingensis_DSM45301	9138787	0.734	99.3	8.633	8624	963.8	0.94	10	1566035
Pseudo_acaciae_DSM45401	9931328	0.723	100	3.597	9556	963.8	0.96	94	5060 97
Pseudo_spinosispora_DSM44797	9537556	0.694	100	7.914	8566	1002.6	0.9	73	3587 47
Pseudo_autotrophica_DSM43083	5830096	0.742	100	5.755	5422	978.9	0.93	30	3456 77
Pseudo_15845	8568280	0.691	98.6	10.791	7975	977.8	0.93	81	2857 12
Pseudo_autotrophica_NRRLB16064	8023843	0.735	98.6	5.036	7672	948.9	0.96	153	2658 65
Pseudo_MH_G8	10179404	0.726	100	11.511	9415	980.3	0.92	80	2595 14
Pseudo_oroxyli_CGMCC	6111570	0.73	99.3	4.317	5879	948.2	0.96	53	2573 44
Pseudo_thermophila_DSM43832	6098214	0.729	100	5.755	5878	944.1	0.96	47	2084 03
Pseudo_N23	6536078	0.725	98.6	14.388	6278	926.8	0.96	173	1834 52
Pseudo_asaccharolytica_DSM44247	5056835	0.718	97.8	9.353	4986	908.6	0.99	72	1731 67
Pseudo_hydroxycarbonoxydans	5290052	0.745	100	5.036	5311	922.3	1	89	1532 06
Pseudo_ammonioxydans_CGMCC	7361511	0.735	100	7.194	7059	938.9	0.96	167	1524 32
Pseudo_CC030328_06ill	6763995	0.719	99.3	8.633	6436	955.9	0.95	196	1501 55

Pseudo_SCN73_27	6952598	0.726	99.3	11.511	6821	924.7	0.98	99	147618
Pseudo_autotrophica_DSM535	7351656	0.73	99.3	10.072	6954	959.4	0.95	117	145790
Pseudo_alaniphila	9377632	0.704	98.6	7.914	8748	943.2	0.93	217	124429
Pseudo_73_21_SCN	7090563	0.726	98.6	5.036	7118	900.5	1	111	110501
Pseudo_CNS_004	9203094	0.726	95	15.827	10490	742.8	1.14	156	110041
Pseudo_AL050513_04	9079058	0.719	100	4.317	8621	960.8	0.95	313	79079
Pseudo_sulfidoxydans_NBRC16205	6576285	0.726	99.3	10.072	6269	947.9	0.95	209	69528
Pseudo_AL041002_03ill	6114861	0.738	100	7.194	6034	919.4	0.99	542	66675
Pseudo_CNS_139	7125388	0.742	94.2	15.108	8523	694.6	1.2	250	652771
Pseudo_hierapolitanaill	8771874	0.725	100	7.914	8694	913.8	0.99	495	437344
Pseudo_ICBG618	8707602	0.724	100	11.511	8769	906	1.01	536	412771
Pseudo_spH69	8610472	0.723	100	10.072	8921	868.6	1.04	584	350445
Pseudo_colA	6485345	0.727	99.3	8.633	6661	868.4	1.03	1487	348532
Pseudo_compacta	6787892	0.729	99.3	5.755	6851	884.7	1.01	767	319340
Pseudo_saturnea_NRRLB16172	7411587	0.728	100	10.791	7359	910.8	0.99	647	312655
Pseudo_10385	6142889	0.744	100	5.036	6142	917.9	1	510	249866
Pseudo_ICBG157	6262428	0.739	99.3	7.194	6373	896	1.02	509	245550
Pseudo_10165	6229613	0.739	100	10.791	6285	897.8	1.01	576	242977
Pseudo_endophyticaill	7487432	0.724	100	12.23	7843	863.8	1.05	1515	227766
Pseudo_SCN72_86	6554597	0.72	41	2.878	6697	876.5	1.02	396	219977
Pseudo_colD	7020216	0.734	100	14.388	6796	937.1	0.97	642	210550
Pseudo_chloro3	8307733	0.722	100	9.353	7867	941.3	0.95	786	192221
Pseudo_kujensis	8002107	0.726	98.6	7.914	8205	859.5	1.03	936	190988
Pseudo_chloroethenivorans	13597444	0.719	100	112.23	14726	817.1	1.08	3227	179866
Pseudo_EC080618_06	7107275	0.733	98.6	6.475	7392	862.4	1.04	990	164633
Pseudo_SCN72_51	5468042	0.724	31.7	6.475	5644	873.8	1.03	408	163866
Pseudo_zijingensis	8615623	0.73	98.6	9.353	8889	887	1.03	1076	161679
Pseudo_EC080610_09ill	7201661	0.732	99.3	6.475	7550	855.3	1.05	1133	150077
Pseudo_ICBG1111	5552254	0.735	100	5.036	5741	867.4	1.03	691	147688
Pseudo_ICBG1043	8160022	0.722	99.3	10.791	8821	843.1	1.08	1285	144778
Pseudo_ICBG101	6366279	0.742	99.3	5.755	6892	845	1.08	812	143551
Pseudo_EC080619_09	7024805	0.733	99.3	7.194	7319	862.7	1.04	1039	141822
Pseudo_P1	6388771	0.732	99.3	7.914	6659	865.9	1.04	875	141449
Pseudo_cypriacaill	8016379	0.725	99.3	7.914	9132	802.9	1.14	1334	141099

Pseudo_halophobica	7844075	0.726	99.3	6.475	8296	843.8	1.06	1121	13613
Pseudo_EC090830_01	6212914	0.739	100	8.633	6531	862.8	1.05	916	13524
Pseudo_EC080524_04	6371034	0.738	99.3	5.755	6506	892.1	1.02	903	13395
Pseudo_ICBG102	6151881	0.742	98.6	5.755	6878	819.6	1.12	895	13141
Pseudo_ICBG1146	6261585	0.737	100	7.194	6648	849.8	1.06	1069	13061
Pseudo_EC080618_05	7101947	0.732	99.3	7.194	7438	857.1	1.05	1127	12991
Pseudo_EC080620_06	6248385	0.747	99.3	5.036	6485	888.2	1.04	990	12843
Pseudo_colB	6495850	0.728	100	15.827	6953	838.8	1.07	1450	12682
Pseudo_ICBG158	6324655	0.742	100	6.475	7116	812.3	1.13	952	12482
Pseudo_EC080529_20	6861062	0.735	99.3	7.914	7305	850.6	1.06	1260	12349
Pseudo_EC080524_13	6209446	0.747	100	5.755	6493	881.6	1.05	1095	12102
Pseudo_ICBG1050	6538952	0.736	99.3	9.353	6841	858.3	1.05	1263	11572
Pseudo_ICBG1042	5542324	0.734	99.3	5.036	5855	847	1.06	911	11542
Pseudo_petroleophila	6304703	0.737	98.6	7.194	7801	746.8	1.24	1109	11327
Pseudo_EC080617_07	6939821	0.733	98.6	6.475	7413	841	1.07	1271	11292
Pseudo_EC080603_07	6481475	0.738	99.3	6.475	6815	864.7	1.05	1109	11157
Pseudo_SID8383	5581429	0.739	98.6	7.194	6328	801.4	1.13	950	11084
Pseudo_EC080529_16	6427229	0.736	98.6	6.475	6833	850.7	1.06	1210	10977
Pseudo_EC080529_15	6433317	0.737	99.3	7.194	6933	842.3	1.08	1259	10854
Pseudo_EC070717_09	6351919	0.738	100	7.194	6788	851.1	1.07	1128	10828
Pseudo_EC080524_14	6195412	0.737	99.3	7.194	6538	863.4	1.06	1087	10374
Pseudo_EC080623_03	6881261	0.734	99.3	7.914	7373	842.8	1.07	1334	10164
Pseudo_EC080525_06	6836013	0.735	99.3	6.475	7397	834.6	1.08	1385	9998
Pseudo_UGM030402_04	6381659	0.741	100	7.914	7060	829.3	1.11	1274	9977
Pseudo_CC060123_03	6002905	0.746	100	4.317	6459	856.1	1.08	1209	9926
Pseudo_P2	7571635	0.731	99.3	33.094	8318	821.5	1.1	1617	9902
Pseudo_EC080617_04	6932605	0.734	98.6	7.914	7484	834.5	1.08	1412	9808
Pseudo_EC061022_05	6342860	0.745	100	5.036	6824	853.5	1.08	1279	9736
Pseudo_EC080524_01	6189181	0.736	99.3	7.194	6701	839.4	1.08	1282	9548
Pseudo_SP030328_02	6174713	0.744	100	4.317	6759	838.4	1.09	1205	9512
Pseudo_EC080625_04ill	6308602	0.736	99.3	6.475	7109	803.3	1.13	1619	9505
Pseudo_EC080525_05	6068818	0.739	100	7.914	6621	834.7	1.09	1349	9480
Pseudo_EC070720_06	6025721	0.746	100	4.317	6566	844.8	1.09	1289	9394
Pseudo_UGM030330_04	5853058	0.732	100	5.036	6490	809.7	1.11	1335	9192

Strept_10815	8118344	0.726	99.3	10.072	8609	824.8	1.06	2165	9188
Pseudo_ICBG1142	6275709	0.729	100	5.036	6824	823.8	1.09	1376	9002
Pseudo_ICBG602	6062549	0.739	97.8	5.036	6636	834.5	1.09	1232	8991
Pseudo_ICBG1143	6463016	0.733	98.6	9.353	6898	837	1.07	1465	8988
Pseudo_EC080620_01	6202257	0.746	100	4.317	6804	839.8	1.1	1413	8773
Pseudo_EC080617_15	6789668	0.734	99.3	7.194	7392	827	1.09	1475	8761
Pseudo_EC080618_04	6789668	0.734	99.3	7.194	7392	827	1.09	1475	8761
Pseudo_EC090828_04	6159993	0.739	100	9.353	6882	812.8	1.12	1379	8729
Pseudo_ICBG1145	6393887	0.736	98.6	8.633	6785	846.6	1.06	1482	8706
Pseudo_ICBG1125	6442816	0.734	98.6	10.072	6889	837.4	1.07	1559	8653
Pseudo_EC080529_19	6745963	0.734	99.3	5.755	7529	811.7	1.12	1626	8650
Pseudo_ICBG103	6072940	0.737	99.3	4.317	5656	975.5	0.93	1277	8642
Pseudo_SP030327_02	6563355	0.741	100	9.353	7143	837.5	1.09	1543	8329
Pseudo_SP020602_02ill	6178162	0.739	100	7.914	6699	843	1.08	1405	8274
Pseudo_ICBG93	6174235	0.738	98.6	5.755	7670	733.2	1.24	1415	8254
Pseudo_EC060123_09	6201792	0.738	100	6.475	6917	816.3	1.12	1463	8217
Pseudo_ICBG161	5970867	0.739	98.6	5.755	7214	756.6	1.21	1259	8214
Pseudo_AL040118_01	6267738	0.736	100	7.194	6980	817.2	1.11	1588	7896
Pseudo_CC020602_01	6053798	0.738	100	6.475	6666	828	1.1	1450	7854
Pseudo_EC080624_04	6961708	0.732	100	7.194	7896	792.9	1.13	1962	7831
Pseudo_EC080620_04	6182267	0.746	100	4.317	6960	818.2	1.13	1591	7816
Pseudo_ICBG1144	5910964	0.731	97.8	7.194	6813	776.7	1.15	1372	7712
Pseudo_ICBG162	6122124	0.738	98.6	6.475	7793	716.1	1.27	1536	7586
Pseudo_EC080618_12	7021027	0.731	99.3	7.194	7949	793.2	1.13	1858	7448
Pseudo_CC031212_01	5880013	0.737	99.3	7.194	6778	790	1.15	1736	7215
Pseudo_EC080618_17	7014064	0.732	99.3	7.914	7924	794.9	1.13	1842	7082
Pseudo_EC070717_12	6293749	0.744	100	5.036	7209	801.8	1.15	1768	6903
Pseudo_EC080525_24	6269987	0.736	99.3	7.194	7067	808.3	1.13	1629	6880
Pseudo_EC080529_09	6348069	0.736	99.3	9.353	7228	795.8	1.14	1714	6770
Pseudo_ICBG98	5955722	0.737	97.8	6.475	7749	699.8	1.3	1644	6657
Pseudo_EC080623_01	6785441	0.733	100	5.755	7844	782.4	1.16	1963	6527
Pseudo_CC031210_09	5748747	0.741	99.3	7.914	6866	765	1.19	1788	6356
Pseudo_EC080617_12	6879092	0.732	99.3	7.914	8020	771.9	1.17	2033	6338
Pseudo_CC030327_02	5921325	0.737	98.6	7.914	6909	782.1	1.17	1740	6235

Pseudo_CC011120_04	6368325	0.736	100	8.633	7322	789.1	1.15	1884	6216
Pseudo_EC080624_07	6855742	0.733	99.3	8.633	7955	779.6	1.16	2110	6063
Pseudo_EC080619_08	6766081	0.733	100	8.633	7861	778.8	1.16	2090	5881
Pseudo_CC011120_01	6369072	0.736	99.3	7.914	7480	771.7	1.17	2178	5458
Pseudo_AL030107_17	5716621	0.74	99.3	7.914	7046	740.5	1.23	2005	5344
Pseudo_JS090511_01	6608484	0.739	99.3	3.597	8055	747.7	1.22	2398	5339
Pseudo_EC080618_16	6923749	0.73	100	7.914	8432	737.2	1.22	2468	5242
Pseudo_EC080610_11	6787135	0.73	98.6	7.194	8293	734.1	1.22	2501	4988
Pseudo_alni_DSM44104	5661871	0.739	99.3	6.475	7128	721.1	1.26	2202	4886
Pseudo_nitrificans	5070148	0.729	94.2	8.633	7076	645.4	1.4	2018	4495
Pseudo_ICBG1052	5363352	0.721	92.8	4.317	6225	758.2	1.16	2041	4321
Pseudo_ICBG1126	5668487	0.718	95	3.597	7283	685.7	1.28	2829	4046
Pseudo_MS02	6188306	0.723	95	5.755	8246	670.2	1.33	3092	3359
Pseudo_ICBG1124	5269945	0.722	92.1	6.475	6562	711.1	1.25	2717	3326
Pseudo_LS100414_046	5095383	0.725	95.7	7.914	7757	583	1.52	3767	2352
Pseudo_HH110414_046	5177098	0.727	95.7	10.791	8247	563.4	1.59	4280	2236
Pseudo_LS100414_076	5168255	0.717	93.5	12.95	8499	534	1.64	5646	1754
Pseudo_CC031210_22	5821909	0.725	97.8	13.669	9693	542.6	1.66	8135	1635
Pseudo_chloroethenivorans_JCM12679	3365015	0.688	59.7	5.036	6473	393.2	1.92	3096	1200
Pseudo_tetrahydrofuranoxydans_JCM14745	2134950	0.688	47.5	5.036	4104	373.7	1.92	1947	1151
Pseudo_antarctica336	6289920	0.699	91.4	42.446	14321	371.6	2.28	9669	1116
Pseudo_kongjuensis_394T	5593391	0.725	92.8	15.108	9199	545.6	1.64	21281	747

Supplemental Table 1 (continued)

Strain	trimmed data set	antismash dataset	Consistent Colonizer	Ant Host	Source	Location	Sequencing Platform	Assembly accession number
Pseudo_hierapolitana	yes	yes	Noncolonizer	Non-ant	Soil	NA	PacBio	GCA_007994075.1
Strept_coelicolor	yes	no	NA	Non-ant	Soil	NA	NA	NA
Pseudo_dioxanivorans_CB1190	yes	yes	Noncolonizer	Non-ant	Industrial sludge	NA	454/Illumina with Newbler	GCA_000196675.1
Pseudo_ICBG100	no	yes	Colonizer	ant	Trachymyrmex	USP campus, Brazil	Illumina	NA
Pseudo_JSC141020_01	yes	yes	Colonizer	ant	Mycetophylax asper	Floresta Nacional de Chapecó, Brazil		NA

Pseudo_CC030328_06	yes	yes	Colonizer	ant	Mycetarotes parallelus	Argentina	PacBio	NA
Pseudo_GB151026_04_3	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_EC080529_01	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_sediminis_DSM45779	yes	yes	Noncolonizer	Non-ant	marine sediment	South China Sea	PacBio	GCA_004217185.1
Pseudo_CC151027_05_1	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_GB151021_07_4	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_GB151026_03_3	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_AL050505_11	yes	yes	Colonizer	ant	Acromyrmex octospinosus	NA	PacBio	NA
Pseudo_EV151028_01_1	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_Ae707_Ps1	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa	PacBio	GCA_001932485.1
Pseudo_JSC111027_01	no	yes	Colonizer	ant	Mycetatorotes parallelus	Brazil	NA	NA
Pseudo_GB151021_02_4	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_SP020602_02	yes	yes	Colonizer	ant	Acromyrmex octospinosus	Gamboa, Panama	PacBio	NA
Pseudo_Ae717_Ps2	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	PacBio	GCA_001932475.1
Pseudo_EV151025_05_1	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_EV151025_04_4	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_CC0031209_02	yes	yes	Colonizer	ant	Acromyrmex echinatior		PacBio	NA
Pseudo_HH130630_07	yes	yes	Colonizer	ant	Apterostigma	La Selva, Costa Rica	PacBio	GCA_001698125.1
Pseudo_TRS120623_01	no	yes	Colonizer	ant	Acromyrmex hispidus fallax	Argentina	NA	NA
Pseudo_EV151025_09_4	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_EV151025_09_1	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA
Pseudo_Ae356_Ps1	yes	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	Illumina	GCA_001932395.1
Pseudo_Ae263Ps1_SC1	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	Illumina	GCA_001932355.1
Pseudo_EC080619_01	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	GCA_001420995.1
Pseudo_EC080529_05	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	NA

Pseudo_AL041005_10	no	yes	Colonizer	ant	Trachymyrmex cornetzi	Peru	PacBio	GCA_001294605.1
Pseudo_EC080610_09	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	GCA_001420975.1
Pseudo_EC080625_04	no	yes	Colonizer	ant	Apterostigma dentigerum	Panama	PacBio	GCA_001294425.1
Pseudo_Ae150A_Ps1	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	Illumina	GCA_001932315.1
Pseudo_Ae706_Ps2	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	Illumina	GCA_001932325.1
Pseudo_UGM030327_02	yes	yes	Colonizer	ant	Acromyrmex hispidus fallax	Peru	PacBio	NA
Pseudo_Ae406_Ps2	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	Illumina	GCA_001932415.1
Pseudo_Ae331Ps2_SC1	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	Illumina	GCA_001932405.1
Pseudo_ICBG1034	no	yes	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_HH130629_09	yes	yes	Colonizer	ant	Apterostigma	NA	PacBio	GCA_001294645.1
Pseudo_Ae505_Ps2	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa, Panama	Illumina	GCA_001932425.1
Pseudo_AL041002_03	no	yes	Colonizer	ant	Trachymyrmex zeteki	Panama	PacBio	NA
Pseudo_UGM030330_05	no	yes	Colonizer	ant	Acromyrmex laticeps	NA	PacBio	NA
Pseudo_alni_PB	yes	yes	Colonizer	Non-ant	root nodule of alder tree	NA	PacBio	GCA_002813375.1
Pseudo_CTL110912_03	yes	yes	Colonizer	ant	NA	Brazil	NA	NA
Pseudo_Ae168_Ps1	no	yes	Colonizer	ant	Acromyrmex echinatior	Gamboa	Illumina	GCA_001932335.1
Gordonia_SID5947	yes	no	NA	Non-ant	NA	NA	NA	NA
Pseudo_ST040116_010	no	yes	Colonizer	ant	Acromyrmex	Panama	PacBio	NA
Pseudo_endophytica	yes	yes	Noncolonizer	Non-ant	Lobelia clavata (plant)	Xishuangbanna, Yunnan Province, China	PacBio	GCA_004339565.1
Pseudo_cypriaca	yes	yes	Noncolonizer	Non-ant	Agricultural estate	Cyprus	PacBio	GCA_006717045.1
Pseudo_kunmingensis_DSM45301	yes	yes	Noncolonizer	Non-ant	roots of the plant Artemisia annua	China, Yunnan Province, Kunming	PacBio	GCA_006716445.1
Pseudo_acaciae_DSM45401	yes	yes	Noncolonizer	Non-ant	Acacia auriculiformis roots	Thailand	Illumina	GCA_000620785.1
Pseudo_spinosispora_DSM44797	yes	yes	Noncolonizer	Non-ant	soil	Korea	Illumina	GCF_000429025.1
Pseudo_autotrophica_DSM43083	no	yes	Colonizer	Non-ant	na	NA	Illumina	GCA_001902615.1
Pseudo_15845	yes	yes	Noncolonizer	Non-ant	Orthoptera	New Mexico	Illumina	NA
Pseudo_autotrophica_NRRLB16064	no	yes	Noncolonizer	Non-ant	na	NA	Illumina	GCA_000717175.1

Pseudo_MH_G8	yes	yes	Noncolonizer	Non-ant	Monostroma hariotii	Antarctica: King George Island	Illumina	GCA_002262885.1
Pseudo_oroxyli_CGMCC	yes	yes	Noncolonizer	Non-ant	Oroxylum indicum root	China	NA	GCA_900102195.1
Pseudo_thermophila_DSM43832	yes	yes	Noncolonizer	Non-ant	fresh horse manure	NA	NA	GCA_900142365.1
Pseudo_N23	yes	yes	Noncolonizer	Non-ant	groundwater	Japan	Illumina	GCA_002583555.1
Pseudo_asaccharolytica_DSM44247	yes	yes	Noncolonizer	Non-ant	tree bark compost	NA	Illumina	GCA_000423625.1
Pseudo_hydroxycarbonoxydans	yes	yes	Noncolonizer	Non-ant	Air contaminant	NA	Illumina	NA
Pseudo_ammonioxydans_CGMCC	yes	yes	Colonizer	Non-ant	coastal sediment	China	NA	GCA_900115005.1
Pseudo_CC030328_06ill	no	no	Colonizer	ant	Mycetarotes parallelus	Argentina	Illumina	NA
Pseudo_SCN73_27	yes	yes	Noncolonizer	Non-ant	thiocyanate bioreactor	South Africa: University of Cape Town, Rondebosch	Illumina	GCA_001725415.1
Pseudo_autotrophica_DSM535	yes	yes	Colonizer	Non-ant	lab phosphate buffer	NA	Illumina	GCA_002119215.1
Pseudo_alaniphila	yes	yes	Noncolonizer	Non-ant	primval forest soil	China	Illumina	NA
Pseudo_73_21_SCN	yes	yes	Noncolonizer	Non-ant	Ammonium sulfate bioreactor	South Africa: University of Cape Town	Illumina	GCA_001899645.1
Pseudo_CNS_004	yes	yes	Noncolonizer	Non-ant	marine sediment	Palau	IonTorrent	GCA_001942185.1
Pseudo_AL050513_04	yes	no	Colonizer	ant	Trachymyrmex	Panama	NA	NA
Pseudo_sulfidoxydans_NBRC16205	yes	no	Noncolonizer	Non-ant	biofiltermaterial/tree bark compost	NA	Illumina HiSeq	GCA_007989085.1
Pseudo_AL041002_03ill	no	no	Colonizer	ant	Trachymyrmex zeteki	Panama	Illumina	NA
Pseudo_CNS_139	yes	no	Noncolonizer	Non-ant	marine sediment	Palau	IonTorrent	GCA_001942415.1
Pseudo_hierapolitanaill	no	no	Noncolonizer	Non-ant	Thermal occurrence	Pamukkale, Turkey	Illumina	NA
Pseudo_ICBG618	yes	no	Colonizer	ant	Acromyrmex	Itatiaia, Rio de Janeiro, Brazil	Illumina	NA
Pseudo_spH69	yes	no	Noncolonizer	Non-ant	Atacama Desert soil	ALMA Observatory site, Atacama Desert, Chile	Illumina	NA

Pseudo_colA	yes	no	Colonizer	ant	Acromyrmex	NA	Illumina	NA
Pseudo_compacta	yes	no	Noncolonizer	Non-ant	garden soil	Wohra, Henssen, Germany	NA	NA
Pseudo_saturnea_NRRLB16172	yes	no	Colonizer	Non-ant	aerosol	NA	Illumina	GCA_006539585.1
Pseudo_10385	yes	no	Colonizer	Non-ant	Caterpillar	Biocore Prarie, Wisconsin, USA	Illumina	NA
Pseudo_ICBG157	yes	no	Colonizer	ant	Acromyrmex	USP campus, Brazil	Illumina	NA
Pseudo_10165	yes	no	Colonizer	Non-ant	Arachnida	Minnesota, USA	Illumina	NA
Pseudo_endophyticaill	no	no	Noncolonizer	Non-ant	Lobelia clavata (plant)	Xishuangbanna, Yunnan Province	Illumina	NA
Pseudo_SCN72_86	no	no	Noncolonizer	Non-ant	Cyanide and thiocyanate bioreactor	South Africa: University of Cape Town, Rondebosch	Illumina	GCA_001724645.1
Pseudo_colD	yes	no	Colonizer	ant	Acromyrmex volcanus	NA	Illumina	NA
Pseudo_chloro3	yes	no	Noncolonizer	Non-ant	lab enrichment from soil	USA	Illumina	NA
Pseudo_kujensis	yes	no	Noncolonizer	Non-ant	soil	Nigeria	Illumina	NA
Pseudo_chloroethenivorans	no	no	Noncolonizer	Non-ant	lab enrichment from soil	USA	NA	NA
Pseudo_EC080618_06	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_SCN72_51	no	no	Noncolonizer	Non-ant	Cyanide and thiocyanate bioreactor	South Africa: University of Cape Town, Rondebosch	Illumina	GCA_001725125.1
Pseudo_zijingensis	yes	no	Noncolonizer	Non-ant	soil	China	Illumina	NA
Pseudo_EC080610_09ill	no	no	Colonizer	ant	Apterostigma dentigerum	Panama	Illumina	NA
Pseudo_ICBG1111	no	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_ICBG1043	no	no	Colonizer	ant	Acromyrmex	Itatiaia, Rio de Janeiro, Brazil	Illumina	NA
Pseudo_ICBG101	no	no	Colonizer	ant	Trachymyrmex	USP campus , Brazil	Illumina	NA

Pseudo_EC080619_09	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_P1	no	no	Colonizer	ant	Acromyrmex octospinosus	NA	NA	GCA_000178675.1
Pseudo_cypriacaill	no	no	Noncolonizer	Non-ant	Agricultural estate	Cyprus	Illumina	NA
Pseudo_halophobica	yes	no	Noncolonizer	Non-ant	soil	NA	NA	NA
Pseudo_EC090830_01	no	no	Colonizer	ant	Trachymyrmex	Panama	NA	NA
Pseudo_EC080524_04	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_ICBG102	no	no	Colonizer	ant	Trachymyrmex	USP campus, Sao Paulo, Panama	Illumina	NA
Pseudo_ICBG1146	yes	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_EC080618_05	no	no	Colonizer	ant	Apterostigma	Panama	NA	NA
Pseudo_EC080620_06	no	no	Colonizer	ant	Apterostigma dentigerum	Buena Vista Peninsula, Panama		NA
Pseudo_colB	yes	no	Colonizer	ant	Acromyrmex	NA	Illumina	NA
Pseudo_ICBG158	no	no	Colonizer	ant	Trachymyrmex	USP campus, Brazil	Illumina	NA
Pseudo_EC080529_20	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_EC080524_13	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_ICBG1050	no	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_ICBG1042	yes	no	Colonizer	ant	Acromyrmex	Riberao Preto, Campus, Brazil	Illumina	NA
Pseudo_petroleophila	yes	no	Noncolonizer	Non-ant	soil	NA	Illumina	NA
Pseudo_EC080617_07	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_EC080603_07	no	no	Colonizer	ant	Apterostigma dentigerum	near Fortuna, Panama	NA	NA
Pseudo_SID8383	no	no	Colonizer	ant	Acromyrmex echinatior	NA	NA	NA
Pseudo_EC080529_16	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_EC080529_15	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA

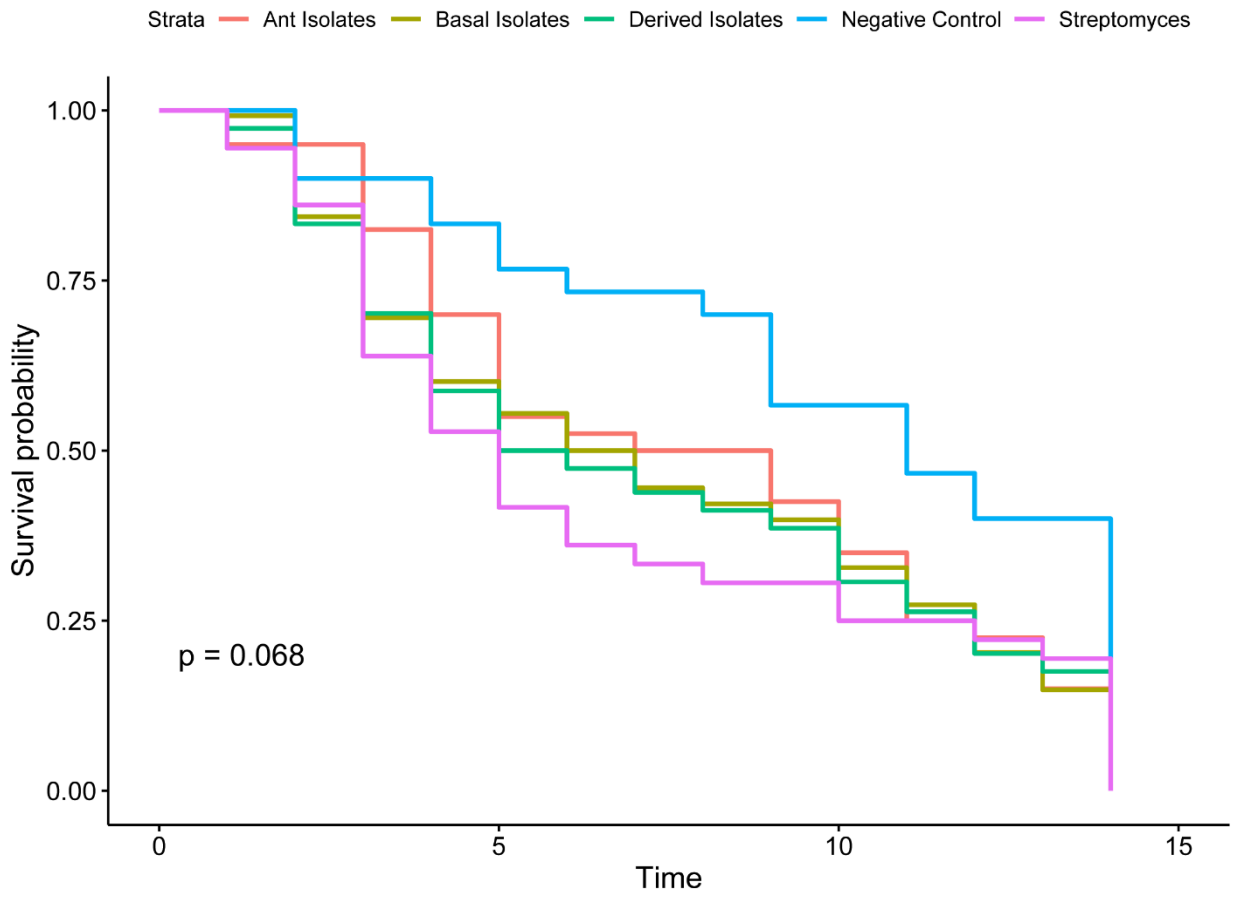
Pseudo_EC070717_09	no	no	Colonizer	ant	Apterostigma dentigerum	La Selva, Camino Experimental Ser 180m, Costa Rica	NA	NA
Pseudo_EC080524_14	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_EC080623_03	no	no	Colonizer	ant	Apterostigma dentigerum	Gamboa Forest, Panama	NA	NA
Pseudo_EC080525_06	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_UGM030402_04	yes	no	Colonizer	ant	Acromyrmex hisidus falis	Argentina	NA	NA
Pseudo_CC060123_03	no	no	Colonizer	ant	Apterostigma dentigerum	La Selva, Lindero Occidental 1300m, Costa Rica	NA	NA
Pseudo_P2	yes	no	Colonizer	ant	Acromyrmex octospinosus	NA	NA	GCA_000179835.2
Pseudo_EC080617_04	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_EC061022_05	yes	no	Colonizer	ant	Atta cephalotes	La Selva, Camino Experimental Ser 200m, Costa Rica	NA	NA
Pseudo_EC080524_01	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_SP030328_02	no	no	Colonizer	ant	Atta sexdens	Argentina	NA	NA
Pseudo_EC080625_04ill	no	no	Colonizer	ant	Apterostigma dentigerum	Panama	Illumina	NA
Pseudo_EC080525_05	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_EC070720_06	no	no	Colonizer	ant	Apterostigma dentigerum	La Selva, Sendero Oriental 550m, Costa Rica	NA	NA
Pseudo_UGM030330_04	yes	no	Colonizer	ant	Acromyrmex laticeps	Argentina	NA	NA
Strept_10815	yes	no	NA	Non-ant	Bee species unidentified	NA	NA	NA
Pseudo_ICBG1142	yes	no	Colonizer	ant	Acromyrmex	Itatiaia, Rio de Janeiro, Brazil	Illumina	NA
Pseudo_ICBG602	no	no	Colonizer	ant	Apterostigma	Amazonas, Anavilhanas, Brazil	Illumina	NA

Pseudo_ICBG1143	no	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_EC080620_01	no	no	Colonizer	ant	Apterostigma dentigerum	Buena Vista Peninsula, Panama	NA	NA
Pseudo_EC080617_15	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_EC080618_04	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_EC090828_04	no	no	Colonizer	ant	Trachymyrmex	Peru	NA	NA
Pseudo_ICBG1145	no	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_ICBG1125	no	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_EC080529_19	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_ICBG103	yes	no	Colonizer	ant	Acromyrmex	USP campus, Brazil	PacBio	NA
Pseudo_SP030327_02	yes	no	Colonizer	ant	Apterostigma	Argentina	NA	NA
Pseudo_SP020602_02ill	no	no	Colonizer	ant	Acromyrmex octospinosus	Gamboa, Panama	Illumina	NA
Pseudo_ICBG93	no	no	Colonizer	ant	Unknown Attine	USP campus, Sao Paulo, Brazil	Illumina	NA
Pseudo_EC060123_09	no	no	Colonizer	ant	Apterostigma dentigerum	La Selva, Lindero Occidental 2200m, Costa Rica	NA	NA
Pseudo_ICBG161	no	no	Colonizer	ant	Trachymyrmex	USP campus, Brazil	Illumina	NA
Pseudo_AL040118_01	no	no	Colonizer	ant	Acromyrmex	Resort Road	NA	NA
Pseudo_CC020602_01	no	no	Colonizer	ant	Acromyrmex octospinosus	Gamboa; Panama	NA	NA
Pseudo_EC080624_04	yes	no	Colonizer	ant	Apterostigma dentigerum	Ridge, Panama	NA	NA
Pseudo_EC080620_04	no	no	Colonizer	ant	Apterostigma dentigerum	Buena Vista Peninsula, Panama	NA	NA
Pseudo_ICBG1144	no	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA

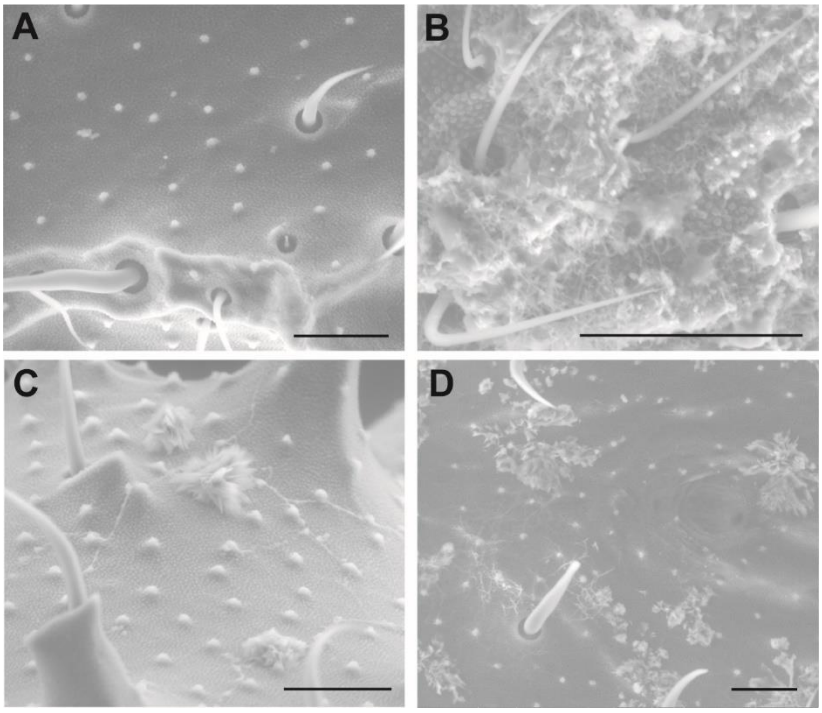
Pseudo_ICBG162	no	no	Colonizer	ant	Trachymyrmex	USP campus, Brazil	Illumina	NA
Pseudo_EC080618_12	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_CC031212_01	no	no	Colonizer	ant	Acromyrmex echinatior	Panama	NA	NA
Pseudo_EC080618_17	yes	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_EC070717_12	no	no	Colonizer	ant	Apterostigma dentigerum	La Selva, Camino Experimental Ser 700m, Costa Rica	NA	NA
Pseudo_EC080525_24	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_EC080529_09	no	no	Colonizer	ant	Apterostigma dentigerum	Pipeline Road, Panama	NA	NA
Pseudo_ICBG98	no	no	Colonizer	ant	Trachymyrmex	USP campus, Sao Paulo, Brazil	Illumina	NA
Pseudo_EC080623_01	yes	no	Colonizer	ant	Apterostigma dentigerum	Gamboa Forest, Panama	NA	NA
Pseudo_CC031210_09	yes	no	Colonizer	ant	Cyphomyrmex costatus	NA	NA	NA
Pseudo_EC080617_12	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_CC030327_02	no	no	Colonizer	ant	Acromyrmex niger	Argentina	NA	NA
Pseudo_CC011120_04	no	no	Colonizer	ant	Apterostigma dentigerum	Panama	NA	NA
Pseudo_EC080624_07	no	no	Colonizer	ant	Apterostigma dentigerum	Ridge, Panama	NA	NA
Pseudo_EC080619_08	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_CC011120_01	no	no	Colonizer	ant	Apterostigma auriculatum	Gamboa, Panama	NA	NA
Pseudo_AL030107_17	no	no	Colonizer	ant	*Trachymyrmex	Panama	NA	NA
Pseudo_JS090511_01	yes	no	Colonizer	ant	Atta cephalotes	NA	NA	NA
Pseudo_EC080618_16	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_EC080610_11	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA

Pseudo_alni_DSM44104	no	no	Colonizer	Non-ant	root nodule of alder tree	NA	NA	NA
Pseudo_nitrificans	yes	no	Colonizer	Non-ant	soil	NA	Illumina	NA
Pseudo_ICBG1052	yes	no	Colonizer	ant	Cyphomyrmex	Itatiaia, Rio de Janeiro, Brazil	Illumina	NA
Pseudo_ICBG1126	no	no	Colonizer	ant	Acromyrmex	Itatiaia, Rio de Janeiro, Brazil	Illumina	NA
Pseudo_MS02	no	no	Colonizer	ant	Apterostigma dentigerum	Barro Colorado Island, Panama	NA	NA
Pseudo_ICBG1124	no	no	Colonizer	ant	Trachymyrmex	Amazonas, Anavilhanas, Brazil	Illumina	NA
Pseudo_LS100414_046	no	no	Colonizer	ant	Trachymyrmex septentrionalis	Archbold Biological Station; Venus, FL; USA	NA	NA
Pseudo_HH110414_046	yes	no	Colonizer	ant	Trachymyrmex septentrionalis	Archbold Biological Station; Venus, FL; USA	NA	NA
Pseudo_LS100414_076	yes	no	Colonizer	ant	Trachymyrmex septentrionalis	Archbold Biological Station; Venus, FL; USA	NA	NA
Pseudo_CC031210_22	yes	no	Colonizer	ant	Acromyrmex	Peru	NA	NA
Pseudo_chloroethenivorans_JCM12679	no	no	Noncolonizer	Non-ant	lab enrichment from soil	USA	NA	NA
Pseudo_tetrahydrofuranoxydans_JCM14745	no	no	Noncolonizer	Non-ant	wastewater	Germany	NA	GCA_001313405.1
Pseudo_antarctica336	yes	no	Colonizer	Non-ant	soil	Antarctica	NA	NA
Pseudo_kongjuensis_394T	yes	no	Colonizer	Non-ant	Gold mine cave soil	Kongju, Republic of Korea	Illumina	NA

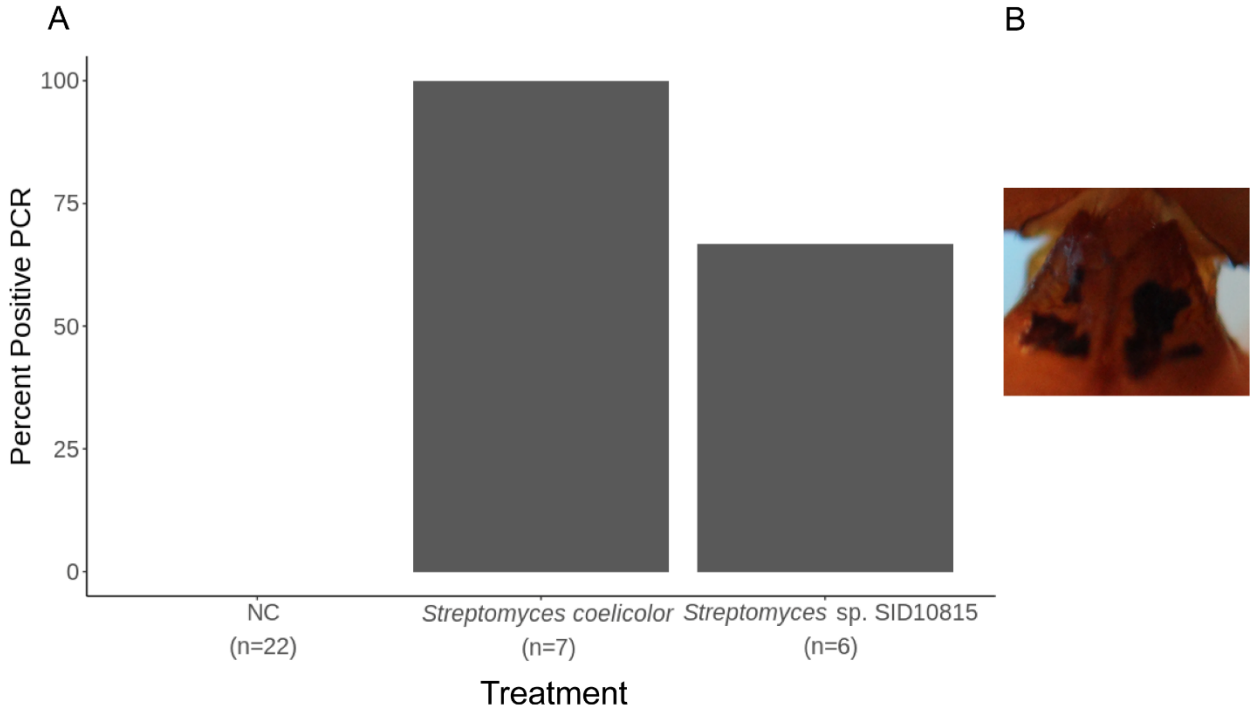
Supplemental Table 1. All strains used in this study with metadata including isolation source, estimated percent completeness and redundancy, estimated number of genes, sequencing technology used if available.



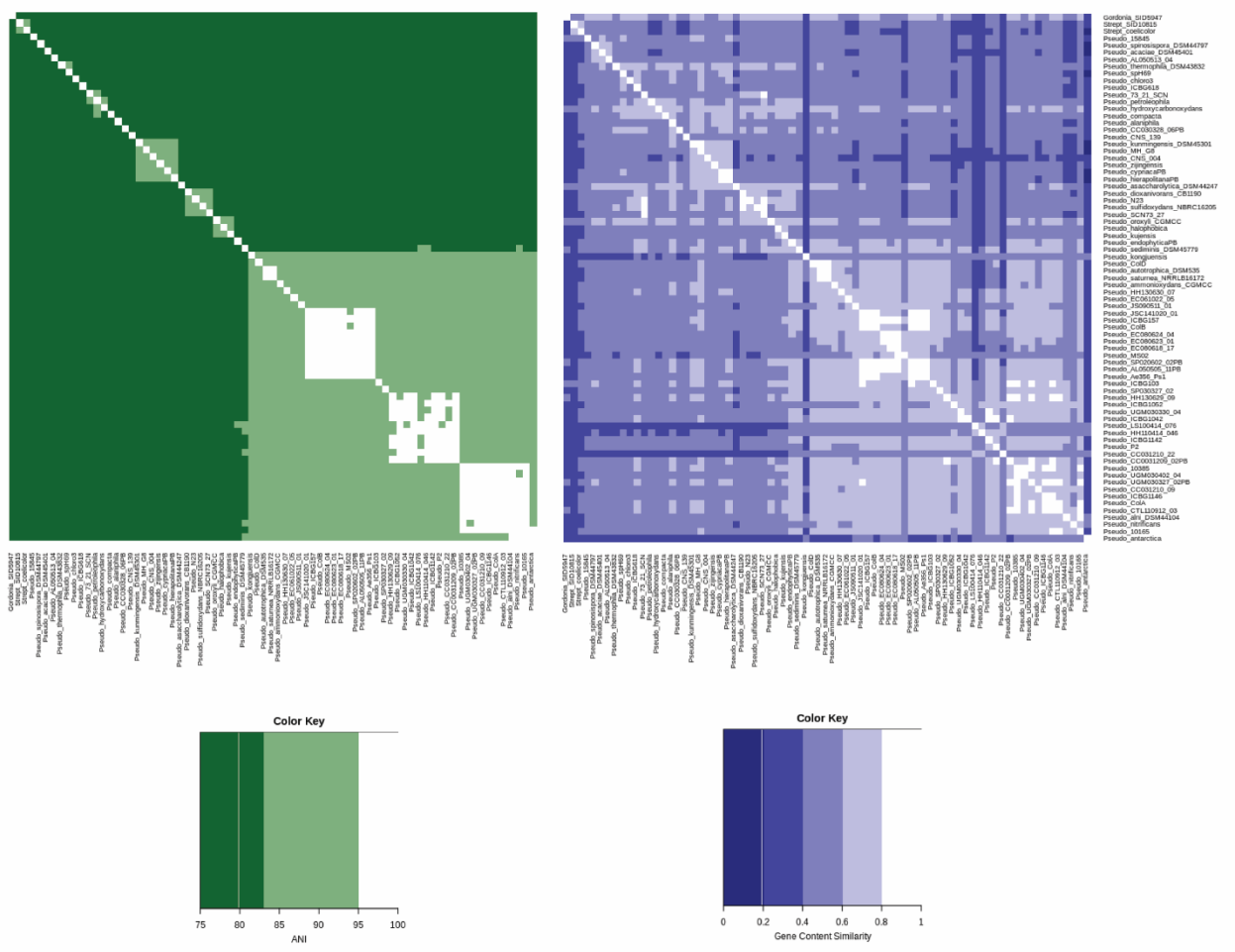
Supplemental Figure 1. Ant survival probability grouped by treatment groups. Time survived measured in days. Ant isolates (red) include *Pseudonocardia* sp. AL050505-11, *Pseudonocardia* CC031209_02; basal isolates (dark yellow) include *P. spinosispora, P. chlorothenenvirons, P. petroleophila, P. compacta, P. alaniphila, P. zijingensis, P. cypriaca;* derived isolates (green) include *P. alni, P. antarctica, P. saturnea, P. kujensis, P. endophytica,* and *P. nitrificans*; Streptomyces (pink) include *S. coelicolor* and *Streptomyces* sp. SID10815.



Supplemental Figure 2. eSEM of worker ants heads. Black scale bar represents 50 μm. A. Negative control. B. Ant naturally reared with native symbiont. C. Ant treated with *P. spinosispora*. D. Ant treated with *P. alni*.



Supplemental Figure 3. A. Percentage of ants colonized with *Streptomyces* strains and negative controls. B. Propleural plate of ant treated with *Streptomyces coelicolor*.



Supplemental Figure 4. Overall genome similarity for all strains in trimmed dataset. A. ANI of strains B. Percentage of shared gene content.

Homolog	EggNOG description	EggNOG COG category	PFAM annotation	FET Consistent vs Inconsistent Odds Ratio	FET pvalue bonferroni corrected	FET consistent v inconsistentEnrichmentCategory	TreeWAS consistent v inconsistent	FET Odds RatioantVSother	FET pvalue bonferroni AntVSother	FET enrichment category	TreeWAS antVother
group_00008	HNH nucleases	V	DUF222	2.409	0.032	Consistent	NA	NS	NS	NA	NA
group_00017	Transposase DDE domain	L	DDE_Tnp_1_2	6.231	0	Consistent	NA	0.179	0	Ant	NA

group_00020	PFAM Transposase, mutator type	L	NA	NA	NS	NA	NA	0.345	0.002	Ant	NA
group_00023	Integrase core domain	L	rve,HTH_21	3.24	0.003	Consistent	NA	0.289	0	Ant	NA
group_00025	Major facilitator Superfamily	EGP	MFS_1	0.287	0	Inconsistent	NA	NS	NS	NA	NA
group_00047	Belongs to the sigma-70 factor family	K	Sigma70_r2,Sigma70_r4_2	0.192	0	Inconsistent	NA	3.858	0	Other	NA
group_00048	Transposase	L	DUF4096,DDE_Tnp_1	3.342	0.026	Consistent	NA	0.327	0.003	Ant	NA
group_00054	luxR family	K	AAA_16,GerE	0.268	0	Inconsistent	NA	3.599	0.001	Other	NA
group_00083	hydrolases or acyltransferases (alpha beta hydrolase superfamily)	I	Abhydrolase_6	0.257	0	Inconsistent	NA	NS	NS	NA	NA
group_00106	transcriptional activator domain	K	BTAD,TPR_MalT,Trans_reg_C,NB-ARC	0.304	0.009	Inconsistent	NA	NS	NS	NA	NA
group_00180	Rieske [2Fe-2S] domain	P	Rieske	0.187	0	Inconsistent	NA	3.77	0.021	Other	NA
group_00186	Protein of unknown function (DUF3558)	S	DUF3558	62.062	0	Consistent	NA	0.25	0	Ant	NA
group_00458	PFAM transposase IS4 family protein	S	DDE_Tnp_4	7.049	0	Consistent	NA	0.14	0	Ant	NA
group_00478	AAA ATPase domain	K	AAA_16,GerE	0.284	0.031	Inconsistent	NA	NS	NS	NA	NA
group_01415	Transcriptional regulator containing an amidase domain and an AraC-type DNA-binding HTH domain	K	DJ-1_PfpI,HTH_18	0.266	0.046	Inconsistent	NA	NS	NS	NA	NA
group_01443	NA	NA	PE	17.446	0	Consistent	NA	NS	NS	NA	NA
group_01532	Belongs to the major facilitator superfamily. Sugar transporter (TC 2.A.1.1) family	U	Sugar_tr	4.818	0.011	Consistent	NA	0.246	0.003	Ant	NA
group_01543	TOBE domain	P	ABC_tran,TOBE_2	0.209	0	Inconsistent	NA	NS	NS	NA	NA
group_01793	Transposase	L	HTH_Tnp_1	14.953	0	Consistent	NA	0.054	0	Ant	NA
group_02091	Putative transposase of IS4/5 family (DUF4096)	L	DUF4096	inf	0	Consistent	NA	0.014	0	Ant	NA
group_02663	Belongs to the binding-protein-dependent transport system permease family	P	BPD_transp_2,ABC_tran,BCA_ABC_TP_C,BPD_transp_2	0.182	0.003	Inconsistent	NA	NS	NS	NA	NA
group_02726	reversible hydration of carbon dioxide	P	Pro_CA	7.327	0.003	Consistent	NA	NS	NS	NA	NA
group_02861	PFAM beta-lactamase domain protein	P	NA	NA	NS	NA	NA	5.369	0.042	Other	NA
group_02998	AsnC-type helix-turn-helix domain	K	AsnC_trans_reg,HTH_AsnC-type,HTH_AsnC-	0.197	0.014	Inconsistent	NA	NS	NS	NA	NA

			type,AsnC_trans_reg								
group_03026	NA	NA	NA	0.208	0.044	Inconsistent	NA	NS	NS	NA	NA
group_03067	protein phosphatase 2C domain protein	T	SpoIIE,PAS_4,GAF,HATPase_c_2,PAS_9	0.03	0	Inconsistent	NA	37.587	0	Other	NA
group_03083	CAAX protease self-immunity	S	CPBP	18.691	0	Consistent	NA	0.214	0.024	Ant	NA
group_03114	NA	NA	NA	19.065	0	Consistent	NA	NS	NS	NA	NA
group_03137	Sigma-70 region 2	K	Sigma70_r4_2,Sigma70_r2	0.153	0.001	Inconsistent	NA	NS	NS	NA	NA
group_03167	Bacterial regulatory proteins, tetR family	K	TetR_N	6.878	0.016	Consistent	NA	NS	NS	NA	NA
group_03278	NA	NA	NA	6.579	0.026	Consistent	NA	NS	NS	NA	NA
group_03329	NA	NA	NA	6.579	0.026	Consistent	NA	NS	NS	NA	NA
group_03398	Methyltransferase domain	I	Methyltransf_25	8.224	0.007	Consistent	NA	NS	NS	NA	NA
group_03401	Acyl-CoA dehydrogenase, C-terminal domain	I	Acyl-CoA_dh_1,Acyl-CoA_dh_M,Acyl-CoA_dh_N	7.85	0.019	Consistent	NA	NS	NS	NA	NA
group_03410	cyclic nucleotide binding	T	cNMP_binding,HTH_Crp_2	8.224	0.007	Consistent	NA	NS	NS	NA	NA
group_03412	SCP-2 sterol transfer family	S	NA	NA	NS	NA	NA	0.194	0.024	Ant	NA
group_03453	Transcriptional regulator	K	AraC_binding_2,HTH_18	0.192	0.026	Inconsistent	NA	NS	NS	NA	NA
group_03457	Dehydrogenase	E	ADH_N,ADH_zinc_N	10.467	0.005	Consistent	NA	NS	NS	NA	NA
group_03464	Amidinotransferase	E	Amidinotransf	0.085	0	Inconsistent	NA	10.663	0.003	Other	NA
group_03490	PFAM peptidase S9 prolyl oligopeptidase active site domain protein	E	Peptidase_S9	7.85	0.019	Consistent	NA	0.194	0.024	Ant	NA
group_03490	Dipeptidyl peptidase IV (DPP IV) N-terminal region	E	NA	NA	NS	NA	NA	0.183	0.042	Ant	NA
group_03504	Belongs to the 'phage' integrase family	L	Phage_integrase,Phage_int_SAM_1	16.074	0	Consistent	NA	NS	NS	NA	NA
group_03550	Flavin containing amine oxidoreductase	E	Amino_oxidase	31.4	0	Consistent	NA	NS	NS	NA	NA
group_03555	Putative sugar-binding domain	K	Sugar-bind,HTH_Crp_2	7.663	0.032	Consistent	NA	NS	NS	NA	NA
group_03641	Prolyl oligopeptidase, N-terminal beta-propeller domain	E	Peptidase_S9,Peptidase_S9_N	15.7	0	Consistent	NA	NS	NS	NA	NA
group_03650	Polyketide cyclase / dehydrase and lipid transport	S	Polyketide_cyc2	10.217	0.008	Consistent	NA	NS	NS	NA	NA
group_03652	Sulfate permease family	P	Sulfate_transp,STAS	0.085	0	Inconsistent	NA	NS	NS	NA	NA

group_03662	Binding-protein-dependent transport system inner membrane component	G	BPD_transp_1	0.14	0.006	Inconsistent	NA	NS	NS	NA	NA
group_03672	SCP-2 sterol transfer family	I	SCP2	7.663	0.032	Consistent	NA	NS	NS	NA	NA
group_03675	Acetyltransferase (GNAT) domain	J	Acetyltransf_3	7.663	0.032	Consistent	NA	NS	NS	NA	NA
group_03680	Periplasmic binding protein	P	Peripla_BP_2	16.074	0	Consistent	NA	NS	NS	NA	NA
group_03681	phosphotransferase system	G	PTS_EIIB	9.968	0.008	Consistent	NA	NS	NS	NA	NA
group_03728	PFAM transposase IS3 IS911 family protein	L	HTH_Tnp_1	15.7	0	Consistent	NA	0.178	0.024	Ant	NA
group_03745	aldo keto reductase	C	Aldo_ket_red	0.145	0.012	Inconsistent	NA	NS	NS	NA	NA
group_03770	sequence-specific DNA binding	S	HTH_31	9.968	0.008	Consistent	NA	NS	NS	NA	NA
group_03773	PFAM AIG2 family protein	S	GGACT	15.7	0	Consistent	NA	NS	NS	NA	NA
group_03778	Polyketide cyclase / dehydrase and lipid transport	E	Polyketide_cyc2	10.217	0.008	Consistent	NA	NS	NS	NA	NA
group_03783	NA	NA	NA	10.217	0.008	Consistent	NA	NS	NS	NA	NA
group_03851	Cold shock	K	CSD	15.326	0.001	Consistent	NA	NS	NS	NA	NA
group_03863	Drug exporters of the RND superfamily	F	MMPL,MMPL	0.068	0.023	Inconsistent	NA	14.794	0.003	Other	NA
group_03875	ATPase involved in chromosome	D	CbiA	9.47	0.022	Consistent	NA	NS	NS	NA	NA
group_03878	Methyltransferase domain	S	Methyltransf_11	15.326	0.001	Consistent	NA	NS	NS	NA	NA
group_03883	glyoxalase bleomycin resistance protein dioxygenase	S	Glyoxalase_6	15.326	0.001	Consistent	NA	NS	NS	NA	NA
group_03891	NA	NA	NA	15.326	0.001	Consistent	NA	NS	NS	NA	NA
group_03892	NA	NA	NA	9.968	0.008	Consistent	NA	NS	NS	NA	NA
group_03893	Dodecin	S	Dodecin	15.326	0.001	Consistent	NA	NS	NS	NA	NA
group_03895	NA	NA	NA	31.4	0	Consistent	NA	NS	NS	NA	NA
group_03896	NA	NA	NA	9.968	0.008	Consistent	NA	NS	NS	NA	NA
group_03897	NA	NA	NA	15.326	0.001	Consistent	NA	NS	NS	NA	NA
group_03912	transcriptional regulator	K	MarR_2	0.088	0	Inconsistent	NA	NS	NS	NA	NA
group_03914	membrane	S	UPF0126,UPF0126	0.129	0.017	Inconsistent	NA	NS	NS	NA	NA
group_03963	Citrate transporter	C	CitMHS	14.204	0.004	Consistent	NA	0.16	0.011	Ant	2
group_03966	Protein of unknown function (DUF3445)	S	DUF3445	9.22	0.038	Consistent	NA	NS	NS	NA	NA
group_03993	NA	NA	NA	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_03995	PFAM Methyltransferase domain	MQ	Methyltransf_25	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_03997	helix_turn_helix, mercury resistance	K	MerR_1	9.719	0.013	Consistent	NA	NS	NS	NA	NA
group_03999	ErfK ybiS ycfS ynhG family protein	D	YkuD	14.952	0.001	Consistent	NA	NS	NS	NA	NA

group_04000	DSBA-like thioredoxin domain	O	Thioredoxin_4	29.157	0	Consistent	NA	NS	NS	NA	NA
group_04001	PFAM Methyltransferase domain	MQ	Methyltransf_11	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_04002	NA	NA	NA	9.719	0.013	Consistent	NA	NS	NS	NA	NA
group_04003	Transglycosylase-like domain	S	Transglycosylas	9.22	0.038	Consistent	NA	NS	NS	NA	NA
group_04007	NA	NA	NA	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_04008	NA	NA	NA	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_04009	Protein of unknown function (DUF4232)	S	NA	9.719	0.013	Consistent	NA	NS	NS	NA	NA
group_04010	TIGRFAM PTS system, glucose subfamily, IIA	G	PTS_EIIA_1	9.22	0.038	Consistent	NA	NS	NS	NA	NA
group_04012	NA	NA	NA	30.652	0	Consistent	NA	NS	NS	NA	NA
group_04013	NA	NA	NA	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_04014	NA	NA	NA	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_04015	NA	NA	NA	30.652	0	Consistent	NA	NS	NS	NA	NA
group_04029	N-acetylneuraminate synthase	M	NeuB,SAF	9.22	0.038	Consistent	NA	NS	NS	NA	NA
group_04036	NA	NA	NA	9.719	0.013	Consistent	NA	NS	NS	NA	NA
group_04042	NA	NA	NA	14.952	0.001	Consistent	NA	NS	NS	NA	NA
group_04048	Universal stress protein family	T	Usp	0.121	0.005	Inconsistent	NA	NS	NS	NA	NA
group_04050	Glycosyl hydrolases family 16	G	Glyco_hydro_16	14.204	0.004	Consistent	NA	NS	NS	NA	NA
group_04068	ABC-2 family transporter protein	CP	ABC2_membrane_3	inf	0	Consistent	NA	NS	NS	NA	NA
group_04079	NADH dehydrogenase	C	Pyr_redox_2	0.1	0.002	Inconsistent	NA	NS	NS	NA	NA
group_04102	PFAM glycoside hydrolase, family 10	EG	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04105	Component of the proteasome core, a large protease complex with broad specificity involved in protein degradation	O	Proteasome	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04106	membrane-bound metal-dependent hydrolase	S	YdjM	inf	0	Consistent	NA	NS	NS	NA	NA
group_04109	Belongs to the anti-sigma-factor antagonist family	T	STAS_2	29.904	0	Consistent	NA	NS	NS	NA	NA
group_04112	NA	NA	NA	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04113	Single-stranded DNA-binding protein	L	SSB	29.904	0	Consistent	NA	NS	NS	NA	NA
group_04115	NA	NA	NA	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04117	NA	NA	NA	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04118	NA	NA	NA	14.578	0.002	Consistent	NA	NS	NS	NA	NA

group_04119	NA	NA	NA	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04126	Dihydropyrimidinase	F	Amidohydro_1	0.07	0	Inconsistent	NA	NS	NS	NA	NA
group_04134	Belongs to the binding-protein-dependent transport system permease family	G	BPD_transp_2	0.066	0	Inconsistent	NA	NS	NS	NA	NA
group_04147	Ubiquinone biosynthesis O-methyltransferase	H	Methyltransf_23	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04153	NA	NA	NA	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04154	NA	NA	NA	14.578	0.002	Consistent	NA	NS	NS	NA	NA
group_04186	GDSL-like Lipase/Acylhydrolase family	E	Lipase_GDSL_2	inf	0	Consistent	NA	NS	NS	NA	NA
group_04187	Spore coat polysaccharide biosynthesis protein F, CMP-KDO synthetase	M	CTP_transf_3	9.22	0.038	Consistent	NA	NS	NS	NA	NA
group_04190	Acetyltransferase (GNAT) domain	K	Acetyltransf_1	13.831	0.008	Consistent	NA	NS	NS	NA	NA
group_04192	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04193	NA	NA	NA	29.157	0	Consistent	NA	NS	NS	NA	NA
group_04194	NA	NA	NA	14.204	0.004	Consistent	NA	NS	NS	NA	NA
group_04195	NA	NA	NA	14.204	0.004	Consistent	NA	NS	NS	NA	NA
group_04196	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04197	NA	NA	NA	14.204	0.004	Consistent	NA	NS	NS	NA	NA
group_04213	transporter mgtE	P	MgtE,MgtE_N,CBS,CBS	9.22	0.038	Consistent	NA	NS	NS	NA	NA
group_04227	ErfK ybiS ycfS ynhG family protein	D	YkuD	29.157	0	Consistent	NA	NS	NS	NA	NA
group_04231	NA	NA	NA	14.204	0.004	Consistent	NA	NS	NS	NA	NA
group_04261	phosphatase	P	PhoD,PhoD_N	28.409	0.001	Consistent	NA	NS	NS	NA	NA
group_04270	transcriptional regulator	K	WYL,HTH_11	0.103	0.003	Inconsistent	NA	NS	NS	NA	NA
group_04273	Sigma-70 region 2	K	Sigma70_r2,Sigma70_r4_2,SnoaL_2	0.045	0	Inconsistent	NA	13.195	0.017	Other	NA
group_04277	NA	NA	NA	0.075	0	Inconsistent	NA	NS	NS	NA	NA
group_04279	transglycosylase associated protein	S	Transgly_assoc	0.05	0	Inconsistent	NA	13.195	0.017	Other	NA
group_04296	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04298	NA	NA	NA	13.831	0.008	Consistent	NA	NS	NS	NA	NA
group_04300	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04303	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04311	TIGRFAM YihY family protein (not ribonuclease BN)	S	Virul_fac_BrkB	0.05	0	Inconsistent	NA	13.195	0.017	Other	NA
group_04321	NA	NA	NA	28.409	0.001	Consistent	NA	NS	NS	NA	NA
group_04333	NA	NA	NA	13.831	0.008	Consistent	NA	NS	NS	NA	NA

group_04334	NA	NA	NA	13.831	0.008	Consistent	NA	NS	NS	NA	NA
group_04353	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04359	NA	NA	NA	27.661	0.001	Consistent	NA	0.15	0.019	Ant	NA
group_04369	RNA polymerase sigma factor, sigma-70 family	K	Sigma70_r2,Sigma70_r4,Sigma70_r3	inf	0	Consistent	NA	0.155	0.035	Ant	NA
group_04370	Glutaminase	E	Glutaminase	27.661	0.001	Consistent	NA	NS	NS	NA	NA
group_04372	GXWXG protein	S	DUF4334,GXWXG	26.914	0.001	Consistent	NA	NS	NS	NA	NA
group_04373	DUF218 domain	V	DUF218	13.083	0.024	Consistent	NA	NS	NS	NA	NA
group_04379	CobQ/CobB/MinD/ ParA nucleotide binding domain	D	CbiA	0.077	0.001	Inconsistent	NA	NS	NS	NA	NA
group_04382	Proline dehydrogenase	E	Pro_dh	0.08	0.002	Inconsistent	NA	NS	NS	NA	NA
group_04385	Belongs to the TPP enzyme family	EH	TPP_enzyme_N,TPP_enzyme_C,TPP_enzyme_M	0.08	0.002	Inconsistent	NA	12.795	0.029	Other	NA
group_04389	Major facilitator Superfamily	EGP	MFS_1	0.08	0.002	Inconsistent	NA	NS	NS	NA	NA
group_04392	Sugar (and other) transporter	EGP	Sugar_tr	27.661	0.001	Consistent	NA	NS	NS	NA	NA
group_04393	Predicted ATPase of the ABC class	S	ABC_ATPase	13.457	0.013	Consistent	NA	NS	NS	NA	NA
group_04394	cheY-homologous receiver domain	T	GerE	inf	0	Consistent	NA	NS	NS	NA	NA
group_04395	PFAM Glyoxalase bleomycin resistance protein dioxygenase	E	Glyoxalase	inf	0	Consistent	NA	0.155	0.035	Ant	NA
group_04398	Pfam SNARE associated Golgi protein	S	SNARE_assoc	inf	0	Consistent	NA	NS	NS	NA	NA
group_04399	Bacterial PH domain	S	bPH_1	27.661	0.001	Consistent	NA	NS	NS	NA	NA
group_04402	Protein of unknown function (DUF2795)	S	DUF2795	13.457	0.013	Consistent	NA	NS	NS	NA	NA
group_04405	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04406	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04407	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04421	Iron-containing alcohol dehydrogenase	C	Fe-ADH	12.709	0.043	Consistent	NA	NS	NS	NA	NA
group_04426	4-amino-4-deoxy-L-arabinose transferase and related glycosyltransferases of PMT family	M	NA	13.457	0.013	Consistent	NA	NS	NS	NA	NA
group_04433	Transcriptional regulator	K	TetR_N	27.661	0.001	Consistent	NA	NS	NS	NA	NA
group_04438	regulation of fungal-type cell wall biogenesis	G	SMI1_KNR4	13.457	0.013	Consistent	NA	NS	NS	NA	NA
group_04439	NUDIX domain	F	NUDIX	26.914	0.001	Consistent	NA	NS	NS	NA	NA
group_04443	Selenoprotein, putative	S	Sel_put	inf	0	Consistent	NA	NS	NS	NA	NA

group_04445	4Fe-4S single cluster domain of Ferredoxin I	C	Fer4_15	13.457	0.013	Consistent	NA	NS	NS	NA	NA
group_04446	NA	NA	ACT_5	inf	0	Consistent	NA	NS	NS	NA	NA
group_04469	transposase activity	L	NA	NA	NS	NA	NA	0.05	0	Ant	NA
group_04472	Uncharacterized protein conserved in bacteria (DUF2236)	S	DUF2236	12.709	0.043	Consistent	NA	NS	NS	NA	NA
group_04476	PFAM Aldehyde dehydrogenase	C	Aldedh	0.083	0.004	Inconsistent	NA	NS	NS	NA	NA
group_04479	Asp Glu hydantoin racemase	Q	Amdase	0.111	0.014	Inconsistent	NA	NS	NS	NA	NA
group_04486	protein conserved in bacteria	S	TctC	26.914	0.001	Consistent	NA	NS	NS	NA	NA
group_04487	Antibiotic biosynthesis monooxygenase	S	ABM	inf	0	Consistent	NA	NS	NS	NA	NA
group_04489	Acetyltransferase (GNAT) domain	K	Acetyltransf_1	inf	0	Consistent	NA	0.155	0.035	Ant	NA
group_04492	3'(2'),5'-bisphosphate nucleotidase	P	NA	25.418	0.003	Consistent	NA	NS	NS	NA	NA
group_04493	NA	NA	CBS,CBS	12.709	0.043	Consistent	NA	NS	NS	NA	NA
group_04517	Belongs to the TrpF family	E	PRAI	13.083	0.024	Consistent	NA	NS	NS	NA	NA
group_04519	Bacterial regulatory proteins, tetR family	K	TetR_N	26.914	0.001	Consistent	NA	NS	NS	NA	NA
group_04523	NA	NA	Colicin_V	13.083	0.024	Consistent	NA	NS	NS	NA	NA
group_04527	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04536	Dicarboxylate carrier protein MatC N-terminus	P	MatC_N,CitMHS	12.709	0.043	Consistent	NA	NS	NS	NA	NA
group_04567	LexA-binding, inner membrane-associated putative hydrolase	S	YdjM	0.115	0.028	Inconsistent	NA	NS	NS	NA	NA
group_04571	6-O-methylguanine DNA methyltransferase, DNA binding domain	L	DNA_binding_1,Methyltransf_1N	0.083	0.004	Inconsistent	NA	NS	NS	NA	NA
group_04573	CDP-alcohol phosphatidyltransferase	I	CDP-OH_P_transf	0.083	0.004	Inconsistent	NA	NS	NS	NA	NA
group_04583	PFAM Fatty acid desaturase	I	FA_desaturase	12.709	0.043	Consistent	NA	NS	NS	NA	NA
group_04586	Catalyzes the epimerization of the C3' and C5'positions of dTDP-6-deoxy-D-xylo-4-hexulose, forming dTDP-6-deoxy-L-lyxo-4-hexulose	M	dTDP_sugar_isom	inf	0	Consistent	NA	NS	NS	NA	NA
group_04589	NA	NA	NA	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04591	NA	NA	NA	25.418	0.003	Consistent	NA	0.133	0.031	Ant	NA
group_04592	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04593	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA

group_04596	Sigma factor PP2C-like phosphatases	T	SpoIIE,Response_reg	0.027	0	Inconsistent	NA	24.79	0.007	Other	NA
group_04600	Cell wall-associated hydrolase, invasion-associated protein	M	NLPC_P60	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_04614	VWA domain containing CoxE-like protein	S	VWA_CoxE	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04617	Belongs to the SEDS family	D	FTSW_RODA_SPOVE	inf	0	Consistent	NA	NS	NS	NA	NA
group_04620	methyltransferase	Q	Methyltransf_25	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04629	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04633	NA	NA	NA	25.418	0.003	Consistent	NA	NS	NS	NA	NA
group_04636	NA	NA	NA	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04638	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04642	O-methyltransferase	Q	Methyltransf_31	0.107	0.007	Inconsistent	NA	NS	NS	NA	NA
group_04679	'Cold-shock' DNA-binding domain	K	CSD	0.028	0	Inconsistent	NA	23.99	0.013	Other	NA
group_04685	Belongs to the sigma-70 factor family. ECF subfamily	K	Sigma70_r4_2,Sigma70_r2	0.086	0.009	Inconsistent	NA	NS	NS	NA	NA
group_04688	NAD(P)H-binding	GM	NAD_binding_10	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_04691	PFAM Polyketide cyclase dehydrase and lipid transport	S	Polyketide_cyc2	0.057	0.001	Inconsistent	NA	NS	NS	NA	NA
group_04699	NA	NA	NA	0.08	0.002	Inconsistent	NA	NS	NS	NA	NA
group_04701	transcriptional regulator	K	FCD,GntR	0.052	0	Inconsistent	NA	NS	NS	NA	NA
group_04711	Cleaves peptides in various proteins in a process that requires ATP hydrolysis. Has a chymotrypsin-like activity. Plays a major role in the degradation of misfolded proteins	OU	CLP_protease	24.67	0.006	Consistent	NA	NS	NS	NA	NA
group_04719	NA	NA	NA	24.67	0.006	Consistent	NA	NS	NS	NA	NA
group_04720	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04733	Polysaccharide biosynthesis protein	S	Polysacc_synt	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04736	Belongs to the FPP GGPP synthase family	H	polyprenyl_synt	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04741	AAA domain, putative AbiEii toxin, Type IV TA system	V	ABC_tran	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04746	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04751	NA	NA	DUF3040	inf	0	Consistent	NA	NS	NS	NA	NA
group_04752	NA	NA	PepSY	12.335	0.043	Consistent	NA	NS	NS	NA	NA
group_04755	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04798	Carbon starvation protein CstA	T	CstA,CstA_5TM	inf	0.001	Consistent	NA	NS	NS	NA	NA

group_04811	Alpha amylase, catalytic domain	G	Alpha-amylase,Malt_amylase_C	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_04812	peptidase U62 modulator of DNA gyrase	S	PmbA_TldD	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_04821	Bacterial protein of unknown function (DUF948)	S	DUF948	0.057	0.001	Inconsistent	NA	NS	NS	NA	NA
group_04824	amine acid ABC transporter, permease protein, 3-TM region, His Glu Gln Arg opine family	E	BPD_transp_1	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_04831	protein deglycation	S	DJ-1_PfpI	inf	0	Consistent	NA	NS	NS	NA	NA
group_04834	transcriptional	K	MarR_2	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_04835	NA	NA	NA	inf	0.001	Consistent	NA	0.11	0.011	Ant	NA
group_04843	ATP-dependent carboxylate-amine ligase which exhibits weak glutamate--cysteine ligase activity	S	GCS2	0.06	0.002	Inconsistent	NA	23.19	0.024	Other	NA
group_04844	Metallo-beta-lactamase superfamily	S	Lactamase_B	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_04850	Alpha/beta hydrolase family	S	Abhydrolase_1	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_04851	cobalamin (vitamin B12) biosynthesis CbiX	S	CbiX,CbiX	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_04867	NA	NA	NA	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_04870	NA	NA	NA	23.923	0.011	Consistent	NA	NS	NS	NA	NA
group_04873	NA	NA	NA	23.923	0.011	Consistent	NA	NS	NS	NA	NA
group_04877	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04880	NA	NA	NA	inf	0	Consistent	NA	NS	NS	NA	NA
group_04883	Part of the twin-arginine translocation (Tat) system that transports large folded proteins containing a characteristic twin-arginine motif in their signal peptide across membranes. TatA could form the protein-conducting channel of the Tat system	U	MttA_Hcf106	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_04890	Catalyzes the formation of acetyl phosphate from acetate and ATP. Can also catalyze the reverse reaction	C	Acetate_kinase	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_04899	ErfK ybiS ycfS ynhG family protein	D	YkuD	23.175	0.02	Consistent	NA	NS	NS	NA	NA
group_04900	PFAM Glyoxalase bleomycin	E	Glyoxalase_6	23.175	0.02	Consistent	NA	NS	NS	NA	NA

	resistance protein dioxygenase										
group_04907	Polyketide cyclase / dehydrase and lipid transport	S	NA	NA	NS	NA	NA	23.99	0.013	Other	NA
group_04919	transcriptional regulators	K	HTH_3	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_04924	PFAM Glycosyl transferase family 2	M	Glyco_tranf_2_3	0.053	0	Inconsistent	NA	NS	NS	NA	NA
group_04925	NA	NA	NA	24.67	0.006	Consistent	NA	NS	NS	NA	NA
group_04931	PFAM Glucose Sorbosone dehydrogenase	G	GSDH	0.093	0.035	Inconsistent	NA	23.19	0.024	Other	NA
group_04935	GDSL-like Lipase/Acylhydrolase family	E	Lipase_GDSL_2	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_04936	Acyl CoA acetate 3-ketoacid CoA	I	CoA_trans	0.086	0.009	Inconsistent	NA	NS	NS	NA	NA
group_04939	[2Fe-2S] binding domain	C	Fer2_2,Fer2	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_04944	Bacterial regulatory proteins, tetR family	K	TetR_N	0.055	0	Inconsistent	NA	NS	NS	NA	NA
group_04945	Adenosyl cobinamide kinase adenosyl cobinamide phosphate guanylyltransferase	H	CobU	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_04946	Domain of unknown function (DUF4349)	M	DUF4349	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_04951	Sigma-70 region 2	K	Sigma70_r2,Sigma70_r4_2	0.055	0	Inconsistent	NA	NS	NS	NA	NA
group_04954	Peptidoglycan polymerase that catalyzes glycan chain elongation from lipid-linked precursors	M	Transgly	23.175	0.02	Consistent	NA	NS	NS	NA	NA
group_04957	helix_turn_helix, Lux Regulon	K	GerE,PAS_4	23.175	0.02	Consistent	NA	NS	NS	NA	NA
group_04964	Putative modulator of DNA gyrase	S	PmbA_TldD	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_04967	integral membrane protein	S	NA	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_04971	Flavin transferase that catalyzes the transfer of the FMN moiety of FAD and its covalent binding to the hydroxyl group of a threonine residue in a target flavoprotein	H	ApbE	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_04976	Helix-turn-helix domain protein	K	MLTR_LBD,HTH_31	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05032	Secreted repeat of unknown function	S	Lipoprotein_15,Lipoprotein_15	0.029	0	Inconsistent	NA	22.391	0.045	Other	NA
group_05033	Pyridoxamine 5'-phosphate oxidase	S	Pyridox_ox_2	0.031	0.001	Inconsistent	NA	22.391	0.045	Other	NA
group_05042	Single-strand binding protein family	L	SSB	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_05048	NA	NA	DUF5336	0.086	0.009	Inconsistent	NA	NS	NS	NA	NA

group_05049	Amino acid permease	E	AA_permease_2	0.086	0.009	Inconsistent	NA	NS	NS	NA	NA
group_05051	ABC transporter	S	ABC_tran,DUF4162	22.427	0.036	Consistent	NA	NS	NS	NA	NA
group_05052	Oxidoreductase molybdopterin binding domain	S	Oxidored_molyb	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_05058	Glycosyltransferase family 87	S	GT87	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_05060	Acyl CoA acetate 3-ketoacid CoA transferase, alpha subunit	I	CoA_trans	0.057	0.001	Inconsistent	NA	NS	NS	NA	NA
group_05063	Belongs to the sigma-70 factor family. ECF subfamily	K	Sigma70_r4_2,Sigma70_r2	0.057	0.001	Inconsistent	NA	NS	NS	NA	NA
group_05064	cytochrome p450	Q	p450,p450	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_05065	RibD C-terminal domain	H	RibD_C	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_05069	PFAM Multicopper oxidase	Q	Cu-oxidase_2,Cu-oxidase_3,Cu-oxidase	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_05071	molybdenum cofactor guanylyltransferase activity	H	NA	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_05072	NA	NA	NA	0.09	0.018	Inconsistent	NA	NS	NS	NA	NA
group_05077	lactoylglutathione lyase activity	E	Glyoxalase_6	22.427	0.036	Consistent	NA	NS	NS	NA	NA
group_05084	HxlR-like helix-turn-helix	K	HxlR,SCP2	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_05088	belongs to the nudix hydrolase family	F	NYN_YacP	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05094	transferase activity, transferring acyl groups other than amino-acyl groups	I	Acyl_transf_3	22.427	0.036	Consistent	NA	NS	NS	NA	NA
group_05097	phosphoribosyl-ATP diphosphatase activity	E	NA	inf	0.003	Consistent	NA	NS	NS	NA	NA
group_05100	Nicotianamine synthase protein	E	NAS	23.175	0.02	Consistent	NA	NS	NS	NA	NA
group_05103	NA	NA	NA	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_05118	cytochrome P450	Q	p450,p450	inf	0.001	Consistent	NA	0.118	0.042	Ant	NA
group_05120	helix_turn_helix, mercury resistance	K	MerR_1	inf	0.001	Consistent	NA	0.118	0.042	Ant	NA
group_05122	NA	NA	NA	inf	0.001	Consistent	NA	0.118	0.042	Ant	NA
group_05123	Belongs to the anti-sigma-factor antagonist family	T	STAS_2	22.427	0.036	Consistent	NA	0.118	0.042	Ant	NA
group_05127	Phosphate transporter family	P	PHO4	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05128	Protein of unknown function DUF47	P	PhoU_div	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05149	Bacterial regulatory proteins, tetR family	K	TetR_N,TetR_C_13	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_05157	Diguanylate cyclase	T	EAL,GGDEF,PAS	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_05158	ABC transporter, ATP-binding protein	E	ABC_tran	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA

group_05166	Zn-finger in ubiquitin-hydrolases and other protein	O	zf-UBP	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05167	cheY-homologous receiver domain	T	Response_reg	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05170	Tautomerase enzyme	S	Tautomerase,Tautomerase	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_05172	Protein of unknown function (DUF998)	S	DUF998	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_05173	NA	NA	NA	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_05176	NA	NA	NA	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_05177	NA	NA	NA	0.06	0.002	Inconsistent	NA	NS	NS	NA	NA
group_05189	NA	NA	DUF2993	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_05193	NA	NA	NA	inf	0.001	Consistent	NA	NS	NS	NA	NA
group_05206	Protein of unknown function (DUF2537)	S	DUF2537	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05241	NUDIX hydrolase	L	NUDIX	22.427	0.036	Consistent	NA	NS	NS	NA	NA
group_05268	Rieske-like [2Fe-2S] domain	P	Rieske	0.093	0.035	Inconsistent	NA	NS	NS	NA	NA
group_05277	Acetyltransferase (GNAT) domain	J	Acetyltransf_3	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05279	CHASE3 domain	T	CHASE3,HATPase_c,HisKA,HAMP	0.032	0.001	Inconsistent	NA	NS	NS	NA	NA
group_05283	Transcriptional regulator	K	MarR_2	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05288	YCII-related domain	S	YCII	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05292	NA	NA	NA	0.03	0	Inconsistent	NA	NS	NS	NA	NA
group_05293	sigma factor antagonist activity	T	HATPase_c_2	0.062	0.005	Inconsistent	NA	NS	NS	NA	NA
group_05296	NA	NA	NA	0.03	0	Inconsistent	NA	NS	NS	NA	NA
group_05300	Permease MlaE	Q	NA	NA	NS	NA	NA	0.096	0.032	Ant	NA
group_05304	NA	NA	NA	inf	0.011	Consistent	NA	NS	NS	NA	NA
group_05306	NA	NA	NA	inf	0.003	Consistent	NA	NS	NS	NA	NA
group_05309	Belongs to the IlvD Edd family	EG	ILVD_EDD	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05322	MerR HTH family regulatory protein	K	MerR_1,B12-binding,B12-binding_2	0.031	0.001	Inconsistent	NA	inf	0.008	Other	NA
group_05346	NA	NA	NA	inf	0.002	Consistent	NA	NS	NS	NA	NA
group_05348	NA	NA	NA	inf	0.002	Consistent	NA	NS	NS	NA	NA
group_05350	NA	NA	NA	inf	0.003	Consistent	NA	NS	NS	NA	NA
group_05351	NA	NA	NA	inf	0.003	Consistent	NA	NS	NS	NA	NA
group_05355	TIGRFAM proton-translocating NADH-quinone oxidoreductase, chain M	C	NA	NA	NS	NA	NA	inf	0.002	Other	NA
group_05362	conserved protein, contains double-stranded beta-helix domain	S	Cupin_2	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_05376	NA	NA	NA	inf	0.002	Consistent	NA	NS	NS	NA	NA
group_05401	Peptidase family M48	O	NA	NA	NS	NA	NA	inf	0.004	Other	NA

group_05432	Belongs to the alpha-IPM synthase homocitrate synthase family	E	HMGL-like,LeuA_dimer	0.071	0.049	Inconsistent	NA	NS	NS	NA	NA
group_05435	Putative neutral zinc metallopeptidase	S	Zn_peptidase	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05436	Acetyltransferase (GNAT) family	K	Acetyltransf_3,ACT	0.071	0.049	Inconsistent	NA	NS	NS	NA	NA
group_05438	Nitrilase cyanide hydratase and apolipoprotein N-acyltransferase	S	CN_hydrolase	0.039	0.036	Inconsistent	NA	NS	NS	NA	NA
group_05444	TIGRFAM amine acid ABC transporter, permease protein, 3-TM region, His Glu Gln Arg opine family	E	BPD_transp_1	0.071	0.049	Inconsistent	NA	NS	NS	NA	NA
group_05446	NA	NA	NA	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05447	Glycosyl hydrolase family 76	G	Glyco_hydro_76	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_05450	PFAM Glyoxalase bleomycin resistance protein dioxygenase	E	Glyoxalase_6	0.031	0.001	Inconsistent	NA	inf	0.014	Other	NA
group_05456	PFAM YCII-related	S	YCII	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05466	NA	NA	NA	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05468	SCP-2 sterol transfer family	S	SCP2	inf	0.011	Consistent	NA	NS	NS	NA	NA
group_05496	iron ion transport	P	HemS	inf	0.006	Consistent	NA	0.096	0.032	Ant	NA
group_05502	Tripartite tricarboxylate transporter TctB family	S	TctB	inf	0.003	Consistent	NA	NS	NS	NA	NA
group_05506	CAAX protease self-immunity	S	CPBP	inf	0.006	Consistent	NA	NS	NS	NA	NA
group_05508	Protein of unknown function, DUF485	S	DUF485	inf	0.006	Consistent	NA	NS	NS	NA	NA
group_05510	NA	NA	NA	inf	0.006	Consistent	NA	NS	NS	NA	NA
group_05516	Nitrate reductase gamma subunit	C	Nitrate_red_gam	0.039	0.036	Inconsistent	NA	inf	0.014	Other	NA
group_05518	Nitrate reductase delta subunit	C	Nitrate_red_del	0.039	0.036	Inconsistent	NA	inf	0.014	Other	NA
group_05530	membrane	S	NA	inf	0.006	Consistent	NA	NS	NS	NA	NA
group_05534	NA	NA	Usp,Usp,Usp	inf	0.038	Consistent	NA	0.07	0.037	Ant	NA
group_05536	NA	NA	NA	inf	0.006	Consistent	NA	NS	NS	NA	NA
group_05544	Belongs to the long-chain O-acyltransferase family	Q	WES_acyltransf,DUF1298	0.071	0.049	Inconsistent	NA	NS	NS	NA	NA
group_05548	NA	NA	NA	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05559	NA	NA	NA	0.065	0.011	Inconsistent	NA	NS	NS	NA	NA
group_05578	NA	NA	NA	0.031	0.001	Inconsistent	NA	NS	NS	NA	NA
group_05580	PFAM Integrase catalytic	L	rve,HTH_21	inf	0.02	Consistent	NA	0.032	0.001	Ant	NA
group_05584	Belongs to the class-III pyridoxal-phosphate-dependent aminotransferase family	E	Aminotran_3	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA

group_05591	Sigma-70 region 2	K	Sigma70_r2,Sigma70_r4_2,SnoaL_2	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_05601	NA	NA	Usp	0.032	0.001	Inconsistent	NA	NS	NS	NA	NA
group_05604	NA	NA	NA	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_05605	NA	NA	NA	0.032	0.001	Inconsistent	NA	NS	NS	NA	NA
group_05613	Heat shock 70 kDa protein	O	NA	NA	NS	NA	NA	0.067	0.019	Ant	NA
group_05614	NA	NA	NA	inf	0.011	Consistent	NA	NS	NS	NA	NA
group_05620	NA	NA	NA	inf	0.02	Consistent	NA	NS	NS	NA	NA
group_05624	Transcriptional regulator	K	Aminotran_1_2,GntR	0.037	0.016	Inconsistent	NA	NS	NS	NA	NA
group_05626	Transport permease protein	V	ABC2_membrane	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_05627	With LigD forms a non-homologous end joining (NHEJ) DNA repair enzyme, which repairs dsDNA breaks with reduced fidelity. Binds linear dsDNA with 5'- and 3'-overhangs but not closed circular dsDNA nor ssDNA. Recruits and stimulates the ligase activity of LigD	L	Ku	0.071	0.049	Inconsistent	NA	NS	NS	NA	NA
group_05655	response regulator	T	ANTAR,PAS_4	inf	0.011	Consistent	NA	NS	NS	NA	NA
group_05660	NA	NA	NA	inf	0.006	Consistent	NA	NS	NS	NA	NA
group_05664	Proton-conducting membrane transporter	C	NA	NA	NS	NA	NA	inf	0.008	Other	NA
group_05684	Branched-chain amino acid transport system / permease component	E	BPD_transp_2	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA
group_05689	Transfers and isomerizes the ribose moiety from AdoMet to the 7-aminomethyl group of 7-deazaguanine (preQ1-tRNA) to give epoxyqueuosine (oQ-tRNA)	J	Queuosine_synth	0.071	0.049	Inconsistent	NA	NS	NS	NA	NA
group_05691	Predicted permease	S	ArsP_1	0	0	Inconsistent	NA	inf	0.027	Other	NA
group_05707	Belongs to the sigma-70 factor family. ECF subfamily	K	Sigma70_r4_2,Sigma70_r2	0.068	0.023	Inconsistent	NA	NS	NS	NA	NA
group_05733	Alanine-glyoxylate amino-transferase	EK	Aminotran_1_2	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_05739	Immunoglobulin-like domain of bacterial spore germination	S	Germane,Gmad2	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA
group_05745	Mycothiol maleylpyruvate	S	MDMPI_N	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA

	isomerase N-terminal domain										
group_05747	AntiSigma factor	K	zf-HC2	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA
group_05757	NA	NA	NA	inf	0.038	Consistent	NA	NS	NS	NA	NA
group_05760	NA	NA	NA	inf	0.02	Consistent	NA	NS	NS	NA	NA
group_05766	AAA ATPase domain	K	AAA_16	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA
group_05767	Hydrolase	S	HAD_2	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA
group_05782	Alpha beta hydrolase	S	Hydrolase_4	inf	0.02	Consistent	NA	NS	NS	NA	NA
group_05792	Nitrate reductase beta subunit	C	NA	NA	NS	NA	NA	inf	0.049	Other	NA
group_05800	NDH-1 shuttles electrons from NADH, via FMN and iron- sulfur (Fe-S) centers, to quinones in the respiratory chain	C	NA	NA	NS	NA	NA	inf	0.027	Other	NA
group_05814	Protein of unknown function (DUF1275)	S	DUF1275	inf	0.02	Consistent	NA	NS	NS	NA	NA
group_05830	helix_turn_helix, Lux Regulon	K	GerE	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA
group_05832	Specifically methylates the pseudouridine at position 1915 (m3Psi1915) in 23S rRNA	J	NA	0.034	0.003	Inconsistent	NA	NS	NS	NA	NA
group_05838	F420H(2)-dependent quinone reductase	S	F420H2_quin_red	inf	0.02	Consistent	NA	NS	NS	NA	NA
group_05841	NA	NA	NA	inf	0.038	Consistent	NA	NS	NS	NA	NA
group_05842	NA	NA	NA	inf	0.038	Consistent	NA	NS	NS	NA	NA
group_05919	NA	NA	NA	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_05922	Domain of unknown function (DUF5134)	S	DUF5134	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_05931	Lsr2	S	Lsr2	inf	0.038	Consistent	NA	NS	NS	NA	NA
group_05934	Membrane dipeptidase (Peptidase family M19)	E	Peptidase_M19	0.037	0.016	Inconsistent	NA	NS	NS	NA	NA
group_05953	NA	NA	NA	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_05972	NA	NA	NA	inf	0.038	Consistent	NA	NS	NS	NA	NA
group_05984	NADH ubiquinone oxidoreductase subunit 1	C	NA	NA	NS	NA	NA	inf	0.049	Other	NA
group_05988	Belongs to the complex I subunit 6 family	C	NA	NA	NS	NA	NA	inf	0.027	Other	NA
group_05990	NDH-1 shuttles electrons from NADH, via FMN and iron- sulfur (Fe-S) centers, to quinones in the respiratory chain. The immediate electron acceptor for the enzyme in this species is believed to be a menaquinone. Couples the redox	C	NA	NA	NS	NA	NA	inf	0.049	Other	NA

	reaction to proton translocation (for every two electrons transferred, four hydrogen ions are translocated across the cytoplasmic membrane), and thus conserves the redox energy in a proton gradient										
group_05991	NA	NA	NA	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06032	Catalyzes the oxidation of glucose 6-phosphate to 6-phosphogluconolactone	G	G6PD_C,G6PD_N	0	0.001	Inconsistent	NA	NS	NS	NA	NA
group_06034	Psort location CytoplasmicMembrane, score	S	NA	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06036	NA	NA	NA	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06090	Domain in cystathionine beta-synthase and other proteins.	S	CBS,CBS,BON	0	0	Inconsistent	NA	NS	NS	NA	NA
group_06099	UDP-glucose/GDP-mannose dehydrogenase family, central domain	M	UDPG_MGDP_dh,UDPG_MGDP_dh_N,UDPG_MGDP_dh_N	0.036	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06156	Protein of unknown function (DUF3311)	S	DUF3311	0.037	0.016	Inconsistent	NA	NS	NS	NA	NA
group_06235	Acetoacetate decarboxylase (ADC)	Q	ADC	0.037	0.016	Inconsistent	NA	NS	NS	NA	NA
group_06240	Pfam Response regulator receiver	T	Response_reg	0.037	0.016	Inconsistent	NA	NS	NS	NA	NA
group_06255	NA	NA	NA	NA	NS	NA	NA	0.036	0.007	Ant	NA
group_06268	NADH ubiquinone oxidoreductase subunit 5 (chain L) Multisubunit Na H antiporter, MnhA subunit	CP	NA	NA	NS	NA	NA	inf	0.049	Other	NA
group_06290	Universal stress protein	T	Usp,Usp	0	0.001	Inconsistent	NA	NS	NS	NA	NA
group_06291	NA	NA	NA	0	0.001	Inconsistent	NA	NS	NS	NA	NA
group_06292	NA	NA	NA	0	0.001	Inconsistent	NA	NS	NS	NA	NA
group_06294	NA	NA	NA	0.037	0.016	Inconsistent	NA	NS	NS	NA	NA
group_06326	Domain present in phytochromes and cGMP-specific phosphodiesterases.	T	HisKA_3,HATPase_c,GAF	0.039	0.036	Inconsistent	NA	NS	NS	NA	NA
group_06334	NA	NA	NA	NA	NS	NA	NA	0.04	0.034	Ant	NA
group_06381	PFAM Extracellular ligand-binding receptor	E	Peripla_BP_6	0.039	0.036	Inconsistent	NA	NS	NS	NA	NA
group_06415	Major Facilitator Superfamily	EGP	MFS_1	0	0.003	Inconsistent	NA	NS	NS	NA	NA
group_06420	Domain of unknown function (DUF1980)	S	DUF1980	0	0.001	Inconsistent	NA	NS	NS	NA	NA

group_06480	Glyoxalase/Bleomycin resistance protein/Dioxygenase superfamily	E	Glyoxalase, Glyoxalase	0.039	0.036	Inconsistent	NA	NS	NS	NA	NA
group_06482	NA	NA	NA	0.039	0.036	Inconsistent	NA	NS	NS	NA	NA
group_06488	Forms part of the polypeptide exit tunnel	J	NA	0.039	0.036	Inconsistent	NA	NS	NS	NA	NA
group_06538	ATPases associated with a variety of cellular activities	S	ABC_tran,ABC_tran,ABC_tran_Xtn	0	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06607	NAD dependent epimerase/dehydratase family	IQ	adh_short	0	0.003	Inconsistent	NA	NS	NS	NA	NA
group_06611	Sigma factor PP2C-like phosphatases	KT	SpoIIE	0	0.017	Inconsistent	NA	NS	NS	NA	NA
group_06613	Protein of unknown function (DUF541)	S	SIMPL	0	0.003	Inconsistent	NA	NS	NS	NA	NA
group_06616	helix_turn_helix, Lux Regulon	K	Response_reg,GerE	0	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06641	NA	NA	DUF4097	0	0.003	Inconsistent	NA	NS	NS	NA	NA
group_06727	D-isomer specific 2-hydroxyacid dehydrogenase, NAD binding domain	EH	2-Hacid_dh_C,2-Hacid_dh	0	0.039	Inconsistent	NA	NS	NS	NA	NA
group_06740	acyl-CoA dehydrogenase	I	NA	NA	NS	NA	NA	0	0.006	Ant	NA
group_06767	Extracellular solute-binding protein, family 5	E	SBP_bac_5	0	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06922	ABC-type transport system involved in resistance to organic solvents, ATPase component	Q	NA	NA	NS	NA	NA	0	0.014	Ant	NA
group_06969	NA	NA	NA	0	0.007	Inconsistent	NA	NS	NS	NA	NA
group_06973	Belongs to the anti-sigma-factor antagonist family	T	STAS_2	0	0.017	Inconsistent	NA	NS	NS	NA	NA
group_06982	C-terminal PDZ domain	O	Trypsin_2,PDZ_2	0	0.017	Inconsistent	NA	NS	NS	NA	NA
group_07000	NA	NA	NA	0	0.017	Inconsistent	NA	NS	NS	NA	NA
group_07002	NA	NA	NA	0	0.017	Inconsistent	NA	NS	NS	NA	NA
group_07003	Acetyltransferase (GNAT) domain	S	Acetyltransf_9	0	0.017	Inconsistent	NA	NS	NS	NA	NA
group_07015	Capsule synthesis protein PGA_cap	M	PGA_cap	0	0.039	Inconsistent	NA	NS	NS	NA	NA
group_07198	NA	NA	NA	0	0.017	Inconsistent	NA	NS	NS	NA	NA
group_07254	NA	NA	NA	0	0.039	Inconsistent	NA	NS	NS	NA	NA
group_07409	Domain of unknown function (DUF397)	S	DUF397	0	0.039	Inconsistent	NA	NS	NS	NA	NA
group_07525	PFAM Transposase, IS4-like	L	DDE_Tnp_1_3	0	0.039	Inconsistent	NA	NS	NS	NA	NA
group_07712	Transposase	L	DDE_Tnp_1,DUF4096	25.418	0.003	Consistent	NA	NS	NS	NA	NA

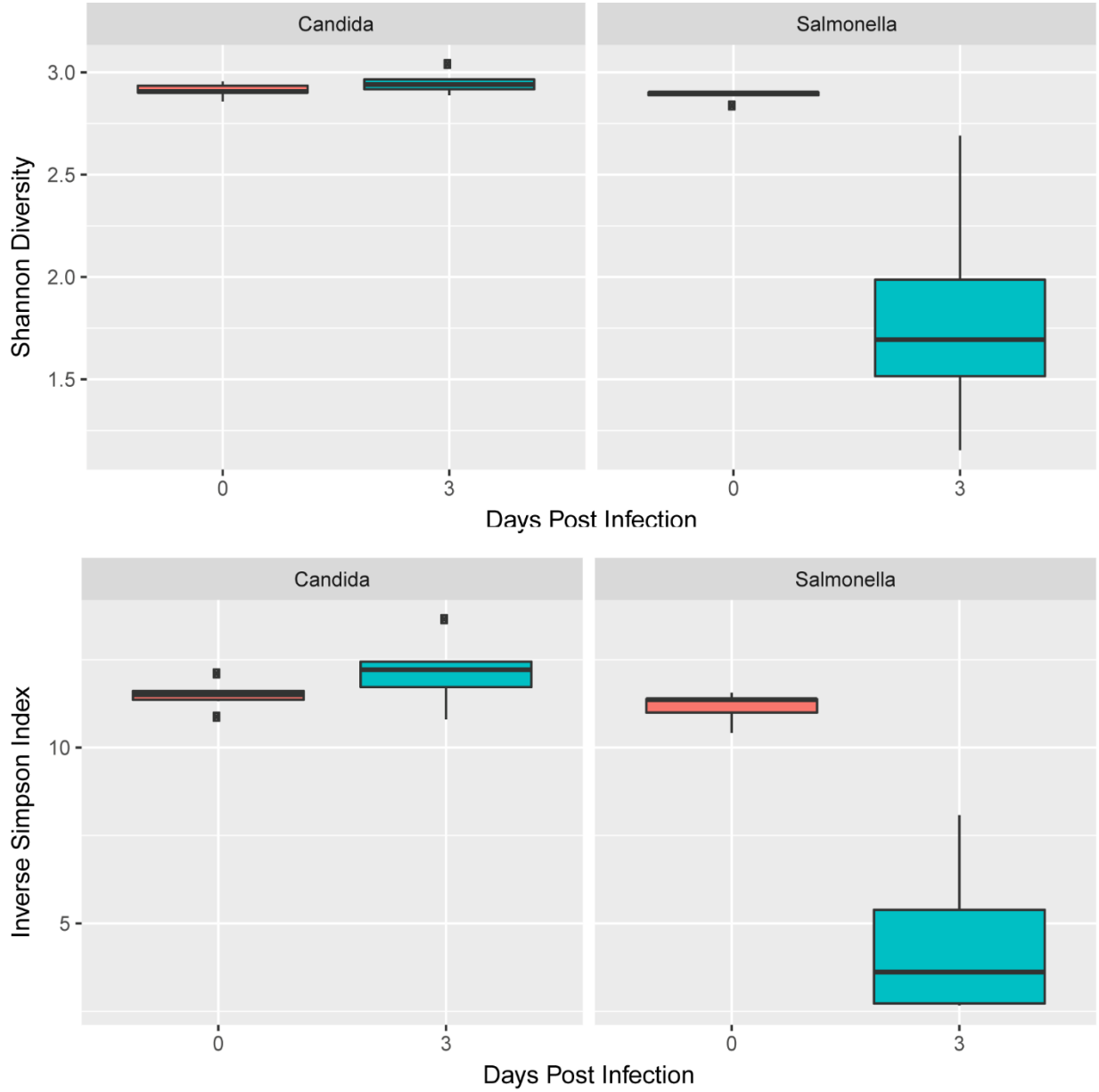
Supplemental Table 2. Gene enrichment in ant versus non-ant strains and consistent versus inconsistent colonizing strains with TreeWAS and Fisher's Exact Test. Values rounded to three decimal places.

Strain	Sequencing	Source	N50	Genome_Length	Clusters	totalPKS/NRPS
Pseudo_AL041002_03	PacBio	Ant	6143341	5978138	10	2
Pseudo_AL041002_03	Illumina	Ant	66675	6114861	11	3
Pseudo_CC030328_06	PacBio	Ant	6654308	6739435	20	7
Pseudo_CC030328_06	Illumina	Ant	150155	6763995	20	7
Pseudo_cypriaca	Illumina	Non-ant	14109	8016379	8	3
Pseudo_cypriaca	PacBio	Non-ant	2964811	8279222	9	1
Pseudo_EC080610_09	PacBio	Ant	6138223	7131853	17	6
Pseudo_EC080610_09	Illumina	Ant	15007	7201661	33	18
Pseudo_EC080625_04	PacBio	Ant	6135769	6554452	15	4
Pseudo_EC080625_04	Illumina	Ant	9505	6308602	17	7
Pseudo_endophytica	Illumina	Non-ant	22766	7567994	36	27
Pseudo_endophytica	PacBio	Non-ant	4021098	7487432	17	8
Pseudo_hierapolitana	Illumina	Non-ant	43734	8856958	9	2
Pseudo_hierapolitana	PacBio	Non-ant	8856958	8771874	10	2
Pseudo_SP020602_02	PacBio	Ant	6322523	6322523	14	5
Pseudo_SP020602_02	Illumina	Ant	8274	6178162	23	10
Pseudo_alni_PB	PacBio	Non-ant	5686562	5994807	11	1
Pseudo_alni_DSM44104	Illumina	Non-ant	4886	5661871	10	0

Supplemental Table 3. BGC abundance in strains sequenced with PacBio or Illumina technology.

Appendix 2: Supplemental Materials for Chapter 3

Supplemental Figure 1 Community diversity before and 3 days postinfection.



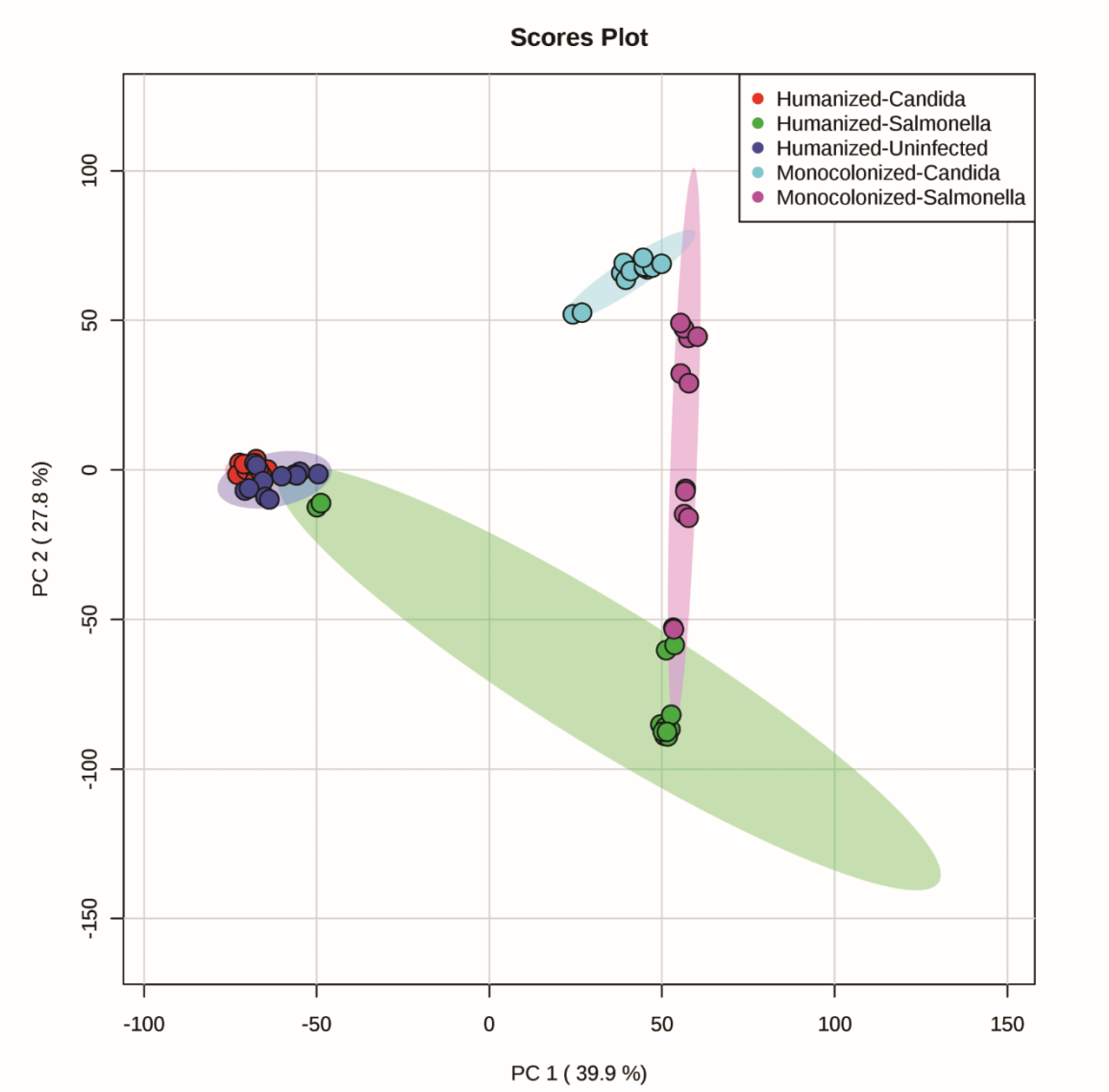
Supplemental Figure 1. Community diversity before and 3 days post-infection.

Supplemental Table 1. Biosynthetic gene clusters predicted in humanized microbiota

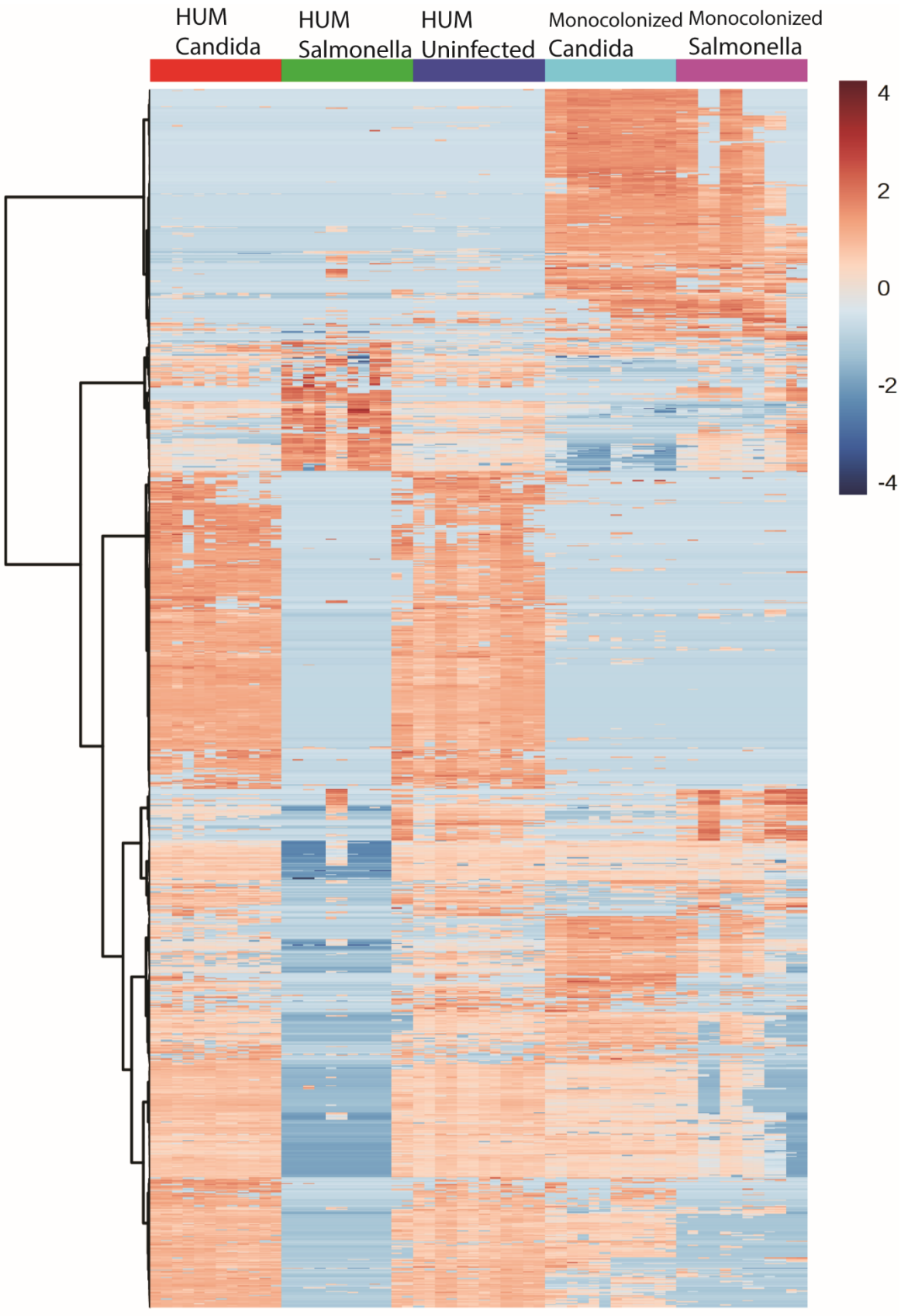
Cluster Name	Count
putative	486
saccharide	345

fatty_acid	117
sactipeptide	34
fatty acid -saccharide	22
nrps	19
arylpolyene	14
thiopeptide	12
bacteriocin	10
siderophore	4
lantipeptide	3
other	3
resorcinol	3
terpene	2
hserlactone	2
sactipeptide-cf_saccharide	1
sactipeptide-nrps	1
bacteriocin-proteusin	1
sactipeptide-lantipeptide	1
t1pks-nrps	1

Supplemental Figure 2. Principal component analysis of all metabolites detected.



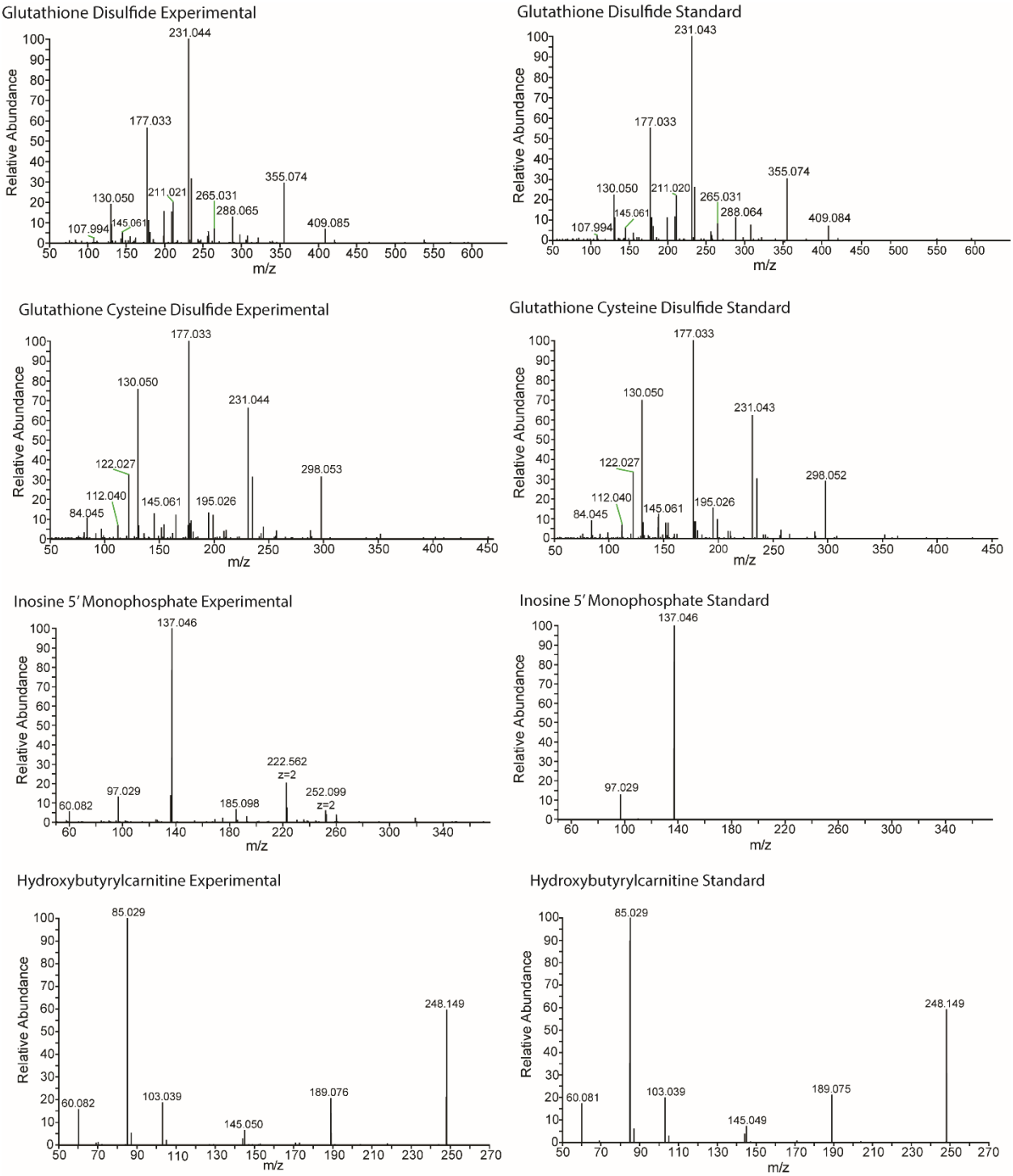
Supplemental Figure 1. Principal component analysis of all metabolites detected.



Supplemental Figure 2. Heatmap of all m/z detected.

Supplemental Figure 3. Heatmap of all m/z detected.

Supplemental Figure 4. MS/MS spectra of experimental compounds and matching standards.



Supplemental Figure 3. MS/MS spectra of experimental compound (left) and matching standard (right).

Supplemental Table 2. Features of interested in humanized infected mice

Molecular weight	**Retention time (min)**	**charge**	**MS/MS Identification**	**Overabundant in *Salmonella* or *Candida***
247.1417	1.391	1	Hydroxybutyrylcarnitine	*Salmonella*
247.1417	1.625	1	Hydroxybutyrylcarnitine	*Salmonella*
256.1401	0.799	1		*Salmonella*
261.0303	1.772	1		*Salmonella*
285.1143	1.002	1		*Salmonella*
336.0561	1.16	1		*Salmonella*
347.0626	1.052	1		Both
348.0467	1.08	1	Inosine monophosphate	*Salmonella*
363.0575	1.093	1		*Salmonella*
371.2516	13.487	1		*Salmonella*
426.0879	0.832	1	Glutathione-cysteine disulfide	*Salmonella*
483.1088	1.048	2		*Salmonella*
503.8396	0.741	1		*Salmonella*
508.3609	22.75	1		*Salmonella*
555.13	1.437	2		*Salmonella*
612.1513	1.789	1	Glutathione disulfide	*Salmonella*
635.3753	13.747	2		*Salmonella*
691.8275	0.733	1		*Salmonella*
701.492	22.63	1		*Salmonella*
726.3701	13.302	2		*Salmonella*
759.8148	0.731	1		*Salmonella*
837.8299	0.72	1		*Salmonella*
205.0773	1.597	1		*Candida*
263.0904	1.592	1		*Candida*
268.0518	9.114	1		*Candida*
336.1797	13.536	1		*Candida*
457.2576	14.042	1		*Candida*
487.2318	18.094	1		*Candida*
487.268	14.205	1		*Candida*
514.3227	11.698	2		*Candida*
577.2232	1.429	1		*Candida*

Supplemental Table 3. Strains used in humanized community

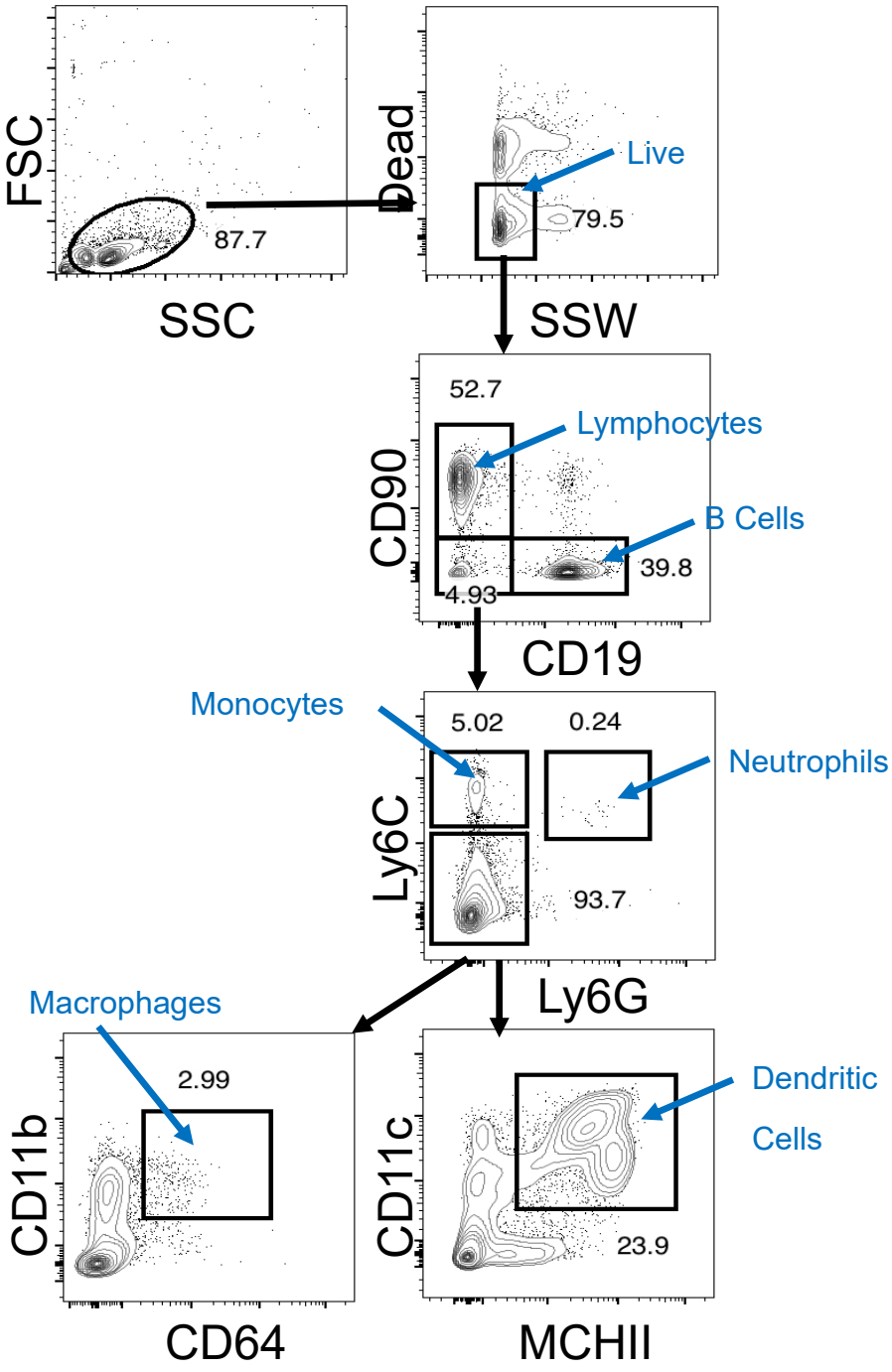
Genus	**Species**	**ATCC**	**DSMZ**	Also known as
Akkermansia	*muciniphila*	BAA-835	22959	
Alistipes	*indistinctus*	NA	22520	
Anaerococcus	*hydrogenalis*	49630	7454	
Anaerotruncus	*colihominis*	na	17241	
Bacteroides	*caccae*	43185	19024	
Bacteroides	*cellulosilyticus*	na	14838	
Bacteroides	*coprophilus*	na	18228	
Bacteroides	*dorei*	na	17855	
Bacteroides	*eggerthii*	27754	20697	
Bacteroides	*finegoldii*	na	17565	
Bacteroides	*intestinalis*	na	17393	
Bacteroides	*ovatus*	8483	na	
Bacteroides	*plebeius*	na	17135	
Bacteroides	*stercoris*	43183	na	
Bacteroides	*thetaiotaomicron3731*	na	na	
Bacteroides	*thetaiotaomicron7330*	na	na	
Bacteroides	*thetaiotaomicronVPI-5482*	29148	na	
Bacteroides	*uniformis*	8492	na	
Bacteroides	*vulgatus*	8482	na	
Bacteroides	*WH2*	na	na	*Bacteroides thetaiotamicron, Bacteroides cellulolyticus*
Bacteroides	*xylanisolvens*	na	18836	
Bifidobacterium	*adolescentis*	15703	na	
Bifidobacterium	*angulatum*	27535	20098	
Bifidobacterium	*bifidum*	29521	20456	
Bifidobacterium	*dentium*	27678	na	
Bifidobacterium	*pseudocatenulatum*	27919	20438	
Blautia	*hansenii*	27752	20583	
Blautia	*luti*	na	14534	
Catenibacterium	*mitsuokai*	na	15897	

Citrobacter	*youngae*	29220	na	
Clostridium	*asparagiforme*	na	15981	
Clostridium	*bartlettii*	na	16795	*Intestinibacter bartlettii*
Clostridium	*bolteae*	BAA-613	15670	
Clostridium	*hathewayi*	na	13479	
Clostridium	*hiranonis*	na	13275	
Clostridium	*hylemonae*	na	15053	
Clostridium	*leptum*	29065	753	
Clostridium	M62_1	na	na	
Clostridium	*nexile*	27757	1787	
Clostridium	*nexile-related*	na	na	*Tyzzerella nexilis*
Clostridium	*ramosum*	25582	1402	
Clostridium	*scindens*	35704	5676	
Clostridium	*spiroforme*	29900	1552	
Clostridium	*sporogenes*	15579	na	
Clostridium	*symbiosum*	14940	934	
Collinsella	*aerofaciens*	25986	3979	
Collinsella	aerofaciens	25986	3979	
Collinsella	*intestinalis*	na	13280	
Collinsella	*stercoris*	na	13279	
Coprococcus	*comes*	27758	na	
Coprococcus	*eutactus*	27759	na	
Desulfovibrio	*piger*	na	na	GOR1
Dorea	*formicigenerans*	27755	3992	
Dorea	*longicatena*	na	13814	
Edwardsiella	*tarda*	23685	na	
Edwardsiella	tarda	23685	na	
Enterobacter	*cancerogenus*	35316	na	
Escherichia	*coliK12*	na		
Escherichia	*fergusonii*	35469	13698	
Eubacterium	*biforme*	27806	3989	
Eubacterium	*cylindroides*	na	na	
Eubacterium	*dolichum*	29143	3991	
Eubacterium	*eligens*	27750	3376	
Eubacterium	*hallii*	27751	3353	
Eubacterium	*plautii*	29863	na	*Clostridium orbscindens; Flavonifractor plautii*
Eubacterium	rectale	33656		
Eubacterium	*ventriosum*	27560	na	

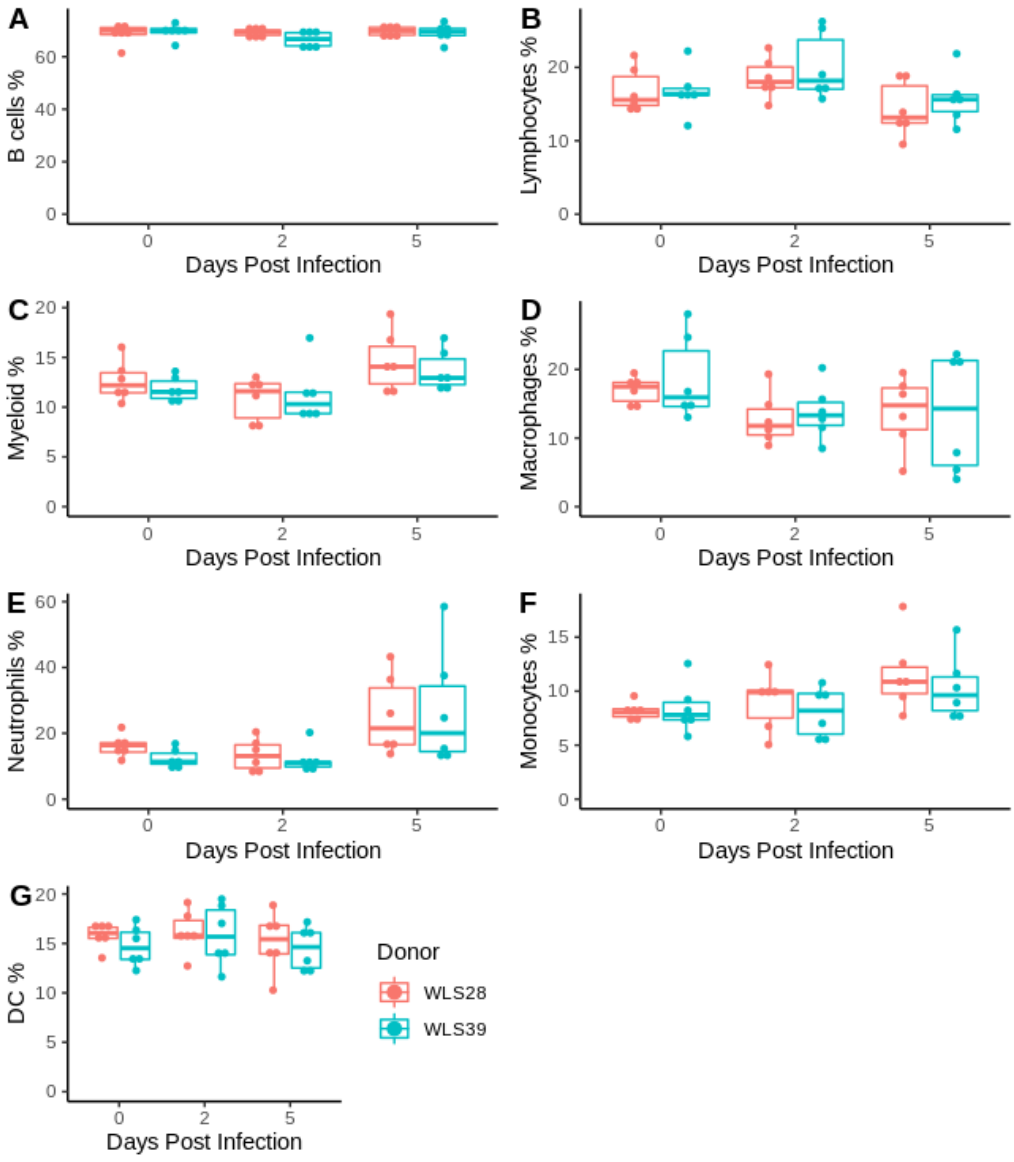
Faecalibacterium	*prausnitzii M21/2*	na	na	
Fusobacterium	*varium*	na	na	JCM6320
Holdemania	*filiformis*	51649	12042	
Lactobacillus	*reuteri*	na	20016	
Lactobacillus	*ruminis*	27780	20403	
Marvinbryantia	*formatexigens*	na	14469	
Megamonas	*funiformis*	na	19343	
Mitsuokella	*multacida*	27723	20544	
Parabacteroides	*distasonis*	8503	20701	
Parabacteroides	*johnsonii*	na	18315	
Parabacteroides	*merdae*	43184	19495	
Proteus	*penneri*	35198	na	
Providencia	*alcalifaciens*	na	na	
Providencia	*rettgeri*	na	1131	
Providencia	*rustigianii*	33673	4541	
Providencia	*stuartii*	25827	na	*Clostridium* sp. GM2/1
Roseburia	*intestinalis*	na	14610	
Ruminococcus	*gnavus*	29149	na	
Ruminococcus	*hydrogenotrophicus*	na	10507	*Blautia hydrogenotrophicus*
Ruminococcus	*lactaris*	29176	na	
Ruminococcus	*obeum*	na	na	
Ruminococcus	*torques*	27756	na	
Streptococcus	*infantarius*	BAA-102	na	
Subdoligranulum	*variabile*	na	15176	
Victivallis	*vadensis*	BAA-548	14823	

Appendix 2: Supplemental Materials for Chapter 4

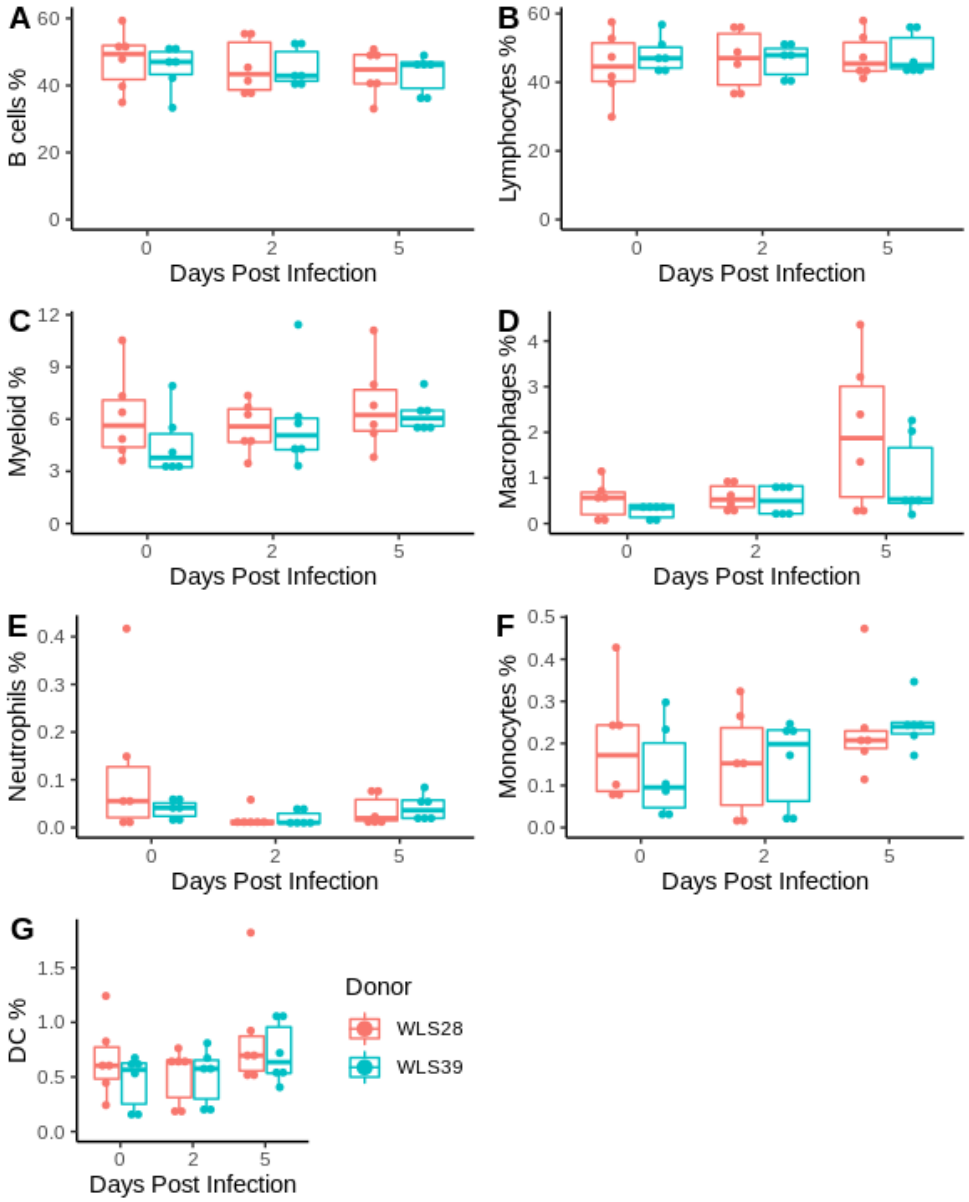
Supplemental Figure 1. Survival plot with 95% confidence intervals.



Supplemental Figure 2. Percentage of immune cells detected in spleen



Supplemental Figure 3. Percentage of immune cells detected in mesenteric lymph nodes


A
B cells %
Days Post Infection
B
Lymphocytes %
Days Post Infection
C
Myeloid %
Days Post Infection
D
Macrophages %
Days Post Infection
E
Neutrophils %
Days Post Infection
F
Monocytes %
Days Post Infection
G
DC %
Days Post Infection
Donor
WLS28
WLS39

Appendix 4: Experimental Microbiomes: Models Not to Scale

Marc G. Chevrette, Jennifer R. Bratburd, Cameron R. Currie, Reed M. Stubbendieck

Reproduced from Chevrette MG, Bratburd JR, Currie CR, Stubbendieck RM. 2019. Experimental microbiomes: models not to scale. mSystems 4:e00175-19. https://doi.org/10.1128/mSystems.00175-19.

M.G.C. and J.R.B. contributed equally to this work.

A4.1 Abstract

Low-cost, high-throughput nucleic acid sequencing ushered the field of microbial ecology into a new era in which the microbial composition of nearly every conceivable environment on the planet is under examination. However, static "screenshots" derived from sequence-only approaches belie the underlying complexity of the microbe-microbe and microbe-host interactions occurring within these systems. Reductionist experimental models are essential to identify the microbes involved in interactions and to characterize the molecular mechanisms that manifest as complex host and environmental phenomena. Herein, we focus on three models (*Bacillus-Streptomyces*, Aliivibrio fischeri-Hawaiian bobtail squid, and gnotobiotic mice) at various levels of taxonomic complexity and experimental control used to gain molecular insight into microbe-mediated interactions. We argue that when studying microbial communities, it is crucial to consider the scope of questions that experimental systems are suited to address, especially for researchers beginning new projects. Therefore, we highlight practical applications, limitations, and tradeoffs inherent to each model.

A4.2 Perspective

Microbiomes shape the fundamental biology of environments and can have substantial impacts on macroscopic ecosystems. Within their hosts, microbiomes alter metabolism, behavior, and disease. Experimental insight into the molecular mechanisms underlying

microbiome interactions remains elusive. High complexity, variable plasticity, and low manipulability of natural systems remain barriers to recapitulating microbiomes in the laboratory.

Distilling the extreme complexity of biology into discrete, functional units remains a difficult challenge. As early as 1662, René Descartes posited that biology could be explained as collectives of self-operating machinery termed “automata” (1). We have dissected the molecular nature of these “machines” into their constituent parts. For example, forward genetic screens, reverse genetics, and complementation aim to connect genomic loci with organism-level effects and are invaluable in understanding how genes function and phenotypes manifest. As we increasingly appreciate how microbes influence ecology and host fitness, models are essential to limit complexity and maximize experimental control, such that we can begin to understand how interactions within microbial communities influence biology (2). From a microbial perspective, understanding the influences of fitness can resolve common and distinct features of microbial interactions in different systems. While microbial fitness is often conceived of as a static property, the dynamics of microbial interactions are shaped by environmental and temporal plasticity and competition. Thus, the phenotypes that shape microbial fitness are the sum of many variables, including but not limited to the presence and regulation of genes, the interspecies interactions of a microbial community, and chemical gradients (3). Further, emergent properties of microbial communities can confound the simplest studies. For example, different combinations of relatively simple ≤5 member communities in *Drosophila* can mediate changes in host life span and fecundity, with some members influencing these traits only in the presence of certain other community members (4). Considering this complexity, model systems

integrating reductionist experimental frameworks are necessary to link the underlying interaction networks of microbiomes to host biology.

For early career researchers and researchers embarking on new projects, it is important to understand the kinds of questions that certain models address well and where reduction can maximize experimental control with minimal loss of biological relevance. Herein, we describe three model systems with different levels of manipulability and complexity which have been used to uncover molecular mechanisms of interactions. First, we discuss *Bacillus-Streptomyces* pairwise interactions to highlight the high experimental control and manipulability of this system used to uncover molecular mechanisms of microbial competition. We then discuss the *Aliivibrio*-squid system, which is uniquely suited for studies of microbial colonization. Finally, we discuss gnotobiotic mice as a model system that can be used to investigate mammalian gut interactions. We highlight where each of these models excels (Fig. 1) and describe limitations within each system to underscore the importance of selecting an appropriate model to address the scientific question at hand.

A4.3 Uncovering Molecular Mechanisms of Interactions Using *Bacillus* and *Streptomyces*

Among the simplest model systems for exploring microbial interactions are pairwise interactions between culturable bacteria. Importantly, these systems intrinsically offer high experimental control to study the molecular underpinnings of interactions that occur between and within microbial communities. As an example, coculture of the soil bacteria Bacillus subtilis and *Streptomyces* spp. demonstrates the power to dissect the molecular mechanisms of competition. Both B. subtilis and *Streptomyces* species are amenable to genetic manipulation, produce antibiotics and other secondary metabolites, and undergo multicellular development (e.g., biofilm formation, motility, and sporulation) on agar plates, providing macroscopic visualization

of interactions. Together, the ability to perform mutagenesis screens, generate targeted gene deletions and complements, extract secondary metabolites in isolation, and easily adjust medium and plating configurations to uncover new macroscopic phenotypes all contribute to this system's high level of experimental manipulability.

Pairwise interactions between *Bacillus* and *Streptomyces* demonstrate that secondary metabolites have multiple roles mediating competition (Fig. 2). For instance, *B. subtilis* produces the lipopeptide surfactin, which triggers its own biofilm formation and multicellular motility (5–7). In contrast, surfactin interferes with the aerial development and sporulation of many *Streptomyces* spp. (8). However, *Streptomyces* sp. strain Mg1 produces a secreted hydrolase that detoxifies surfactin and allows this bacterium to sporulate when cultured with B. subtilis (9). Similarly, *B. subtilis* produces bacillaene that interferes with prodigiosin pigment production in *Streptomyces coelicolor* and *Streptomyces lividans* (10, 11) and protects *B. subtilis* from lysis by linearmycins produced by strain Mg1 (12–14) (Fig. 2C). In addition to bacillaene, *B. subtilis* may protect itself from linearmycin-induced lysis by activating a linearmycin-induced, coupled signaling system and exporter that are necessary and sufficient for linearmycin resistance (12, 15). Finally, as an additional means to escape competition, subinhibitory concentrations of chloramphenicol and several other ribosome-targeting antibiotics induce directional sliding motility in *B. subtilis* away from *Streptomyces* (16) (Fig. 2D).

We highlight the above as examples of multifaceted interactions that can occur between one pair of microbes. Further, even by simply substituting one member of the pair, new interaction dynamics may emerge. For instance, recent work on interactions between *Streptomyces venezuelae* and *Saccharomyces cerevisiae* uncovered a new type of "exploration" motility in *S. venezuelae* induced by the production of volatile trimethylamine (17). However, it

is important to consider the artificial abstraction when microbes are transplanted into the laboratory. Compared to microbes in their natural environments, microbes in growth medium encounter atypical nutrients at inordinate concentrations and grow at unnaturally high cell densities. Consequently, microbes may produce extracellular products (e.g., antibiotics) at concentrations that elicit nonphysiological/hormetic responses in interacting partners (18, 19). Furthermore, the evolutionary implications from pairwise interactions are often unknown or unclear. Nevertheless, microbial coculture allows us to infer mechanisms that are impossible to uncover from sequencing studies alone. Therefore, to gain similar mechanistic insight into interactions occurring in communities, model systems where microbes can be isolated in pure culture and investigated in simplified pairwise interactions are invaluable.

A4.4 Colonization of the Light Organ by *Aliivibrio Fischeri* to Investigate Host-Microbe Interactions

The bacterium *Aliivibrio fischeri* (formerly *Vibrio fischeri*) specifically establishes a symbiosis within the light organ of newly hatched Hawaiian bobtail squid (*Euprymna scolopes*). This symbiosis has proven an excellent system to investigate colonization dynamics and specificity: though the ocean harbors an incredibly complex microbial community (>10^6 bacterial cells/ml), the relatively rare *A. fischeri* (<1 in 5,000 cells) specifically colonizes the light organ (20).

Specialized cilia and mucus recruit *A. fischeri* during early squid development. Bacteria within the mucus are chemotactically attracted toward pores and swim into light organ crypts (21). During the earliest stages of colonization, *A. fischeri* expresses a suite of genes under the "symbiotic colonization-sensor" RscS regulator (22, 23), which promotes polysaccharide production and biofilm formation (24–26) essential for colonization. The bacterially produced,

diaminopimelic acid (DAP) type peptidoglycan tracheal cytotoxin (TCT) and lipid A cause apoptosis of ciliated cells (20). The squid subsequently detoxifies TCT (27) and lipid A (28), followed by hemocyte infiltration and tissue regeneration to form the mature light organ (20). Further, squid nitric oxide (NO) signaling (29, 30) and detoxification (31) are tuned in response to colonization, modulating *A. fischeri* populations and excluding competitors from the light organ (20). When RscS is introduced into *A. fischeri* MJ11, a fish symbiont that naturally lacks RscS, the bacteria gain the ability to colonize E. scolopes (23), despite more than 400 unique genes in the laboratory squid strain ES114 compared to MJ11. Aside from biofilm formation and RscS-controlled responses, bacterial motility (20), type VI secretion systems (32), bacterial stress responses (33), other *A. fischeri* regulatory cascades (34), and host genetic factors (35) play key roles in colonization success.

An implicit and unique strength of the squid-*A. fischeri* light organ system is its simplicity, as one-host, one-microbe studies are experimentally manageable and yield ecologically relevant insights into the molecular mechanisms of this symbiosis. Historically, the majority of mechanistic research describing both host and microbe in the squid-*Aliivibrio* symbiosis has focused on a single strain, *A. fischeri* ES114. As such, assessing the extent to which the molecular insights of ES114 colonization apply to other *A. fischeri* strains remains an ongoing effort in this system. Notably, multiple strains of ecologically and phylogenetically distinct *A. fischeri* have been experimentally evolved within the squid host, selecting for alleles of the regulator *binK* that coordinate symbiosis traits and enhance colonization and growth within the light organ (36). Thus, to better understand how specificity relates to the diversity of both *A. fischeri* and *E. scolopes* that exists in nature, future studies are needed to address the impact of strain- and population-level diversity on colonization success and host-microbe

fidelity. Nevertheless, the many molecular interactions between one host species and one bacterial strain in this system, even when restricting focus to interactions surrounding colonization, make it a promising research area. Furthermore, whether the specialized physical, chemical, and genetic interactions between squid and *A. fischeri* during colonization have broader implications across different microbes and hosts is unknown. However, a newly emerging system involves the squid nidamental gland, which is situated next to the light organ and harbors a more complex community that consists of *Roseobacter*, *Flavobacteriales*, *Rhizobiales*, and *Verrucomicrobia* (37). We envision that comparison between these two adjacent organs within the same animal that recruit a different set of microbial symbionts from the same seawater environment will provide further insight into how host selection affects microbiome composition and function.

A4.5 Levels of Complexity in Germfree Mice

In humans, the gut microbiota is a complex community containing hundreds of species that impact a variety of health outcomes (38, 39). The microbiota is critical for normal development, as germfree animals possess immune, digestive, and behavioral differences compared to conventional counterparts (40). Germfree animals offer a platform for characterizing interactions with the host and defined communities of microbes (together known as gnotobiotics), ranging from monoassociations to complex communities. Arguably, monocolonized and germfree animals represent vast oversimplification. Defined synthetic communities simplify complex microbiotas while maintaining diversity, and the use of genome-sequenced strains facilitates multi-omics studies (41, 42). Further, using a simplified core microbiota with a genetically tractable strain of interest offers a compromise between creating a well-controlled experiment and not relying on monoassociation studies. For example, to

determine the role of the microbial conversion of choline to trimethylamine, mice were colonized with a simplified, six-member gut microbiota containing a single member that could metabolize choline or a mutant of the same strain that was unable to use choline. This approach demonstrates that choline-metabolizing bacteria compete with their hosts for choline and can exacerbate diet-induced metabolic disease in hosts and alter DNA methylation patterns in the brains of offspring (43). Notably, the choline utilization pathway is not taxonomically conserved, and it would be impossible to infer this phenotype from sequencing the 16S rRNA gene from gut communities (44).

To study entire communities, germfree mice can be colonized with complex communities, often from fecal samples. Donor communities can demonstrate a proof of principle of microbiota-mediated effects on a particular phenotype, such as linking the microbiota to obesity (45, 46). However, with increasing community complexity, more reproducibility issues arise. For instance, though donor communities reduce the artificial nature of gnotobiotics, rare strains may be stochastically lost in the transplanted community. When human fecal microbiota are transplanted into germfree mice, 10 to 30% of operational taxonomic units fail to colonize the mouse (47). Strains present at 0.15% of the community can impact phenotypes like choline conversion to trimethylamine (44). Alternatively, using donor microbiota derived from the same species as the germfree animal can be more appropriate for certain ecological questions and better retain members (48). Reproducibility is also an issue for studying some emergent phenotypes of complex communities, as maintenance of certain members may depend on diet or even water pH (49), and social, coprophagous animals like mice may necessitate cages as biological units of replication, rather than individuals (50, 51). Though reproducibility issues also arise in simplified communities, troubleshooting whether small changes in abiotic or biotic

factors influence phenotypes is more challenging in complex communities and could be limiting in a mouse system with a relatively slow generation time and ethical constraints on animal usage.

Overall, gnotobiotic animals provide an approach to interrogate the role of complex microbiota in emergent phenotypes of interest by reducing the complexity to controllable independent variables (e.g., a single bacterial strain or product). Experimenting with multiple levels of community complexity applies to germfree hosts beyond mice (e.g., *Arabidopsis*, *Danio*, and *Drosophila*), but specific mechanisms may differ. For example, facultatively anaerobic pathogens exploiting inflammation-associated oxidation in the typically anaerobic mouse gut would not be readily apparent in aerobic *Drosophila* guts (52, 53). Further, although mice are often sought as medically relevant models, the ease of producing large numbers of gnotobiotic animals and availability of tools in other models, such as imaging in translucent zebrafish, can reveal alternative mechanisms for microbial proteins mediating mutualism that may have remained obscure in a mouse model (54). Ultimately, shared insights from different models support broad ecological principles of microbiome interactions.

A4.6 Conclusion

By leveraging the unique features of experimental microbiome systems, important and outstanding questions can be addressed (Fig. 1). Chief among these questions is understanding how interactions between microbes and hosts influence behavior and health and how communities respond to perturbations, such as invasion or abiotic stresses. Although it is well understood that microbiomes influence the health of hosts and macroscopic ecosystems, the specific molecular mechanisms remain elusive. For instance, what interactions differentiate “healthy” and “dysbiotic” microbial communities are often unresolved. Further, communities can exhibit emergent phenotypes that are not seen when members are grown in isolation, such as

catabolism of recalcitrant materials (55, 56), biofilm formation (57), or antibiotic production (58–60).

As microbiome research continues, new frameworks for characterizing the interactions that occur within microbial communities will emerge from novel systems spanning the spectra of complexity and tractability and developments enabling established systems to address new questions. As examples, two particular systems that we are especially interested in are the cheese rind microbial community and the gardens of fungus-growing ants. The cheese rind microbial community is an emerging system particularly suitable for characterization of multipartite interactions and simulating ecological phenomena through control of abiotic factors (61–64), yet the unclear evolutionary relationships between members may limit its applicability to coevolved, natural systems. In contrast, because the microbial symbionts of fungus-growing ants provide a coevolutionary framework from which to investigate microbial population dynamics (65, 66), nutrient flow (67), host-pathogen interactions (68–70), and defensive symbiosis (71), further characterizations of these microbiomes may provide broader implications for other natural systems (59, 60).

In conclusion, delineating community states that contribute to emergent properties and complex interactions will require experimental models, and the ideal balance between a model's complexity, ease of manipulation, and overall biological relevance will depend upon the scientific questions posed.

A4.7 Acknowledgments

We apologize to those whose work we have not discussed due to the space constraints.

We thank Mark Mandel for critical appraisal of the manuscript.

This project was supported through National Institutes of Health (NIH) U19 Al109673 and NIH U19 TW009872. Additional support was provided to M.G.C. by NIH T32 GM008505, to R.M.S.

by an NLM training grant to the Computation and Informatics in Biology and Medicine Training Program (NLM 5T15LM007359), and to J.R.B. by NIH T32 AI55397.

A4.8 Figures

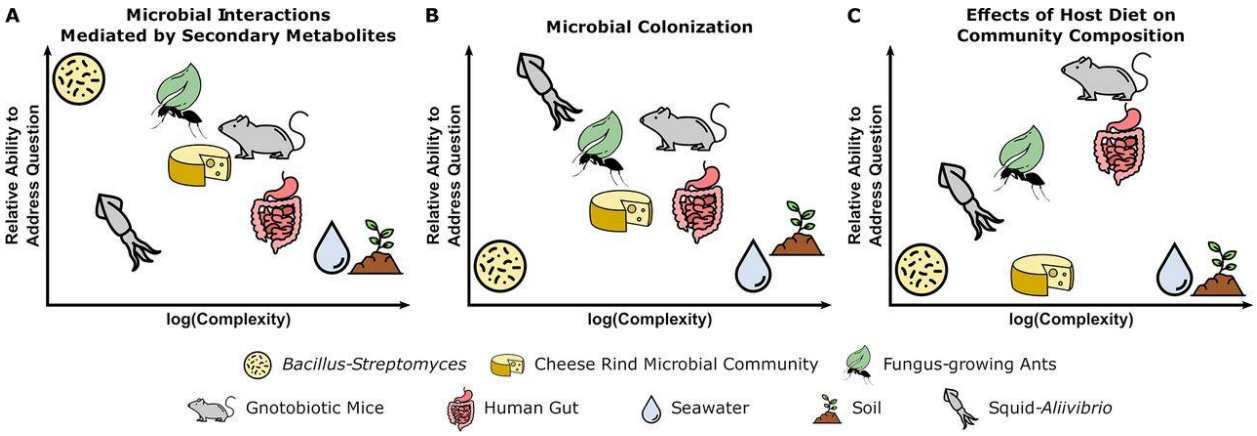


Figure 1. Tradeoffs between experimental questions and complexity of microbiome systems. Each microbiome system is suited to address different types of questions based on the culturability of microbes, genetic tractability of microbes and host (where relevant), ability to maintain system in laboratory setting, and ability to make host/environment germfree. Three different systems are shown in this figure as examples. (A) Pairwise interactions between B. subtilis and Streptomyces spp. are well-suited for characterizing the functions of secondary metabolites in microbial interactions. (B) The symbiosis between bobtail squid and A. fischeri is fundamental to understanding host and microbial factors that influence colonization. (C) The use of gnotobiotic mice is crucial for making links between host diet and the effects on specific microbial taxa in a community (see the text for specific details). Specific original image credit from the Noun Project (https://thenounproject.com/): Fertile Soil by Ben Davis; Droplet by Focus; Mouse by Iconic; Cheese Wheel by Anniken & Andreas; Bacteria by Arthur Shlain; Squid by Artem Kovyazin; ant by Yugudesign; leaf by Saeful Muslim; all used and modified under the Creative Commons License, Attribution 3.0.

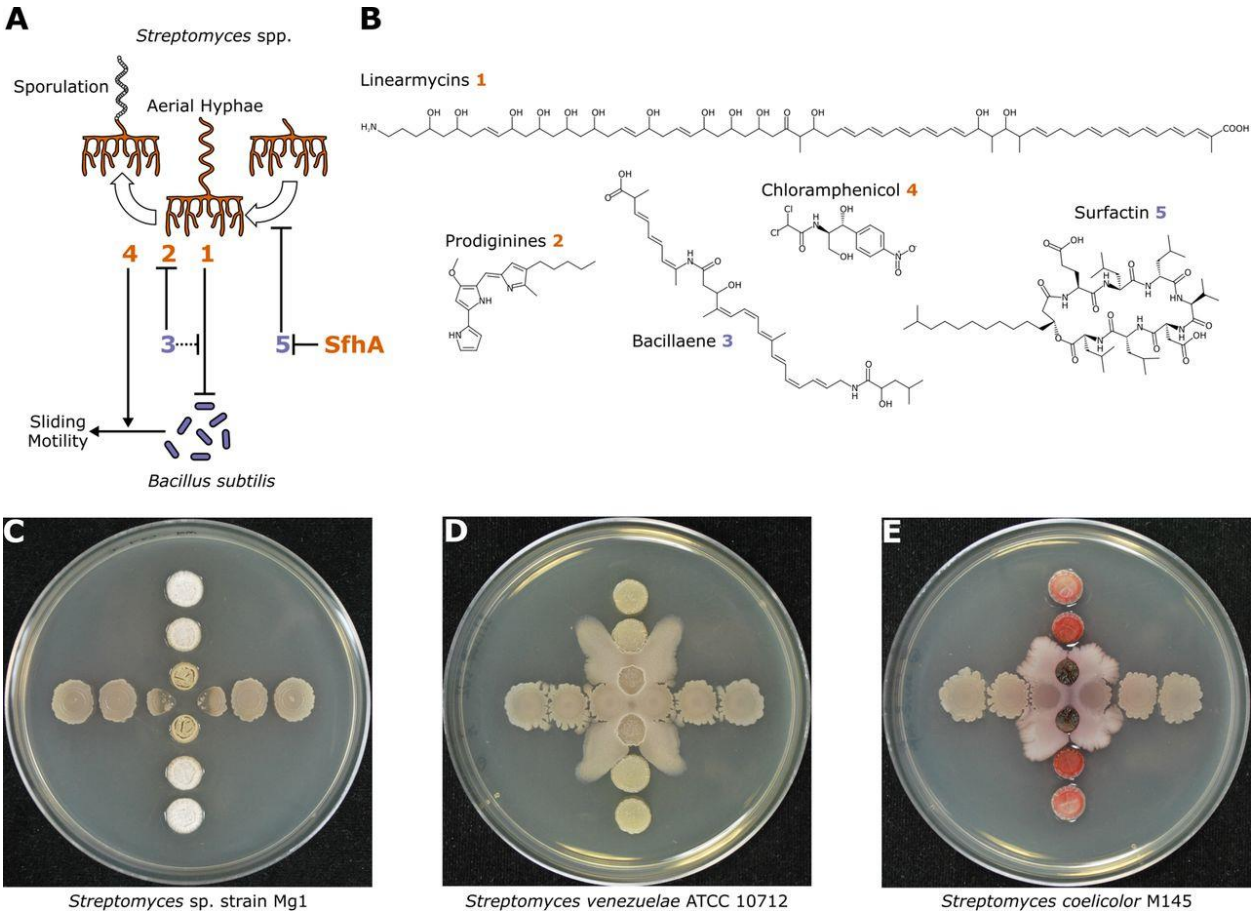


Figure 2. Secondary metabolites mediate interactions between B. subtilis and *Streptomyces* spp. (A) Summary schematic of interactions between B. subtilis and *Streptomyces* spp. The secondary metabolites produced by B. subtilis and *Streptomyces* spp. are represented by the purple and orange numbers, respectively, and the chemical structures are shown in panel B. SfhA refers to surfactin hydrolase produced by *Streptomyces* sp. strain Mg1 that specifically hydrolyzes the ester linkage in surfactin (compound 5). (C to E) *Streptomyces* spp. (vertical) and B. subtilis (horizontal) spotted in a perpendicular pattern on agar plates. (C) B. subtilis colonies proximal to *Streptomyces* sp. strain Mg1 colonies are lysed by linearmycins (compound 1). (Republished from *Frontiers in Microbiology* [3].) (D) Subinhibitory concentrations of chloramphenicol (compound 4) produced by Streptomyces venezuelae induce sliding motility of proximal B. subtilis colonies. (E) Production of the red pigment prodiginine (compound 2) is strongly induced in Streptomyces coelicolor colonies proximal to sliding B. subtilis colonies, which do not produce bacillaene (compound 3). (Images in panels D and E courtesy of Yongjin Liu and Paul Straight, reproduced with permission.)

A4.9 References

1. Descartes R. 1662. De homine figuris et latinatate donatus a Florentio Schuyl. P. Leffen & F. Moyardum, Leyden, The Netherlands.

2. Tecon R, Mitri S, Ciccarese D, Or D, van der Meer JR, Johnson DR. 2019. Bridging the holistic-reductionist divide in microbial ecology mSystems 4:e00265-18. doi:10.1128/mSystems.00265-18.

3. Stubbendieck RM, Vargas-Bautista C, Straight PD. 2016. Bacterial communities: interactions to scale. Front Microbiol 7:1234. doi:10.3389/fmicb.2016.01234.

4. Gould AL, Zhang V, Lamberti L, Jones EW, Obadia B, Korasidis N, Gavryushkin A, Carlson JM, Beerenwinkel N, Ludington WB. 2018. Microbiome interactions shape host fitness. Proc Natl Acad Sci U S A 115:E11951–E11960. doi:10.1073/pnas.1809349115.

5. Kearns DB, Losick R. 2003. Swarming motility in undomesticated Bacillus subtilis. Mol Microbiol 49:581–590.

6. Lopez D, Fischbach MA, Chu F, Losick R, Kolter R. 2009. Structurally diverse natural products that cause potassium leakage trigger multicellularity in Bacillus subtilis. Proc Natl Acad Sci U S A 106:280–285. doi:10.1073/pnas.0810940106.

7. van Gestel J, Vlamakis H, Kolter R. 2015. From cell differentiation to cell collectives: Bacillus subtilis uses division of labor to migrate. PLoS Biol 13:e1002141. doi:10.1371/journal.pbio.1002141.

8. Straight PD, Willey JM, Kolter R. 2006. Interactions between Streptomyces coelicolor and Bacillus subtilis: role of surfactants in raising aerial structures. J Bacteriol 188:4918–4925. doi:10.1128/JB.00162-06.

9. Hoefler BC, Gorzelnik KV, Yang JY, Hendricks N, Dorrestein PC, Straight PD. 2012. Enzymatic resistance to the lipopeptide surfactin as identified through imaging mass spectrometry of bacterial competition. Proc Natl Acad Sci U S A 109:13082–13087. doi:10.1073/pnas.1205586109.

10. Straight PD, Fischbach MA, Walsh CT, Rudner DZ, Kolter R. 2007. A singular enzymatic megacomplex from Bacillus subtilis. Proc Natl Acad Sci U S A 104:305–310. doi:10.1073/pnas.0609073103.

11. Vargas-Bautista C, Rahlwes K, Straight P. 2014. Bacterial competition reveals differential regulation of the pks genes by Bacillus subtilis. J Bacteriol 196:717–728. doi:10.1128/JB.01022-13.

12. Stubbendieck RM, Straight PD. 2015. Escape from lethal bacterial competition through coupled activation of antibiotic resistance and a mobilized subpopulation. PLoS Genet 11:e1005722. doi:10.1371/journal.pgen.1005722.

13. Stubbendieck RM, Brock DJ, Pellois J-P, Gill JJ, Straight PD. 2018. Linearmycins are lytic membrane-targeting antibiotics. J Antibiot 71:372. doi:10.1038/s41429-017-0005-z.

14. Hoefler BC, Stubbendieck RM, Josyula NK, Moisan SM, Schulze EM, Straight PD. 2017. A link between linearmycin biosynthesis and extracellular vesicle genesis connects specialized

metabolism and bacterial membrane physiology. Cell Chem Biol 24:1238–1249.e7. doi:10.1016/j.chembiol.2017.08.008.

15. Stubbendieck RM, Straight PD. 2017. Linearmycins activate a two-component signaling system involved in bacterial competition and biofilm morphology. J Bacteriol 199:e00186-17. doi:10.1128/JB.00186-17.

16. Liu Y, Kyle S, Straight PD. 2018. Antibiotic stimulation of a Bacillus subtilis migratory response. mSphere 3:e00586-17. doi:10.1128/mSphere.00586-17.

17. Jones SE, Ho L, Rees CA, Hill JE, Nodwell JR, Elliot MA. 2017. Streptomyces exploration is triggered by fungal interactions and volatile signals. Elife 6:1–21. doi:10.7554/eLife.21738.

18. Davies J. 2006. Are antibiotics naturally antibiotics? J Ind Microbiol Biotechnol 33:496–499. doi:10.1007/s10295-006-0112-5.

19. Davies J, Spiegelman GB, Yim G. 2006. The world of subinhibitory antibiotic concentrations. Curr Opin Microbiol 9:445–453. doi:10.1016/j.mib.2006.08.006.

20. Mandel MJ, Dunn AK. 2016. Impact and influence of the natural Vibrio-squid symbiosis in understanding bacterial-animal interactions. Front Microbiol 7:1982. doi:10.3389/fmicb.2016.01982.

21. Nyholm SV, Stabb EV, Ruby EG, McFall-Ngai MJ. 2000. Establishment of an animal-bacterial association: recruiting symbiotic vibrios from the environment. Proc Natl Acad Sci U S A 97:10231–10235. doi:10.1073/pnas.97.18.10231.

22. Yip ES, Grublesky BT, Hussa EA, Visick KL. 2005. A novel, conserved cluster of genes promotes symbiotic colonization and σ54-dependent biofilm formation by Vibrio fischeri. Mol Microbiol 57:1485–1498. doi:10.1111/j.1365-2958.2005.04784.x.

23. Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG. 2009. A single regulatory gene is sufficient to alter bacterial host range. Nature 458:215–218. doi:10.1038/nature07660.

24. Singh P, Brooks JF, Ray VA, Mandel MJ, Visick KL. 2015. CysK plays a role in biofilm formation and colonization by Vibrio fischeri. Appl Environ Microbiol 81:5223–5234. doi:10.1128/AEM.00157-15.

25. Visick KL, Skoufos LM. 2001. Two-component sensor required for normal symbiotic colonization of Euprymna scolopes by Vibrio fischeri. J Bacteriol 183:835–842. doi:10.1128/JB.183.3.835-842.2001.

26. Yip ES, Geszvain K, DeLoney-Marino CR, Visick KL. 2006. The symbiosis regulator rscS controls the syp gene locus, biofilm formation and symbiotic aggregation by Vibrio fischeri. Mol Microbiol 62:1586–1600. doi:10.1111/j.1365-2958.2006.05475.x.

27. Troll JV, Bent EH, Pacquette N, Wier AM, Goldman WE, Silverman N, McFall-Ngai MJ. 2010. Taming the symbiont for coexistence: a host PGRP neutralizes a bacterial symbiont toxin. Environ Microbiol 12:2190–2203. doi:10.1111/j.1462-2920.2009.02121.x.

28. Rader BA, Kremer N, Apicella MA, Goldman WE, McFall-Ngai MJ. 2012. Modulation of symbiont lipid A signaling by host alkaline. Mar Biol 3:e00093-12. doi:10.1128/mBio.00093-12.

29. Wang Y, Dufour YS, Carlson HK, Donohue TJ, Marletta MA, Ruby EG. 2010. H-NOX-mediated nitric oxide sensing modulates symbiotic colonization by Vibrio fischeri. Proc Natl Acad Sci U S A 107:8375–8380. doi:10.1073/pnas.1003571107.

30. Wang Y, Ruby EG. 2011. The roles of NO in microbial symbioses. Cell Microbiol 13:518–526. doi:10.1111/j.1462-5822.2011.01576.x.

31. Tucker NP, Le Brun NE, Dixon R, Hutchings MI. 2010. There's NO stopping NsrR, a global regulator of the bacterial NO stress response. Trends Microbiol 18:149–156. doi:10.1016/j.tim.2009.12.009.

32. Speare L, Cecere AG, Guckes KR, Smith S, Wollenberg MS, Mandel MJ, Miyashiro T, Septer AN. 2018. Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host. Proc Natl Acad Sci U S A 115:E8528–E8537. doi:10.1073/pnas.1808302115.

33. Brooks JF, Gyllborg MC, Cronin DC, Quillin SJ, Mallama CA, Foxall R, Whistler C, Goodman AL, Mandel MJ. 2014. Global discovery of colonization determinants in the squid symbiont Vibrio fischeri. Proc Natl Acad Sci U S A 111:17284–17289. doi:10.1073/pnas.1415957111.

34. Bose JL, Rosenberg CS, Stabb EV. 2008. Effects of luxCDABEG induction in Vibrio fischeri: enhancement of symbiotic colonization and conditional attenuation of growth in culture. Arch Microbiol 190:169–183. doi:10.1007/s00203-008-0387-1.

35. McFall-Ngai M. 2014. Divining the essence of symbiosis: insights from the squid-Vibrio model. PLoS Biol 12:e1001783. doi:10.1371/journal.pbio.1001783.

36. Sabrina Pankey M, Foxall RL, Ster IM, Perry LA, Schuster BM, Donner RA, Coyle M, Cooper VS, Whistler CA. 2017. Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. Elife 6:e24414. doi:10.7554/eLife.24414.

37. Collins AJ, Fullmer MS, Gogarten JP, Nyholm SV. 2015. Comparative genomics of Roseobacter clade bacteria isolated from the accessory nidamental gland of Euprymna scolopes. Front Microbiol 6:123. doi:10.3389/fmicb.2015.00123.

38. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, et al. . 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464:59–65. doi:10.1038/nature08821.

39. Lynch SV, Pedersen O. 2016. The human intestinal microbiome in health and disease. N Engl J Med 375:2369–2379. doi:10.1056/NEJMra1600266.

40. Wostmann BS. 1996. Germfree and gnotobiotic animal models: background and applications, 1st ed CRC Press, Boca Raton, FL.

41. Bratburd JR, Keller C, Vivas E, Gemperline E, Li L, Rey FE, Currie CR. 2018. Gut microbial and metabolic responses to Salmonella enterica serovar Typhimurium and Candida albicans. mBio 9:e02032-18. doi:10.1128/mBio.02032-18.

42. Li H, Limenitakis JP, Fuhrer T, Geuking MB, Lawson MA, Wyss M, Brugiroux S, Keller I, Macpherson JA, Rupp S, Stolp B, Stein JV, Stecher B, Sauer U, McCoy KD, Macpherson AJ. 2015. The outer mucus layer hosts a distinct intestinal microbial niche. Nat Commun 6:8292. doi:10.1038/ncomms9292.

43. Romano KA, Martinez-del Campo A, Kasahara K, Chittim CL, Vivas EI, Amador-Noguez D, Balskus EP, Rey FE. 2017. Metabolic, epigenetic, and transgenerational effects of gut bacterial choline consumption. Cell Host Microbe 22:279–290.e7. doi:10.1016/j.chom.2017.07.021.

44. Romano KA, Vivas EI, Amador-Noguez D, Rey FE. 2015. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. mBio 6:e02481-14. doi:10.1128/mBio.02481-14.

45. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444:1027–1031. doi:10.1038/nature05414.

46. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR, Muehlbauer MJ, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle BJ, Martini MC, Ursell LK, Clemente JC, Van Treuren W, Walters WA, Knight R, Newgard CB, Heath AC, Gordon JI. 2013. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science 341:1241214. doi:10.1126/science.1241214.

47. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. 2009. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 1:6ra14. doi:10.1126/scitranslmed.3000322.

48. Arrieta M-C, Walter J, Finlay BB. 2016. Human microbiota-associated mice: a model with challenges. Cell Host Microbe 19:575–578. doi:10.1016/j.chom.2016.04.014.

49. Barnett JA, Gibson DL. 2019. H2Oh No! The importance of reporting your water source in your in vivo microbiome studies. Gut Microbes 10:261–269. doi:10.1080/19490976.2018.1539599.

50. Hugenholtz F, de Vos WM. 2018. Mouse models for human intestinal microbiota research: a critical evaluation. Cell Mol Life Sci 75:149–160. doi:10.1007/s00018-017-2693-8.

51. Lazic SE, Clarke-Williams CJ, Munafò MR. 2018. What exactly is ‘N’ in cell culture and animal experiments? PLoS Biol 16:e2005282. doi:10.1371/journal.pbio.2005282.

52. Cox CR, Gilmore MS. 2007. Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcus faecalis pathogenesis. Infect Immun 75:1565–1576. doi:10.1128/IAI.01496-06.

53. Rivera-Chávez F, Zhang LF, Faber F, Lopez CA, Byndloss MX, Olsan EE, Xu G, Velazquez EM, Lebrilla CB, Winter SE, Bäumler AJ. 2016. Depletion of butyrate-producing clostridia from the gut microbiota drives an aerobic luminal expansion of Salmonella. Cell Host Microbe 19:443–454. doi:10.1016/j.chom.2016.03.004.

54. Rolig AS, Sweeney EG, Kaye LE, DeSantis MD, Perkins A, Banse AV, Hamilton MK, Guillemin K. 2018. A bacterial immunomodulatory protein with lipocalin-like domains facilitates host–bacteria mutualism in larval zebrafish. Elife 7:e37172. doi:10.7554/eLife.37172.

55. Lewin GR, Johnson AL, Soto RDM, Perry K, Book AJ, Horn HA, Pinto-Tomás AA, Currie CR. 2016. Cellulose-enriched microbial communities from leaf-cutter ant (Atta colombica) refuse dumps vary in taxonomic composition and degradation ability. PLoS One 11:e0151840. doi:10.1371/journal.pone.0151840.

56. Carlos C, Fan H, Currie CR. 2018. Substrate shift reveals roles for members of bacterial consortia in degradation of plant cell wall polymers. Front Microbiol 9:364. doi:10.3389/fmicb.2018.00364.

57. Lozano GL, Bravo JI, Garavito Diago MF, Park HB, Hurley A, Peterson SB, Stabb EV, Crawford JM, Broderick NA, Handelsman J. 2019. Introducing THOR, a model microbiome for genetic dissection of community behavior. mBio 10:e02846-18. doi:10.1128/mBio.02846-18.

58. Adnani N, Chevrette MG, Adibhatla SN, Zhang F, Yu Q, Braun DR, Nelson J, Simpkins SW, McDonald BR, Myers CL, Piotrowski JS, Thompson CJ, Currie CR, Li L, Rajski SR, Bugni TS. 2017. Coculture of marine invertebrate-associated bacteria and interdisciplinary technologies enable biosynthesis and discovery of a new antibiotic, keyicin. ACS Chem Biol 12:3093–3102. doi:10.1021/acschembio.7b00688.

59. Chevrette MG, Currie CR. 2019. Emerging evolutionary paradigms in antibiotic discovery. J Ind Microbiol Biotechnol 46:257–271. doi:10.1007/s10295-018-2085-6.

60. Chevrette MG, Carlson CM, Ortega HE, Thomas C, Ananiev GE, Barns KJ, Book AJ, Cagnazzo J, Carlos C, Flanigan W, Grubbs KJ, Horn HA, Hoffmann FM, Klassen JL, Knack JJ, Lewin GR, McDonald BR, Muller L, Melo WGP, Pinto-Tomás AA, Schmitz A, Wendt-Pienkowski E, Wildman S, Zhao M, Zhang F, Bugni TS, Andes DR, Pupo MT, Currie CR. 2019. The antimicrobial potential of Streptomyces from insect microbiomes. Nat Commun 10:516. doi:10.1038/s41467-019-08438-0.

61. Wolfe BE, Button JE, Santarelli M, Dutton RJ. 2014. Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. Cell 158:422–433. doi:10.1016/j.cell.2014.05.041.

62. Monnet C, Landaud S, Bonnarme P, Swennen D. 2015. Growth and adaptation of microorganisms on the cheese surface. FEMS Microbiol Lett 362:1–9. doi:10.1093/femsle/fnu025.

63. Zhang Y, Kastman EK, Guasto JS, Wolfe BE. 2018. Fungal networks shape dynamics of bacterial dispersal and community assembly in cheese rind microbiomes. Nat Commun 9:336. doi:10.1038/s41467-017-02522-z.

64. Kastman EK, Kamelamela N, Norville JW, Cosetta CM, Dutton RJ, Wolfe BE. 2016. Biotic interactions shape the ecological distributions of Staphylococcus species. mBio 7:e01157-16. doi:10.1128/mBio.01157-16.

65. McDonald BR, Chevrette MG, Klassen JL, Horn HA, Caldera EJ, Wendt-Pienkowski E, Cafaro MJ, Ruzzini AC, Arnam EV, Weinstock GM, Gerardo NM, Poulsen M, Suen G, Clardy J, Currie CR. 2019. Biogeography and microscale diversity shape the biosynthetic potential of fungus-growing ant-associated Pseudonocardia. bioRxiv doi:10.1101/545640. [CrossRef]

66. Caldera EJ, Currie CR. 2012. The population structure of antibiotic-producing bacterial symbionts of Apterostigma dentigerum ants: impacts of coevolution and multipartite symbiosis. Am Nat 180:604–617. doi:10.1086/667886.

67. Steffan SA, Chikaraishi Y, Currie CR, Horn H, Gaines-Day HR, Pauli JN, Zalapa JE, Ohkouchi N. 2015. Microbes are trophic analogs of animals. Proc Natl Acad Sci U S A 112:15119–15124. doi:10.1073/pnas.1508782112.

68. Currie CR, Mueller UG, Malloch D. 1999. The agricultural pathology of ant fungus gardens. Proc Natl Acad Sci U S A 96:7998–8002. doi:10.1073/pnas.96.14.7998.

69. Birnbaum SSL, Gerardo NM. 2016. Patterns of specificity of the pathogen Escovopsis across the fungus-growing ant symbiosis. Am Nat 188:52–65. doi:10.1086/686911.

70. Gerardo NM, Mueller UG, Currie CR. 2006. Complex host-pathogen coevolution in the Apterostigma fungus-growing ant-microbe symbiosis. BMC Evol Biol 6:88. doi:10.1186/1471-2148-6-88.

71. Currie CR, Scott JA, Summerbell RC, Malloch D. 1999. Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. Nature 398:701–704. doi:10.1038/19519. (Erratum, 423:461, 2003.) doi:10.1038/19519.