6-2-2023

# Investigating Eco-evolutionary Interactions between Hosts and Members of Their Gut Microbiota

Logan A. Sauers
*Illinois State University*, lsauers@ilstu.edu

# INVESTIGATING ECO-EVOLUTIONARY INTERACTIONS BETWEEN HOSTS AND MEMBERS OF THEIR GUT MICROBIOTA

LOGAN A SAUERS

169 Pages

Evolutionary and ecological interactions between hosts and their associated microbial communities, their microbiota, and between members of these communities are vital to understand. Host-associated microbial communities are widespread across diverse host taxa and the hosts of these communities receive a variety of well-documented benefits from them. Despite the importance of an appreciation of eco-evolutionary dynamics for colonization outcomes and the benefits these microbial communities provide to their hosts, our current knowledge in this area remains incomplete. For example, we do not know the full extent of coevolution and specific relationships between hosts and microbes, and between the microbes themselves, across host taxa. Questions remain about how host taxonomy, ecology and physiology, and other microbes present within the microbiota influence microbial community membership and function, host and microbe evolution, and specificity in colonization of hosts. I present several studies that aim to shed further light on these topics from an eco-evolutionary perspective utilizing insect pollinators and their gut microbial communities, with a particular focus on bumble bees and their microbiota.

KEYWORDS: Host microbiota, colonization, eco-evolutionary dynamics, *Bombus impatiens,* coevolution, pollinator health

# INVESTIGATING ECO-EVOLUTIONARY INTERACTIONS BETWEEN HOSTS AND MEMBERS OF THEIR GUT MICROBIOTA

LOGAN A SAUERS

A Dissertation Submitted in Partial
Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

School of Biological Sciences

ILLINOIS STATE UNIVERSITY

2023

# INVESTIGATING ECO-EVOLUTIONARY INTERACTIONS BETWEEN HOSTS AND MEMBERS OF THEIR GUT MICROBIOTA

LOGAN A SAUERS

COMMITTEE MEMBERS:

Ben Sadd, Chair

Fredrick Inglis

Steve Juliano

Quinn McFrederick

Nathan Mortimer

ACKNOWLEDGMENTS

This work is dedicated to my late grandmother, Harriet Quandt, she bravely stood in the face of my greatest fear. Her constant reminders of how proud she was of the work I was doing, and how proud my grandfather would have been was able to get me through even the most difficult times during my studies.

Thanks especially to my family, both old and new. To the old, specifically my mom and dad, you the support you provided to me was not only greatly appreciated but serves as a bright and brilliant example of how I hope to support my family in the future. To my mom, your insistence that I pursue my dreams and a career that would make my life fulfilling provided the spark necessary for finding my passion in science. To my dad, I still remember all the times you sacrificed your own time and happiness to ensure that those around you were happy and taken care of, and this served as an essential component in my own willingness to help others, leading me to many fruitful endeavors and collaborations I would not otherwise have. Additionally, your perseverance and success in your fight with cancer and chemotherapy showed me that no matter how difficult things became it was always possible to get through the suck.

Moving on to the family that is new. Nicole, I would not have survived graduate school without you. You are an amazing individual, and your ability to put up with my little brain farts and my sidetracked mind is impressive. You do so much to make things easier on me at home and in the lab, and I will forever appreciate what you have done and sacrificed to help me finish this dissertation. I look forward to the things we can do and the places we can see after I graduate, but I know it will be great because I will have you and Finn by my side.

To my friends, Tony and Matt, your over 10 years of constant support will never be forgotten. The group chat between the three of us has been a blessing, even if it was a curse for my productivity some days. To Kate, your friendship was a stone during tough times, and I will forever cherish our coffee

chats and boardgame nights. To Ian, we may not have been able to spend as much time together as it would've initially seemed, but I am grateful for the many boardgame nights, and I hope you have found your own spark. To Zoey, I am grateful for the time we were able to spend together during our Master's studies and I am thankful we have stayed in touch, it is bittersweet seeing how old Maggie has gotten.

To my lab members, specifically Austin, I am forever thankful we have cultivated such an open and welcoming lab environment. I hope that we can continue working together in the future. To Bryan, your effort in lab was next to none, your perseverance through several failed projects and your constant happy attitude were both examples of dedication. Working with Ben and mentoring you was an experience I will forever be grateful for. To the many other students I had the opportunity to mentor, Earon Grinage, Amelia Mwilambwe, and Teni Shosanya, Mickey Williams, I hope, and know, each of you is able to find your own spark and success.

Finally, to my collaborators, Jilian Li, Laura Avila, and the members of the Wren lab. I am thankful for the fruitful cooperation and papers that have come from our work. It has been a pleasure working with you all, and these collaborations have really allowed me to develop skills I would not otherwise have had the opportunity to learn. To my greatest collaborator these past 7 years, Ben Sadd, your mentoring, help, and kindness will never be forgotten. I would never have had the chance to be the scientist I am today without your help, and I wish you, Cristina, and the SaddLab the best going forward.

L.A.S

CONTENTS


| | Page |
|---|---|
| ACKNOWLEDGMENTS | i |
| TABLES | vii |
| FIGURES | viii |
| CHAPTER I: INTRODUCTION TO AN ECOLOGICAL AND EVOLUTIONARY CONTEXT FOR HOST GUT MICROBIAL COMMUNITIES | 1 |
| Background | 1 |
| Bidirectionality of host-microbiota interactions | 2 |
| Microbial community interactions determining membership and function | 4 |
| Bumble bees as a model system for understanding host gut microbial communities | 10 |
| References | 15 |
| CHAPTER II: THE DIVERGENCE AND CONVERGENCE OF THE GUT MICROBIOMES OF WILD INSECT POLLINATORS | 32 |
| Abstract | 32 |
| Introduction | 33 |
| Results | 36 |
| Identifying colonization of bacterial genera within hosts | 37 |
| Alpha and beta diversity of the pollinator microbiomes | 38 |
| Investigating association between microbiota structure and host genetic distance | 39 |
| Metagenome Assembled Genomes (MAGs) | 40 |
| Phylogenetics and comparative genomics | 41 |
| Metabolic Reconstructions | 42 |
| Discussion | 43 |
| Conclusion | 50 |

Materials and Methods 51

Sample collection and processing 51

Amplicon and metagenomic sequencing 52

Bioinformatic analysis 53

Microbial Community Analysis 54

Metagenomic sequences analysis 55

References 56

CHAPTER III: EXAMINING THE EXTENT OF COLONIZATION SPECIFICITY BETWEEN BUMBLE SPECIES AND A CORE GUT MICROBIOTA MEMBER. 78

Abstract 78

Introduction 78

Materials and Methods 80

Standard bee rearing 80

Snodgrassella alvi strain culturing, identification, and inoculation preparation 81

Rearing and inoculation of germ-free bees 82

qPCR quantification of colonization 83

Statistical analysis 84

Results 85

Discussion 86

References 89

CHAPTER IV: IMMUNE STIMULATION INDUCED CHANGES TO THE NATIVE GUT MICROBIOTA OF BUMBLE BEES 98

Abstract 98

Introduction 98

Materials and Methods 101

Overall experimental design 101

Bumble bee maintenance and microbiota inoculation 102

Immune stimulation treatments of microbiota analysis bees 102

Confirmation of immune stimulation treatments 103

Targeted qPCR 104

Bioinformatics analysis 105

Statistical analysis 105

Results 107

Antibacterial activity following immune stimulation treatment 107

Microbiome analysis and diversity indices 107

Quantification of specific microbiota members 108

Network analysis-based community associations 108

Discussion 109

References 112

CHAPTER V: GUT MICROBIOME AND HOST METABOLOME CHANGES ASSOCIATED WITH DIFFERENT BUMBLE BEE QUEEN (*BOMBUS LANTSCHOUENSIS*) LIFE STAGES 128

Abstract 128

Introduction 128

Materials and Methods 130

Sample collection 130

Queen gut DNA extraction 131

Metagenomic sequencing and data processing 131

Queen hemolymph extraction 132

Statistical analysis 133

Results 134

Profiling of the queen gut microbiota 134

Hemolymph metabolome differences across queen status 135

Discussion 136

References 139

CHAPTER VI: INVESTIGATING THE ROLE OF INTERMICROBIAL INTERACTIONS IN DETERMINING HOST HEALTH 151

Abstract 151

Introduction 151

Materials and Methods 153

Experimental Design 153

Bacterial Culturing 154

Germ-free Procedure, Treatment Feeding and Health Measures 155

Statistical Analyses 155

Results 156

Discussion 156

References 159

CHAPTER VII: CONCLUSIONS 167

TABLES

Table Page

1. Metagenome assembled genome bins from the bee, fly, and wasp microbiota metagenome sequencing. Completeness and contamination were calculated using the CheckM plugin for MetaWrap. Predicted open reading frames were determined with Prodigal and annotation was done by matching against the KEGG, COG, and tRNA databases with Anvio. 72

2. Information on the strains used in this study and the bee species they were isolated from. 93

3. Factors and relevant statistics for the final models of Snodgrassella alvi colonization. Significant terms ($p<0.05$) are bolded. 94

4. Model results from fitted negative binomial models on adjusted absolute amplicon values. Significant effects are bolded. 122

5. Metagenome assembled genome information. 145

6. Model results from the best fitting models for each dataset. Significant effects are bolded. 163

FIGURES

Figure Page

1. Per host genera relative abundance of microbial families with at least 1% abundance in the data set (>296,547 reads). 73

2. Observed species richness and Shannon alpha diversity index of gut communities for each host genus. Bars represent the medians, and the boxes show the interquartile ranges. Whiskers are the upper and lower values, with outliers shown as individual data points beyond these. 74

3. Ordination plot of the Bray-Curtis beta diversity distance among individual host microbiomes. Host genera are represented by shape and color combinations. The inset represents the same but with the genera *Apis* and *Bombus* removed. 75

4. *Orbaceae* family phylogeny from selected core genes from the fly-associated metagenome assembled genome assigned as *Gilliamella sp*. and existing genomes. Coloring shows different species while the branch length shows amino acid substitution rates. *Pseudomonas aeruginosa* is used as an outgroup. 76

5. Fly-associated *Candidatus Gilliamella eristali* reconstructed metabolic pathways in comparison with other *Gilliamella* species. Solid black lines are genes/pathways shared by *Candidatus G. eristali* and other *Gilliamella* species. Dotted black lines are genes/pathways lacking across *Gilliamella* species. Dotted red lines are genes/pathways found in bee-associated *Gilliamella* but not *Candidatus G. eristali*. Blue solid lines are genes/pathways found in *Candidatus G. eristali* but no in bee-associated *Gilliamella* species. 77

6. *Snodgrassella* colonization abundance for *Bombus griseocollis* and *Bombus impatiens* recipients inoculated with honeybee isolated strains of *Snodgrassella* or strains isolated from the conspecific host of the respective recipient. Means and standard errors were estimated from truncated negative binomial models. 95

7. Colonization abundance of *Snodgrassella* in *Bombus griseocollis* and *Bombus impatiens* recipients of strains isolated from a conspecific source to the recipient or other bumble bee species. The means and standard errors were estimated from a truncated negative binomial model. 96

8. The presence or absence of colonization for *Bombus impatiens* samples across colonies inoculated with *Snodgrassella* strains isolated from bumble bee species. Estimated marginal means and standard errors are estimated from a binomial model. When standard errors appear to not be present it is because the errors are extremely small and represent cases where *Snodgrassella* either colonized every sample or colonized no sample 97

9. The effect of immune stimulation treatment on hemolymph antibacterial activity (standardized to units of the antibiotic tetracycline). 123

10. The relative abundance of bacterial genera making up at least 1% of the total reads (>161,161 reads) in the amplicon sequencing data set across **A)** immune stimulation treatment, **B)** treatment day, and **C)** bumble bee host colonies. 124

11. Bray-Curtis ordination plots of the composition of bumble bee microbiomes across immune stimulation treatments and treatment days. 125

12. The abundances of *Schmidhempelia* **(A)** and *Gilliamella* **(B)** across treatment initiation day and immune stimulation treatments, respectively. Error bars represent the standard error 126

13. Spearman correlation-based networks for immune stimulation treatments of naïve (A), wounding (B), Gram –'ve injected (C), and Gram +'ve injected (D). Node color is related to the cluster the bacterial genus belongs to in the network, black lines indicate negative correlations, green lines indicate positive correlations. Node size indicates normalized abundance and line thickness the strength of the interaction. 127

14. Spearman correlation-based networks for immune stimulation treatments of naïve **(A)**, wounding **(B)**, heat-killed *E. coli* injected **(C)**, and *S. epidermidis* injected **(D)**. Node color is related to the cluster the bacterial genus belongs to in the network, black lines indicate negative correlations, green lines indicate positive correlations. Node size indicates normalized abundance and line thickness the strength of the interaction. 146

15. Ordination plot of calculated Bray-Curtis distances. Points correspond to individual samples with point shape and color corresponding to queen status. Ellipses demark individual points within each queen status. 147

16. Ordination plot created by STAMP from the gene copy abundance of each gene annotated by KEGG or COG annotation. Individual points correspond to samples while the coloration and shape correspond to different queen statuses. 148

17. Ordination plot of metabolite profiles across queen statuses. Individual points correspond to samples while the coloration and shape correspond to different statuses. The colored regions show the 95% confidence regions. 149

18. Partial least square-discriminant analysis (PLS-DA) identifying the important features among the queen statuses. The metabolites with Variable Importance in Projection (VIP) scores > 1 are listed. The colored boxes on the right show the relative abundance levels of the corresponding metabolite in each group with the mated queen group (MQ) as the reference. 150

19. Protein levels in media depleted by single microbe and two-microbe community structures. Protein measures were taken as optical densities using the ThermoFisher Pierce BCA Protein Assay kit. Points represent the estimated marginal means and bars represent standard errors for the fitted model with n=6 samples/community structure. The control represents unconditioned, sterile media incubated for the same duration as conditioned media. 164

20. Consumption of depleted media treatments for **A)** germ-free bees immediately after emerging and **B)** germ-free bees starting at 3 days after emerging. The points are the estimated marginal means while the error bars are standard errors. Control shows individuals fed a mixture of sugar water with unconditioned, sterile media incubated for the same duration as the conditioned media. 165

21. The effect of depleted media on bee survival as shown by estimated survival hazards for **A)** starved bees and **B)** IAPV infected bees. Points represent the model-estimated survival hazard and error bars show represent the 95% confidence intervals. Control shows individuals fed a mixture of sugar water with unconditioned, sterile media incubated for the same duration as the conditioned media. Sugar water are bees fed only a sugar water mixture with no media. 166

# CHAPTER I: INTRODUCTION TO AN ECOLOGICAL AND EVOLUTIONARY CONTEXT FOR HOST GUT MICROBIAL COMMUNITIES

## Background

In the Costa Rican rainforest, a line of leaf cutting ants can be found carrying freshly cut leaves. These ants bear their heavy loads onward to their colonies, not to feed themselves, or their young, but rather the symbiotic fungal cultivars living within their colonies (1). As the cultivated fungus grows and matures, it produces digestive enzymes aiding in leaf decomposition and specialized structures that are fed to developing ant larvae. In this relationship, neither ant colony nor fungus can survive without the other, and the fungus is so precious to the leafcutter ants, that the ants will even groom the fungus (2). However, there are more characters in this play than those that can be seen with the naked eye. Upon the cuticles of these ants grows a biofilm containing an assemblage of actinobacteria species. These bacterial species secrete variants of antifungal nystatins, which aid the ants in removing fungal pathogens from their fungal crops (3). Both the bacteria and the farmed fungi are vertically transmitted from the founding queen and are actively maintained by the ants. If the bacterial biofilms fail, the fungal gardens deteriorate, and the ant colony collapses. This story of ants, their associated bacteria, and their fungal gardens is just one of the many complex, interacting communities of host-associated microbial symbionts that exist within nature. Studies examining these communities and their relationships not only provide us with crucial information related to the health of our natural world and our understanding of interactions within it but allow us to paint beautiful pictures to demonstrate curiosity capturing complexity.

Most multicellular organisms harbor microbial communities, termed microbiota. The gut bacteria microbial community of metazoans has received special attention, and there is an increasing appreciation of its formation and function. These communities aid their hosts in the realization of many fitness-related traits, including digestion (4–7), development (8–11), and defense against pathogens (12–14). Well-studied examples of these relationships include the microbial community within the gut of termites that plays a central role in the digestion of their high cellulose diet (5), the colonization dependent maturation of the Hawaiian bobtail squid's light organ (10), and the microbe-assisted defense against *Clostridium difficile* in humans (15, 16). Additionally, hosts usually experience detrimental outcomes when their native associated microbial communities are disturbed, further supporting the critical roles these microbes play in host health (17–19). For example, when gut communities in hamsters are perturbed by antibiotics, mortality from *C. difficile* infection is increased (20); a pattern mirrored in honey bees, where disturbed microbial communities are associated with increased mortality to *Serratia marcescens* infection (19). However, despite major recent advances in the study and characterization of these beneficial microbial symbiont communities, questions remain regarding their function, assembly, and stability.

**Bidirectionality of host-microbiota interactions**

The examples highlighted above demonstrate the influence of individual microbes and microbiota on their hosts, but the host also plays an important role in determining microbiota establishment, development, and function. Hosts may set ecological filters on their microbiota through diet, their environment, selective feeding of microbes, and immune function (21–28). Therefore, while much attention has been given to the influence of microbiota on host health, studies elucidating the role hosts

play in selecting and altering their microbiota are needed to fully understand host-microbe interactions and resulting phenotypes.

Diet plays a critical role in selecting both the microbiota of the host (21, 22, 29), but also works bidirectionally where microbiota function influences the host niche (30). Within humans breast milk derived oligosaccharides promoted a diversity of *Bifidiobacterium* strains through cross-feeding interactions and partitioning of resources (31). These findings in humans are further supported by experiments in honeybees finding that niche partitioning between gut microbes on common dietary components such as sugars and pollen facilitated the coexistence of microbial species (32). Additional work done on vulture bees (*Trigona spp.*) demonstrates that carrion feeding bees possessed a more variable microbiome than bee species that are strictly pollinovorous (33). Thus, the host diet represents a critical mechanism through which the host may influence the structure, stability, and functionality of its microbiota.

The environment of the host can influence its microbiota in a number of mechanisms. These mechanisms can be direct, by serving as the inoculation source of the microbial community, or indirect by providing sources of perturbation that the community must respond to. Within *Drosophila* the microbiota is horizontally transmitted through the environment and food sources (34, 35). Therefore, the environment in which the larvae emerge determines which microbes are present for colonization. In mosquito species the larval environment was found to significantly influence the adult microbiota (28). Additionally, host exposures to environmental factors, such as temperature or environmental antibiotics may perturb the microbiota structure. Within bees, the temperature experienced by the host has been

shown to influence the colonization success of a common gut microbiota member (36), while the effects of antibiotic exposure on host microbiota have been well documented (18–20). Thus, whether directly or indirectly, hosts may in many ways select for their microbial communities because of the environmental conditions they experience.

Finally, hosts may select specific microbial community structures through physiological responses. Conceptual modelling combined with empirical approaches has demonstrated that changes in host mucosal glycans and immunoglobulins can select against certain microbial species (26, 27). Both plants and mammals have been shown to secrete substances that the microbiota feed upon (25, 31, 37), and therefore it is possible that hosts can modify the rate of secretion or even completely halt secretion of these food sources to control their microbial communities (38). Additionally, hosts may mount immune responses to control or manipulate the microbes present in the microbiota (39–41). Although, most of the work on immune modulation of the microbiota focuses on adaptive and mammalian systems, with work in other systems often focusing on the role of microbiota in developing the immune response. It is also possible that host immunity may disrupt beneficial microbiota as a side-effect of the response to infection (42), which could constrain host immune evolution (43). In other cases, it has been shown that pathogens co-opt host immunity to disrupt the hosts native microbiota to facilitate infection (44).

**Microbial community interactions determining membership and function**

Owing to the evolutionary history between hosts, their microbiota, and between members of the microbiota themselves, complex ecological interactions may exist within these communities. These interactions may include beneficial interactions such as cross-feeding, biofilm formation, and horizontal

gene transfer, or competitive interactions through direct or indirect killing. Therefore, the structure of these ecological communities, succession of varying taxa, and evolutionary histories likely determines the functionality and stability of the microbiota (45–51). For example, multiple members of the microbiota may be required to breakdown host ingested glycans, especially considering hosts often possess only a fraction of the enzymatic repertoire required to utilize these compounds (29, 52). Work in humans has shown a vast array of microbial enzymes targeting specific linkages of the glycans but also species-specificity in the capacity of different gut microbial members (52–54). Such a division of the enzymatic arsenal among community members can lead to cross-feeding, with sequential steps with microbes feeding on the metabolic byproducts from microbes in prior steps, further breaking down ingested glycans. Cross-feeding relationships may determine associations between microbial community members, as the colonization of a cross-feeder will necessarily depend on the presence of other microbes (7, 29). Community members can also influence the gene expression and functioning of co-occurring members through quorum sensing or metabolic cues. Interactive effects between microbial community members is seen in *in vitro* four-member microbial communities where the bacterium *Mycobacterium oxydans* makes up only 0.5% of biomass but greatly enhances *Paenibacillus amylolyticus* biofilm formation (55). Moreover, the difficulty in culturing many microbiota community members has been attributed to microbe codependence (51).

Determining the processes that drive microbiota variation across host individuals, populations, and species is essential to understanding the benefit these communities provide their hosts and the ecological and evolutionary underpinnings of them. This includes not only investigating benefits and host

interactions with individual microbes, but in naturally co-occurring pairs and communities to uncover important pairwise or higher-order interactions that determine emergent properties of microbial communities and community functionality. Emergent properties are properties that are not able to be determined or predicted parts of a system in isolation. For example, the consciousness is an emergent property of the human brain, that is not able to be predicted or understood from examining the neurons alone. In terms of community ecology, emergent properties often result from dynamic interactions between community members and the interactions with abiotic factors. For example, the methane flux of a microbial ecosystem is likely determined by the microbe genotypes, community interactions, the redox chemistry conditions of the environment and other physiochemical variables (56). Microbial communities possess many emergent properties that make them unique when compared to traditional animal or plant communities (57). These differences include functional redundancy that is typically low in eukaryotic communities but high in the microbiota, exchange of genetic material between community members common in microbiota, and the secretion and utilization of public goods (57, 58).

Organisms in eukaryotic communities are expected to segregate into roles or specific niches, as seen in the communities of Mediterranean woody species (59), freshwater fish (60), and mammals such as the fur seal (61). While members of host microbiota do show signs of niche segregation (51), microbial communities display a significantly greater degree of functionally redundancy (62–64). Examples of redundancy in microbial communities include communities of bacterioplankton, but also within host examples such as microbial short chain fatty acid production in the human gut or the fermentation properties of the ruminal gut microbial community (62, 65, 66). Redundancy of host-associated

communities likely results in resilience and stability to perturbations, but what is less understood are the evolutionary drivers behind this stability.

Relative high rates of genetic exchange through horizontal gene transfer between microbes may cloud traditional taxonomy-function relationships and affect how hosts and other factors affect microbiota membership and subsequent function. Because individuals within microbial communities can exchange important functional genes, including for metabolism (67), antibiotic resilience (68), or even virulence (69), species-function relationships may be difficult to determine and has consequences for comparisons of microbial communities between host individuals, populations, and species. Horizontal gene transfer between community members also raises questions about how hosts may monitor and select for their microbial communities, as taxonomic level surveillance may not always select for beneficial communities.

Interactions between microbes introduced above, including their communication and use of public goods also represent relevant features of microbial communities. Individuals within communities are able to communicate and synchronize gene expression via quorum sensing (70, 71), and catabolite repression (29), changing process and function in response to publicly secreted goods from other individuals. Interestingly, the evolutionary response to the secretion of these public goods in microbial communities is the possible reduction of genome size, another unique aspect of host-associated microbes (58). Such genome size reductions may reinforce dependencies and cross-feeding dynamics in specialized host-associated community members, while the change in expression associated with metabolic and quorum sensing may also serve to maintain functional redundancy through plastic responses.

Work investigating host-associated microbiota needs to account for the above unique attributes of microbial communities. Yet, many of the experimental investigations of these communities have relied upon amplicon sequencing to identify abundant taxa without further consideration. These approaches are powerful investigative tools (72–76), but may miss important features and dynamic interactions of microbial communities because the unique properties of these communities or technical issues of sequencing biases, poor accuracy of automatic classification systems in previously undescribed systems, and statistical difficulties analyzing large datasets (74, 77–81). For example, amplicon sequencing is predictive of only relative abundance and may poorly estimate the absolute colonization of these microbes (82). The limitations of these experimental approaches feed into a framework in which microbiota studies fail to account for the dynamics of these communities or provide incomplete pictures of these communities (83, 84). Thus, studies integrating measures of absolute colonization abundance, community traits, and metabolic function provide a better understanding of these communities. Recently, studies have begun to appreciate the separation of hosts from their microbiota, focusing on interactions between not only host and microbiota, but also the members of the community, with explicit attempts to establish causation between these components, host disease states, and the response to external influences (45, 46, 82, 85–87).

Specific issues with many current frameworks for microbiota research are rooted in evolutionary theory. First, the evolution of traits beneficial to hosts in these communities would postulate some reciprocal benefit for the microbial community members. While often demonstrated for hosts, the benefit for microbes can often be more elusive and is often simply assumed. Even the evolution of traits in

microbes that are beneficial for their hosts may be problematic. When in a diverse community, if a microbe evolves a host-beneficial trait at a cost it may be out competed. Thus, the microbiota and host should not be assumed to be one unit with aligned interests, although these interests may be more aligned in cases of predominantly vertical transmission of microbes within host lineages (88). From the host perspective, is how do they surveil and maintain their microbiota, which often contain hundreds to thousands of species and where taxonomy and function may be disentangled. Addressing these and other issues is a central goal for the ecosystem on a leash model of hosts and their microbial communities (46), postulating that the microbiota is not only driven by bottom-up species interactions, but also that the host is under strong selection to shape its microbiota from the top down, fostering a community that is beneficial (46). This framework generates some distinguishing and testable hypotheses and predictions compared to more traditional frameworks. The host presents an ecosystem in which microbes may interact to fill the niche space. Also, the host will select for specific microbial assemblies that provide benefit to the host. Finally, the host will punish communities that fail to provide beneficial functions while rewarding those that do. Using these initial hypotheses about how microbial communities evolve and interact it is possible to extend these hypotheses to systems with vertically transmitted microbiota. First, if the microbiota is indeed an ecosystem on a leash it should display some characteristics of microbial communities such as facilitation and emergent properties of processing chains, but these will likely differ between communities of different host lineages. Second, hosts likely respond to communities that fail to provide beneficial function through immune responses, as such policing cheating microbial strains. Third,

the unique relationship between plasticity and redundancy may be an evolved response allowing microbes to maintain beneficial functioning during transient changes to avoid host punishment.

**Bumble bees as a model system for understanding host gut microbial communities**

The studies carried out in this dissertation focus on the gut microbiota of insects, and predominantly bumble bee (*Bombus spp.*) gut associated microbial communities. Bumble bees are excellent model systems for studying host-microbe interactions. They have long been a system to study physiological responses to environmental variation (89–93). Moreover, recently, multiple species of bumble bee have had their genomes fully sequenced (94, 95), allowing for functional genomic approaches to measure specific host responses. From the perspective of understanding the role of associated microbes, axenic bees are easy to rear in the lab (96). Many species of bees mate singly, as opposed to honeybees which will mate many times, resulting in a high relatedness between offspring workers, reducing the genetic variation between samples within a colony (97, 98). Temperate bumble bee species (except the social parasitic sub genus *Psithyrus*) have a life cycle that consists of solitary and social stages (99). Established social colonies will produce new offspring queens and males towards the end of the colony cycle. Queens leave the colony and seek out a mate and subsequently queens begin a solitary phase of their lives by entering a period of diapause. Following emergence from diapause, queens search for nest sites, forage and initiate new colonies by laying eggs and rearing larvae. This ultimately starts the social phase of cycle again, as adult workers emerge and participate in rearing of their siblings. This life cycle is conductive to the vertical transmission of host-associated microbes. Because of this vertical

transmission (100), many of the factors influencing microbe-microbe-host interactions likely have parallels to mammalian systems.

Corbiculate bees, including bumble bees, harbor a relatively simple gut community (101), containing about 10-15 members. While many studies focus on the microbiota of honeybees, bumble bees share many of the honeybee core species, thus these studies can begin to demonstrate the importance of the bumble bee microbiota while generating predictions for these microbes in bumble bee systems. Many members are largely vertically transmitted and show some degree of coevolution with their bee hosts (102). Coevolution between the gut community of these bees and their hosts has likely created a codependence between the two, with the microbiota now aiding the host in many aspects of fitness including digestion (7, 103), defense against parasites and pathogens (13, 19, 104), and even development (8, 96, 105, 106). Within the core microbiota, three members, *Snodgrassella alvi*, *Gilliamella spp.*, and *Lactobacillus spp.*, have been the central focus of recent research, and *S. alvi* and *G. apicola* have sequenced genomes (107).

*Snodgrassella alvi* is the primary colonizer of the bumble bee gut (100). This gram-negative, beta-proteobacterium forms a thick biofilm on the bumble bee's hindgut epithelium (108). The biofilm formed by this microbe has been implicated in host weight gain and ecosystem engineering, by maintaining oxygen gradients in the hindgut (6). Additionally, previous research has shown that strains are generally phylogenetically distributed across corniculate bees (109). Furthermore, there is evidence that specificity in colonization exists at the genus level between bumble bee and honey bee hosts (102, 107), and that

colonization in bumble bees may be further determined by interactions between host and microbe genotype (110).

*Gilliamella spp.* are commonly occurring gram-negative, Gamma-proteobacterium within the gut of social bees. *Gilliamella* has been implicated as an important microbe for bee digestion. When grown on agar plates with polygalacturonic acid, the polysaccharide backbone of pectin, *Gilliamella* strains caused zones of clearance (103). The presence of these zones signifies the secretion of extra-cellular enzymes targeting these pectin backbones. Thus, members of *Gilliamella* genus are likely to be crucial for the degradation of more recalcitrant resources in the bumble bee hindgut (111).

*Lactobacillus* firm-4 and firm-5 are also commonly found within the social bee hindgut (112, 113). These bacteria are implicated as important fermenters, aiding the host in digesting many of the specialized glycans found in plant nectar and pollen, as well as possibly aiding in pathogen defense (96, 104). Interestingly, analyses of firm-5 *Lactobacillus* not only support the bacterium's importance in digestion, but also demonstrates genus-level specificity like that observed in *S. alvi*. While sublineages of firm-5 *Lactobacillus* from bumble bees were able to colonize germ-free honeybees they were outcompeted by the native honeybee sublineages (114).

The unique evolutionary history of social bees and their microbiota, the observed inter-dependencies and processing chains between microbiota members, and the specificity of sublineages of the microbes is strongly suggestive that the ecosystem on a leash type model may be at play in this system. First, the fact that both *S. alvi* and *Lactobacillus* Firm-5 sublineages that are native to their host genus outcompete foreign sublineages has strong resemblance to one of the primary predictions from the

ecosystem on a leash model, which predicts that microbes should interact through community dynamics to adapt to the available niches presented by their hosts (45). Thus, strains of a microbial species that have not adapted to their specific host environment should be less able to colonize or be outcompeted by native strains. Additionally, there are significant processing chains and cross-feeding dynamics in the social bee community leading to breakdown of recalcitrant resources. This provides a benefit to the host through increased energetic gain, but also detoxification of metabolic byproducts and toxic sugars (111). For example, *Lactobacillus* and *Gilliamella* strains likely work redundantly to break down pollen using polygalacturonases, fermenting the by-products and producing lactate and acetate. *Snodgrassella alvi* can then utilize that lactate and acetate to perform its uniquely modified TCA cycle (115), using this energy to create and secrete amino acids and vitamins that the host and other microbiota can feed on. This processing chain exemplifies the redundancy that is characteristic to microbiota, and this redundancy likely buffers the bee against transient changes to community composition. If *Lactobacillus* or *Gilliamella* populations are perturbed, the other species may be able to still fulfill the function, preventing the host from responding with punishment through immune stimulation. Additionally, the processing chain may create situations where the colonization of microbes is dependent on the presence of the other co-occurring community members.

The central goal of my dissertation is to investigate the evolutionary and ecological interactions that determine the membership and functionality of host-associated microbial communities, interpreting outcomes within the ecosystem on a leash and other frameworks. In **Chapter 2** I report the investigation of microbial communities across a diverse range of insect pollinators that is based on their community

membership and general structure. Surprisingly, we found a high abundance of bacteria belonging to the previously bee-associated genus *Gilliamella* in pollinating flies. Utilizing metagenomic sequencing we investigate differences between the fly and bee associated *Gilliamella*, reconstructing phylogenetic relationships and metabolic pathways. In **Chapter 3** I investigate the extent of colonization specificity in the core gut microbiota member of bumble bees, *Snodgrassella alvi*. Previously, a small number of strains of this microbe were found to be specific at the genus level across honey bees and a single bumble bee species. I expand on this previous work by investigating specificity across several strains isolated from honey bee and different bumble bee host species inoculated into two bumble bee species. In **Chapter 4** I examine the influence of host innate immune stimulation on the gut microbiota of bumble bees, addressing a hypothesis that infection-induced responses have the potential to perturb the beneficial microbiota. In **Chapter 5** I report an investigation into the relationship between the bumble bee queen life cycle stage, and her metabolome and microbiota. Finally, in **Chapter 6** I test whether intrinsic properties of microbial communities influence the health of bumble bee hosts.

**References**


1. Hölldobler B, Wilson E. 1998. *Journey to the ants: a story of scientific exploration*. Harvard University Press.

2. Abramowski D, Currie CR, Poulsen M. 2011. Caste specialization in behavioral defenses against fungus garden parasites in *Acromyrmex octospinosus* leaf-cutting ants. *Ins. Soc.* **58**:65–75.

3. Heine D, Holmes NA, Worsley SF, Santos ACA, Innocent TM, Scherlach K, Patrick EH, Yu DW, Murrell JC, Vieria PC, Boomsma JJ, Hertweck C, Hutchings MI, Wilkinson B. 2018. Chemical warfare between leafcutter ant symbionts and a co-evolved pathogen. *Nat. Comm.* **9**:2208.

4. Cleveland LR. 1923. Symbiosis between termites and their intestinal protozoa. *Proc. Natl. Acad. Sci. USA.* **9**:424–428.

5. Brune A. 2014. Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbol.* **12**:168–180.

6. Zheng H, Powell JE, Steele MI, Dietrich C, Moran NA. 2017. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc. Natl. Acad. Sci. USA.* **114**:4775–4780.

7. Kešnerová L, Mars RAT, Ellegaard KM, Troilo M, Sauer U, Engel P. 2017. Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol.* **15**:e2003467.

8. Engel Philipp, Bartlett Kelsey D., Moran Nancy A. 2015. The bacterium *Frischella perrara* causes scab formation in the gut of its honeybee host. *mBio*. **6**:e00193-15.

9. Maynard CL, Elson CO, Hatton RD, Weaver CT. 2012. Reciprocal interactions of the intestinal microbiota and immune system. *Nature.* **489**:231–241.

10. Visick KL, Foster J, Doino J, McFall-Ngai M, Ruby EG. 2000. *Vibrio fischeri* lux genes play an important role in colonization and development of the host light organ. *J. Bacteriol.* **182**:4578–86.

11. Palm NW, de Zoete MR, Flavell RA. 2015. Immune-microbiota interactions in health and disease. *Clin. Immunol*. **159**:122–127.

12. Jaenike J, Unckless R, Cockburn SN, Boelio LM, Perlman SJ. 2010. Adaptation via symbiosis: recent spread of a *Drosophila* defensive symbiont. *Science*. **329**:212—215.

13. Koch H, Schmid-Hempel P. 2011. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc. Natl. Acad. Sci. USA*. **108**:19288–92.

14. Oliver KM, Noge K, Huang EM, Campos JM, Becerra JX, Hunter MS. 2012. Parasitic wasp responses to symbiont-based defense in aphids. *BMC Biol*. **10**:11.

15. Buffie CG, Pamer EG. 2013. Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol*. **13**:790–801.

16. Buonomo EL, Cowardin CA, Wilson MG, Saleh MM, Pramoonjago P, Petri WA Jr. 2016. Microbiota-regulated IL-25 Increases eosinophil number to provide protection during *Clostridium difficile* infection. *Cell Rep*. **16**:432–443.

17. Hagan T, Cortese M, Rouphael N, Boudreau C, Linde C, Maddur MS, Das J, Wang H, Guthmiller J, Zheng N-Y, Huang M, Uphadhyay AA, Gardinassi L, Petitdemange C, McCullough MP, Johnson SJ, Gill K, Cervasi B, Zou J, Bretin A, Hahn M, Gewirtz AT, Bosinger SE, Wilson PC, Li S, Alter G, Khurana S, Golding H, Pulendran B. 2019. Antibiotic-driven gut microbiome perturbation alters immunity to vaccines in humans. *Cell*. **178**:1313-1328.e13.

18. Li JH, Evans JD, Li WF, Zhao YZ, DeGrandi-Hoffman G, Huang SK, Li ZG, Hamilton M, Chen YP. 2017. New evidence showing that the destruction of gut bacteria by antibiotic treatment could increase the honey bee's vulnerability to *Nosema* infection. *PLoS One*. **12**:e0187505.

19. Raymann K, Shaffer Z, Moran NA. 2017. Antibiotic exposure perturbs the gut microbiota and elevates mortality in honeybees. *PLoS Biol*. **15**:e2001861.

20. Burdet C, Sayah-Jeanne S, Nguyen TT, Hugon P, Sablier-Gallis F, Saint-Lu N, Corbel T, Ferreira S, Pulse M, Weiss W, Andremont A, Mentré F, de Gunzburg J. 2018. Antibiotic-induced dysbiosis predicts mortality in an animal model of *Clostridium difficile* infection. *Antimicrob. Agents Chemother*. **62**.

21. Brunetti M, Magoga G, Gionechetti F, De Biase A, Montagna M. 2022. Does diet breadth affect the complexity of the phytophagous insect microbiota? The case study of Chrysomelidae. *Environ. Microbiol.* **24**:3565–3579.

22. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, Henrissat B, Knight R, Gordon JI. 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science.* **332**:970–4.

23. Kolodziejczyk AA, Zheng D, Elinav E. 2019. Diet–microbiota interactions and personalized nutrition. *Nat. Rev. Microbiol*. **17**:742–753.

24. Marcobal A, Southwick AM, Earle KA, Sonnenburg JL. 2013. A refined palate: bacterial consumption of host glycans in the gut. *Glycobiology.* **23**:1038–1046.

25. Tamura K, Brumer H. 2021. Glycan utilization systems in the human gut microbiota: a gold mine for structural discoveries. *Curr. Opin. Struct. Biol.* **68**:26–40.

26. McLoughlin K, Schluter J, Rakoff-Nahoum S, Smith AL, Foster KR. 2016. Host selection of microbiota via differential adhesion. *Cell Host Microbe*. **19**:550–559.

27. Schluter J, Foster KR. 2012. The evolution of mutualism in gut microbiota via host epithelial selection. *PLoS Biol*. **10**:e1001424.

28. Saab SA, Dohna H zu, Nilsson LKJ, Onorati P, Nakhleh J, Terenius O, Osta MA. 2020. The environment and species affect gut bacteria composition in laboratory co-cultured *Anopheles gambiae* and *Aedes albopictus* mosquitoes. *Sci. Rep*. **10**:3352.

29. Tuncil YE, Xiao Y, Porter NT, Reuhs BL, Martens EC, Hamaker BR. 2017. Reciprocal prioritization to dietary glycans by gut bacteria in a competitive environment promotes stable coexistence. *mBio*. **8**.

30. Cornwallis CK, van ’t Padje A, Ellers J, Klein M, Jackson R, Kiers ET, West SA, Henry LM. 2023. Symbioses shape feeding niches and diversification across insects. *Nat. Ecol. Evol.*

31. Lawson MAE, O’Neill IJ, Kujawska M, Gowrinadh Javvadi S, Wijeyesekera A, Flegg Z, Chalklen L, Hall LJ. 2020. Breast milk-derived human milk oligosaccharides promote *Bifidobacterium* interactions within a single ecosystem. *ISME J*. **14**:635–648.

32. Brochet S, Quinn A, Mars RAT, Neuschwander N, Sauer U, Engel P. 2021. Niche partitioning facilitates coexistence of closely related honey bee gut bacteria. *Elife*.

33. Figueroa LL, Maccaro JJ, Krichilsky E, Yanega D, McFrederick QS. 2021. Why Did the Bee eat the chicken? Symbiont gain, loss, and retention in the vulture bee microbiome. *mBio*. **12**:e0231721.

34. Davoodi S, Foley E. 2020. Host-microbe-pathogen interactions: A review of *Vibrio cholerae* pathogenesis in *Drosophila*. *Front. Immunol.* **10**.

35. Wong AC, Chaston JM, Douglas AE. 2013. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *ISME J*. **7**:1922–32.

36. Hammer TJ, Le E, Moran NA. 2021. Thermal niches of specialized gut symbionts: the case of social bees. *Proc. Roy. Soc. Biol. Sci*. **288**:20201480.

37. Nguema-Ona E, Vicré-Gibouin M, Cannesan M-A, Driouich A. 2013. Arabinogalactan proteins in root-microbe interactions. *Trends. Plant. Sci.* **18**:440–449.

38. Kiers ET, Rousseau RA, West SA, Denison RF. 2003. Host sanctions and the legume–rhizobium mutualism. *Nature*. **425**:78–81.

39. Levy M, Kolodziejczyk AA, Thaiss CA, Elinav E. 2017. Dysbiosis and the immune system. *Nat. Rev. Immunol.* **17**:219–232.

40. Yoo JY, Groer M, Dutra SVO, Sarkar A, McSkimming DI. 2020. Gut microbiota and immune system interactions. *Microorganisms*. **8**.

41. Knoop KA, McDonald KG, McCrate S, McDole JR, Newberry RD. 2015. Microbial sensing by goblet cells controls immune surveillance of luminal antigens in the colon. *Mucosal Immunol.* **8**:198–210.

42. Armitage SA, Genersch E, McMahon DP, Rafaluk-Mohr C, Rolff J. 2022. Tripartite interactions: how immunity, microbiota and pathogens interact and affect pathogen virulence evolution. *Curr. Opin. Insect. Sci.* **50**:100871.

43. Gerardo NM, Hoang KL, Stoy KS. 2020. Evolution of animal immunity in the light of beneficial symbioses. *Phil. Trans. Roy. Soc. Biol. Sci.* **375**:20190601.

44. Abraham NM, Liu L, Jutras BL, Yadav AK, Narasimhan S, Gopalakrishnan V, Ansari JM, Jefferson KK, Cava F, Jacobs-Wagner C, Fikrig E. 2017. Pathogen-mediated manipulation of arthropod microbiota to promote infection. *Proc. Natl. Acad. Sci. USA*. **114**:E781–E790.

45. Coyte KZ, Schluter J, Foster KR. 2015. The ecology of the microbiome: Networks, competition, and stability. *Science*. **350**:663–666.

46. Foster KR, Schluter J, Coyte KZ, Rakoff-Nahoum S. 2017. The evolution of the host microbiome as an ecosystem on a leash. *Nature*. **548**:43–51.

47. Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, Beman JM, Abell G, Philippot L, Prosser J, Foulquier A, Yuste JC, Glanville HC, Jones DL, Angel R, Salminen J, Newton RJ, Bürgmann H, Ingram LJ, Hamer U, Siljanen HMP, Peltoniemi K, Potthast K, Bañeras L, Hartmann M, Banerjee S, Yu R-Q, Nogaro G, Richter A, Koranda M, Castle SC, Goberna M, Song B, Chatterjee A, Nunes OC, Lopes AR, Cao Y, Kaisermann A, Hallin S, Strickland MS, Garcia-Pausas J, Barba J, Kang H, Isobe K, Papaspyrou S, Pastorelli R, Lagomarsino A, Lindström ES, Basiliko N, Nemergut DR. 2016. Microbes as Engines of Ecosystem Function: When does community structure enhance predictions of ecosystem processes? *Front. Microbiol.* **7**:214.

48. Guittar J, Shade A, Litchman E. 2019. Trait-based community assembly and succession of the infant gut microbiome. *Nat. Comm.* **10**:512.

49. Lin XB, Wang T, Stothard P, Corander J, Wang J, Baines JF, Knowles SCL, Baltrūnaitė L, Tasseva G, Schmaltz R, Tollenaar S, Cody LA, Grenier T, Wu W, Ramer-Tait AE, Walter J. 2018. The evolution of ecological facilitation within mixed-species biofilms in the mouse gastrointestinal tract. *ISME J.* **12**:2770–2784.

50. Mori AS, Osono T, Cornelissen JHC, Craine J, Uchida M. 2017. Biodiversity–ecosystem function relationships change through primary succession. *Oikos*. **126**:1637–1649.

51. Plichta DR, Juncker AS, Bertalan M, Rettedal E, Gautier L, Varela E, Manichanh C, Fouqueray C, Levenez F, Nielsen T, Doré J, Machado AMD, de Evgrafov MCR, Hansen T, Jørgensen T, Bork P, Guarner F, Pedersen O, Sommer MOA, Ehrlich SD, Sicheritz-Pontén T, Brunak S, Nielsen HB. 2016. Transcriptional interactions suggest niche segregation among microorganisms in the human gut. *Nat. Microbiol.* **1**:16152.

52. Koropatkin NM, Cameron EA, Martens EC. 2012. How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol*. **10**:323–335.

53. Cerqueira FM, Photenhauer AL, Pollet RM, Brown HA, Koropatkin NM. 2020. Starch digestion by gut bacteria: crowdsourcing for carbs. *Trends Microbiol*. **28**:95–108.

54. El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B. 2013. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* **11**:497–504.

55. Liu W, Russel J, Røder HL, Madsen JS, Burmølle M, Sørensen SJ. 2017. Low-abundant species facilitates specific spatial organization that promotes multispecies biofilm formation. *Environ. Microbiol*. **19**:2893–2905.

56. Gilbert JA, Henry C. 2015. Predicting ecosystem emergent properties at multiple scales. *Environ. Microbiol. Rep*. **7**:20–22.

57. Koskella B, Hall LJ, Metcalf CJE. 2017. The microbiome beyond the horizon of ecological and evolutionary theory. *Nat. Ecol. Evol*. **1**:1606–1615.

58. Morris JJ. 2015. Black Queen evolution: the role of leakiness in structuring microbial communities. *Trends Genet*. **31**:475–482.

59. de la Riva EG, Marañón T, Violle C, Villar R, Pérez-Ramos IM. 2017. Biogeochemical and ecomorphological niche segregation of Mediterranean woody species along a local gradient. Frontiers in *Plant Sci*. **8**.

60. Comte L, Cucherousset J, Boulêtreau S, Olden JD. 2016. Resource partitioning and functional diversity of worldwide freshwater fish communities. *Ecosphere*. **7**:e01356.

61. de Lima RC, Franco-Trecu V, Vales DG, Inchausti P, Secchi ER, Botta S. 2019. Individual foraging specialization and sexual niche segregation in South American fur seals. *Marine Biol.* **166**:32.

62. Comte J, Fauteux L, Del Giorgio PA. 2013. Links between metabolic plasticity and functional redundancy in freshwater bacterioplankton communities. *Front. Microbiol*. **4**:112.

63. Moya A, Ferrer M. 2016. Functional redundancy-induced stability of gut microbiota subjected to disturbance. *Trends Microbiol*. **24**:402–413.

64. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, Ackermann M, Hahn AS, Srivastava DS, Crowe SA, Doebeli M, Parfrey LW. 2018. Function and functional redundancy in microbial systems. *Nat. Ecol. Evo.* **2**:936–943.

65. Reichardt N, Vollmer M, Holtrop G, Farquharson FM, Wefers D, Bunzel M, Duncan SH, Drew JE, Williams LM, Milligan G, Preston T, Morrison D, Flint HJ, Louis P. 2018. Specific substrate-driven changes in human faecal microbiota composition contrast with functional redundancy in short-chain fatty acid production. *ISME J.* **12**:610–622.

66. Weimer PJ. 2015. Redundancy, resilience, and host specificity of the ruminal microbiota: implications for engineering improved ruminal fermentations. *Front Microbiol.* **6**:296.

67. Pál C, Papp B, Lercher MJ. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet*. **37**:1372–1375.

68. Maeusli M, Lee B, Miller S, Reyna Z, Lu P, Yan J, Ulhaq A, Skandalis N, Spellberg B, Luna B. 2020. Horizontal gene transfer of antibiotic resistance from *Acinetobacter baylyi* to *Escherichia coli* on lettuce and subsequent antibiotic resistance transmission to the gut microbiome. *mSphere*. **5**.

69. Jain H, Mulay S, Mullany P. 2016. Persistence of endodontic infection and *Enterococcus faecalis*: Role of horizontal gene transfer. *Gene Rep*. **5**:112–116.

70. Duddy OP, Bassler BL. 2021. Quorum sensing across bacterial and viral domains. *PLoS Pathog*. **17**:e1009074.

71. Whiteley M, Diggle SP, Greenberg EP. 2017. Progress in and promise of bacterial quorum sensing research. *Nature*. **551**:313–320.

72. Gołębiewski M, Tretyn A. 2020. Generating amplicon reads for microbial community assessment with next-generation sequencing. *J. Appl. Microbiol.* **128**:330–354.

73. Hugerth LW, Andersson AF. 2017. Analysing microbial community composition through amplicon sequencing: from sampling to hypothesis testing. *Front. Microbiol*. **8**.

74. Silva DP, Epstein HE, Vega Thurber RL. 2022. Best practices for generating and analyzing 16S rRNA amplicon data to track coral microbiome dynamics. *Front. Microbiol.* **13**:1007877.

75. Kinoshita Y, Niwa H, Uchida-Fujii E, Nukada T. 2021. Establishment and assessment of an amplicon sequencing method targeting the 16S-ITS-23S rRNA operon for analysis of the equine gut microbiome. *Sci. Rep.* **11**:11884.

76. Ju F, Zhang T. 2015. 16S rRNA gene high-throughput sequencing data mining of microbial diversity and interactions. *Appl. Microbiol. Biotech.* **99**:4119–4129.

77. Schloss Patrick D. 2020. Reintroducing mothur: 10 Years Later. *Appl. Environ. Microbiol.* **86**:e02343-19.

78. Pollock Jolinda, Glendinning Laura, Wisedchanwet Trong, Watson Mick. 2018. The madness of microbiome: attempting to find consensus “best practice” for 16S microbiome studies. *Appl. Environ. Microbiol.* **84**:e02627-17.

79. McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol*. **10**:e1003531.

80. Thorsen J, Brejnrod A, Mortensen M, Rasmussen MA, Stokholm J, Al-Soud WA, Sørensen S, Bisgaard H, Waage J. 2016. Large-scale benchmarking reveals false discoveries and count transformation sensitivity in 16S rRNA gene amplicon data analysis methods used in microbiome studies. *Microbiome*. **4**:62.

81. Tsilimigras MCB, Fodor AA. 2016. Compositional data analysis of the microbiome: fundamentals, tools, and challenges. *Ann. Epidem*. **26**:330–335.

82. Rao C, Coyte KZ, Bainter W, Geha RS, Martin CR, Rakoff-Nahoum S. 2021. Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature*. **591**:633–638.

83. Fischbach MA. 2018. Microbiome: focus on causation and mechanism. *Cell*. **174**:785–790.

84. Widder S, Allen RJ, Pfeiffer T, Curtis TP, Wiuf C, Sloan WT, Cordero OX, Brown SP, Momeni B, Shou W, Kettle H, Flint HJ, Haas AF, Laroche B, Kreft J-U, Rainey PB, Freilich S, Schuster S, Milferstedt K, van der Meer JR, Groβkopf T, Huisman J, Free A, Picioreanu C, Quince C, Klapper I, Labarthe S, Smets BF, Wang H, Soyer OS, Isaac Newton Institute Fellows. 2016. Challenges in microbial ecology: building predictive understanding of community function and dynamics. *ISME J*. **10**:2557–2568.

85. Watson AR, Füssel J, Veseli I, DeLongchamp JZ, Silva M, Trigodet F, Lolans K, Shaiber A, Fogarty E, Runde JM, Quince C, Yu MK, Söylev A, Morrison HG, Lee STM, Kao D, Rubin DT, Jabri B, Louie T, Eren AM. 2023. Metabolic independence drives gut microbial colonization and resilience in health and disease. *Gen. Biol*. **24**:78.

86. Proctor DM, Relman DA. 2017. The landscape ecology and microbiota of the human nose, mouth, and throat. *Cell Host Microbe*. **21**:421–432.

87. Skwara A, Gowda K, Yousef M, Diaz-Colunga J, Raman AS, Sanchez A, Tikhonov M, Kuehn S. 2023. Learning the functional landscape of microbial communities. *bioRxiv*.

88. Leeks A, Dos Santos M, West SA. 2019. Transmission, relatedness, and the evolution of cooperative symbionts. *J. Evol. Biol*. **32**:1036–1045.

89. Cariveau DP, Winfree R. 2015. Causes of variation in wild bee responses to anthropogenic drivers. *Curr. Opin. Insect Sci*. **10**:104–109.

90. Heinrich B. 1975. Thermoregulation in bumblebees. *J. Comp. Physiol*. **96**:155–166.

91. Heinrich B. 1979. Resource heterogeneity and patterns of movement in foraging bumblebees. *Oecologia*. **40**:235–245.

92. Oyen KJ, Giri S, Dillon ME. 2016. Altitudinal variation in bumble bee (*Bombus*) critical thermal limits. *J. Therm. Biol*. **59**:52–57.

93. Woodard SH. 2017. Bumble bee ecophysiology: integrating the changing environment and the organism. *Curr. Opin. Insect Sci*. **22**:101–108.

94. Sadd BM, Barribeau SM, Bloch G, de Graaf DC, Dearden P, Elsik CG, Gadau J, Grimmelikhuijzen CJ, Hasselmann M, Lozier JD, Robertson HM, Smagghe G, Stolle E, Van Vaerenbergh M, Waterhouse RM, Bornberg-Bauer E, Klasberg S, Bennett AK, Câmara F, Guigó R, Hoff K, Mariotti M, Munoz-Torres M, Murphy T, Santesmasses D, Amdam GV, Beckers M, Beye M, Biewer M, Bitondi MM, Blaxter ML, Bourke AF, Brown MJ, Buechel SD, Cameron R, Cappelle K, Carolan JC, Christiaens O, Ciborowski KL, Clarke DF, Colgan TJ, Collins DH, Cridge AG, Dalmay T, Dreier S, du Plessis L, Duncan E, Erler S, Evans J, Falcon T, Flores K, Freitas FC, Fuchikawa T, Gempe T, Hartfelder K, Hauser F, Helbing S, Humann FC, Irvine F, Jermiin LS, Johnson CE, Johnson RM, Jones AK, Kadowaki T, Kidner JH, Koch V, Köhler

A, Kraus FB, Lattorff HMG, Leask M, Lockett GA, Mallon EB, Antonio DSM, Marxer M, Meeus I, Moritz RF, Nair A, Näpflin K, Nissen I, Niu J, Nunes FM, Oakeshott JG, Osborne A, Otte M, Pinheiro DG, Rossié N, Rueppell O, Santos CG, Schmid-Hempel R, Schmitt BD, Schulte C, Simões ZL, Soares MP, Swevers L, Winnebeck EC, Wolschin F, Yu N, Zdobnov EM, Aqrawi PK, Blankenburg KP, Coyle M, Francisco L, Hernandez AG, Holder M, Hudson ME, Jackson L, Jayaseelan J, Joshi V, Kovar C, Lee SL, Mata R, Mathew T, Newsham IF, Ngo R, Okwuonu G, Pham C, Pu L-L, Saada N, Santibanez J, Simmons D, Thornton R, Venkat A, Walden KK, Wu Y-Q, Debyser G, Devreese B, Asher C, Blommaert J, Chipman AD, Chittka L, Fouks B, Liu J, O'Neill MP, Sumner S, Puiu D, Qu J, Salzberg SL, Scherer SE, Muzny DM, Richards S, Robinson GE, Gibbs RA, Schmid-Hempel P, Worley KC. 2015. The genomes of two key bumblebee species with primitive eusocial organization. *Gen. Biol*. **16**:76.

95. Sun C, Huang J, Wang Y, Zhao X, Su L, Thomas GWC, Zhao M, Zhang X, Jungreis I, Kellis M, Vicario S, Sharakhov IV, Bondarenko SM, Hasselmann M, Kim CN, Paten B, Penso-Dolfin L, Wang L, Chang Y, Gao Q, Ma L, Ma L, Zhang Z, Zhang H, Zhang H, Ruzzante L, Robertson HM, Zhu Y, Liu Y, Yang H, Ding L, Wang Q, Ma D, Xu W, Liang C, Itgen MW, Mee L, Cao G, Zhang Z, Sadd BM, Hahn MW, Schaack S, Barribeau SM, Williams PH, Waterhouse RM, Mueller RL. 2021. Genus-wide characterization of bumblebee genomes provides insights into their evolution and variation in ecological and behavioral traits. *Mol. Biol. Evol.* **38**:486–501.

96. Näpflin K, Schmid-Hempel P. 2016. Immune response and gut microbial community structure in bumblebees after microbiota transplants. *Proc. Roy. Soc. Biol. Sci.* **283**.

97. Goulson D, Lye GC, Darvill B. 2008. Decline and conservation of bumble bees. *Annu. Rev. Entomol.* **53**:191–208.

98. Villa JD. 2004. Swarming behavior of honey bees (Hymenoptera: Apidae) in southeastern Louisiana. *Ann. Entomol. Soc. Am.* **97**:111–116.

99. Goulson D. 2003. *Bumblebees: behaviour, ecology, and conservation*. Oxford Biology.

100. Powell J. Elijah, Martinson Vincent G., Urban-Mead Katherine, Moran Nancy A. 2014. Routes of acquisition of the gut microbiota of the honey bee *Apis mellifera*. *App. Environ. Microbiol.* **80**:7378–7387.

101. Kwong WK, Moran NA. 2016. Gut microbial communities of social bees. *Nat. Rev. Microbiol*. **14**:374–84.

102. Kwong WK, Moran NA. 2015. Evolution of host specialization in gut microbes: the bee gut as a model. *Gut Microbes*. **6**:214–20.

103. Engel P, Martinson VG, Moran NA. 2012. Functional diversity within the simple gut microbiota of the honey bee. *Proc. Natl. Acad. Sci. USA*. **109**:11002–7.

104. Palmer-Young EC, Raffel TR, McFrederick QS. 2019. pH-mediated inhibition of a bumble bee parasite by an intestinal symbiont. *Parasitology*. **146**:380–388.

105. Horak RD, Leonard SP, Moran NA. 2020. Symbionts shape host innate immunity in honeybees. *Proc. Roy. Soc. Biol. Sci*. **287**:20201184.

106. Kwong WK, Mancenido AL, Moran NA. 2017. Immune system stimulation by the native gut microbiota of honey bees. *R Soc Open Sci*. **4**:170003.

107. Kwong WK, Engel P, Koch H, Moran NA. 2014. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci. USA*. **111**:11509–14.

108. Powell E, Ratnayeke N, Moran NA. 2016. Strain diversity and host specificity in a specialized gut symbiont of honeybees and bumblebees. *Mol. Ecol*. **25**:4461–4471.

109. Kwong WK, Medina LA, Koch H, Sing KW, Soh EJY, Ascher JS, Jaffé R, Moran NA. 2017. Dynamic microbiome evolution in social bees. *Sci. Adv*. **3**:e1600513.

110. Sauers LA, Sadd BM. 2019. An interaction between host and microbe genotypes determines colonization success of a key bumble bee gut microbiota member. *Evolution*. **73**:2333–2342.

111. Zheng H, Nishida A, Kwong WK, Koch H, Engel P, Steele MI, Moran NA. 2016. Metabolism of toxic sugars by strains of the bee gut symbiont *Gilliamella apicola*. *mBio*. **7**.

112. Olofsson TC, Alsterfjord M, Nilson B, Butler È, Vásquez A. 2014. *Lactobacillus apinorum sp. nov., Lactobacillus mellifer sp. nov., Lactobacillus mellis sp. nov., Lactobacillus melliventris sp. nov., Lactobacillus kimbladii sp. nov., Lactobacillus helsingborgensis sp. nov.* and *Lactobacillus kullabergensis sp. nov.*, isolated from the honey stomach of the honeybee *Apis mellifera*. *Int. J. Syst. Evol. Microbiol*. **64**:3109–3119.

113. Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA. 2011. A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol. Ecol*. **20**:619–628.

114. Ellegaard KM, Brochet S, Bonilla-Rosso G, Emery O, Glover N, Hadadi N, Jaron KS, van der Meer JR, Robinson-Rechavi M, Sentchilo V, Tagini F, Engel P. 2019. Genomic changes underlying host specialization in the bee gut symbiont *Lactobacillus Firm5*. *Mol. Ecol*. **28**:2224–2237.

115. Kwong WK, Zheng H, Moran NA. 2017. Convergent evolution of a modified, acetate-driven TCA cycle in bacteria. *Nat. Microbiol*. **2**:17067.

# CHAPTER II: THE DIVERGENCE AND CONVERGENCE OF THE GUT MICROBIOMES OF WILD INSECT POLLINATORS

**This chapter is modified from the under-review article:** Li, J.*, Sauers, L.* et al. Divergence and convergence of the gut microbiomes of wild insect pollinators. *mBio*. *Joint first authors

## Abstract

Pollination services provided by wild insect pollinators are critical to natural ecosystems and crops around the world. There is an increasing appreciation that the gut microbiota of these insects influences their health and consequently their services. However, pollinator gut microbiota studies have focused on well-described social bees, but rarely include other, more phylogenetically divergent insect pollinators. To expand our understanding, we explored the insect pollinator microbiomes across three insect orders through two DNA sequencing approaches. First, in an exploratory 16S amplicon sequencing analysis of taxonomic community assemblages, we found lineage-specific divergences of dominant microbial genera and microbiota community composition across divergent insect pollinator genera. However, we found no evidence for a strong broad-scale phylogenetic signal, which we see for community relatedness at finer scales. Subsequently, we utilized metagenomic shotgun sequencing to obtain metagenome assembled genomes and assess the functionality of the microbiota from pollinating flies and social wasps. We uncover from pollinating flies a novel gut microbe in the family Orbaceae that is closely related to *Gilliamella spp.* from social bees but with divergent functions. We propose this novel species be named *Candidatus Gilliamella eristali*. Further metagenomes of dominant fly and wasp microbiome members suggest they are largely not host-insect adapted and instead may be environmentally derived. Overall, this study suggests selective processes involving ecology or

physiology, or neutral processes determining microbe colonization may predominate in the turnover of lineages in insect pollinators broadly, while evolution with hosts may occur only under certain circumstances and on smaller phylogenetic scales.

**Introduction**

Wild insect pollinators provide critical ecosystem services, being crucial to the maintenance of both wild and agricultural plant communities (1, 2). Although the focus is often on managed honey bees as pollinators, thousands of wild bee and other insect species provide efficient and under-valued pollination services (3–7). Insect species in general are undergoing significant declines (8–10), but there is particular concern about wild pollinator insects facing threats that could destabilize natural ecosystems (11) and limit important agricultural production (12). Therefore, studies investigating factors linked to insect pollinator health, and thus the services they provide, that extend beyond well-studied bee pollinators are needed.

Host-associated microbes and their specific functions may determine host niche specificity, survival, and fitness (13–15). In insects, beneficial symbiotic bacteria can play important roles in the development and health of their hosts (16–18). In particular, the gut microbiota of insects has received attention due to its potentially high functional importance (13–15). However, the associations between gut microbes and their hosts can be highly variable (19), spanning from highly specialized gut microbial communities, such as those of social bees (20, 21), to insects with transient, environmentally determined communities or hosts with few microbes resident in their guts at all (22, 23). Given their importance for interacting with their host and the environment, it is vital to understand the composition and diversity of gut microbiota and the mechanisms that could shape these communities across relevant groups of host species.

Important questions in microbial ecology remain relating to the relative contributions of host-associated and environmental factors in determining the composition of microbiomes and what mechanisms are driving microbial community structure and function (24). The microbes colonizing the gut may be determined by neutral or selective processes (25), which can depend upon stochastic processes, host ecology or evolutionary history. Diet has been shown to be an important factor determining the gut microbiota of mammals, with similar microbes inhabiting the guts of unrelated host but with similar diets (26). However, phylogeny and hence relatedness of hosts has also been shown to have a strong effect determining similarity in composition (27). Relationships between microbial communities that recapitulate the phylogeny of their hosts are termed examples of phylosymbiosis (28). While phylosymbiosis can emerge due to vertical transmission and co-diversification of symbionts and hosts, it can also emerge from ecological and physiological filtering that establish interactions anew each generation from environmental microbes (29). Defining the distributions of microbes and their predicted functions among host species will help to determine the ecological and evolutionary processes that may underlie the associations. Our understanding of host-associated gut microbiomes and their potential effects can benefit from investigations of similarities and differences in community membership, community structure, and predicted function across host species that share certain ecological characteristics, such as pollination.

Studies of insect pollinator gut microbiota have mainly been in the eusocial honey bees and bumble bees (30). In these species, the gut microbiota plays important roles in nutrition, detoxification, and resistance to parasite infection (13–15, 31). Associated with the corbiculate Apid bees is a relatively small

core set of gut bacterial symbionts, which are thought to have largely undergone coevolution with their hosts, facilitated by vertical transmission in a social setting (20). The dynamic nature of the microbiota composition over evolutionary time, with lineage turnover, is however also apparent in this Apid clade, as the core microbes *Snodgrassella* and *Gilliamella* are absent from the stingless bee genus *Melipona*, which hosts more environmental bacteria and bee-specific yeasts (32). Studies of microbiota structure and function have been rare in other insect pollinator clades, but there is also evidence of a phylogenetic signal underlying differences in the structure of butterfly microbiomes (33). Most effective insect pollinators are either partially pollinivorous, nectarivorous, or both, and such diet sharing could influence broader patterns of microbiome and insect pollinator host associations, as it is known that diet can have important links to the gut microbiota structure and function (34). Despite an increased appreciation for diverse insect pollinators, studies documenting broad-scale patterns of gut microbiome and host associations that may be suggestive of general patterns driving host microbial community composition and functioning are warranted.

In this study, using high throughput sequencing of 16S rRNA amplicons, we investigated the associated gut microbiomes across a wide phylogenetic representation of insect pollinators from China, spanning three holometabolous insect orders (Diptera, Lepidoptera, Hymenoptera). To derive patterns of community structure and infer potential mechanisms shaping the gut microbial communities, from this approach we determined: i) bacterial genera colonization with each analyzed host genus, ii) patterns of microbiome diversity within (alpha diversity) and between (beta diversity) the host insect pollinator genera, iii) if a broad phylogenetic signal underlies differences in the community composition across host

genera signifying broad-scale phylosymbiosis. Additionally, utilizing metagenomic shotgun sequencing we report three high quality metagenome assembled genomes (MAGs) for dominant microbial community members associated with the pollinator fly *Eristalis tenax* and the wasp *Vespa bicolor*, and infer functions based on their gene repertoires. These species were picked as they possessed clear, representative community members from the exploratory 16s analysis, and are taxonomically distant from well-studied Apidae species. Based upon these MAGs we propose that pollinating flies harbor a unique species of *Gilliamella*, which we propose to be called *Candidatus Gilliamella eristali*.

**Results**

The gut microbiomes of a total of 861 individuals across 34 insect pollinator species (belonging to three orders: Hymenoptera, Lepidoptera, and Diptera) and two outgroup comparison species (Hemiptera: *Aphis craccivora* and *Halyomorpha halys*) were analyzed. Species were identified by morphology and by COI gene barcoding, and unless specific phylogenetic distances based on COI sequences were calculated for analyses, phylogenetic relatedness refers to previous studies (35–41). Amplicon sequencing of the hypervariable V3-V4 region of the bacterial 16S rRNA was performed on individual whole-gut samples. We obtained a total of 41,656,064 high-quality reads, which were passed through the DADA2 assembly and filtering with a resulting 34,897±10,024 reads (mean ± s.d.) per sample. This resulted in reads being assigned to 27,887 amplicon sequence variants (ASVs). After filtering ASVs with taxonomic assignments belonging to Eukaryotes, chloroplast, mitochondria, or no successful assignment 26,669 ASVs remained for further analysis. Additionally, for subsequent analyses, host genera with fewer than 10 samples were excluded. This resulted in excluding 3 samples each from

the genera *Ceratina* and *Eristalinus*, 2 samples each from the genera *Eucera*, *Lasioglossum*, and *Nomia*, and 1 sample from the genus *Sapyga*. Thus, the family Halictidae is excluded from the analysis because it consisted of only 4 samples, and the family Sapygidae is excluded as it consisted only of a single sample.

***Identifying colonization of bacterial genera within hosts***

We identified associations between insect pollinator host genera and particular bacterial taxa, indicative of distinct ecological or evolutionary associations. Microbial OTUs were grouped into microbial genera and visualized based on relative abundance (Figure 1). We find results that conform with some previously described relationships between hosts and their microbiomes, confirming our 16s amplicon sequencing is capturing expected native associations. First, we find low diversity within the two outgroup phloem feeding Hemiptera, with the microbial genus *Pantoea* (family *Erwiniaceae*) dominating in the host genus *Halymorpha*, which has been noted in previous work (42). Within the insect pollinators we also observe some prior documented relationships. First, both bumble bee (*Bombus spp.*) and honey bee (*Apis sp.*) host microbiomes contain the core bacterial genera *Gilliamella*, *Snodgrassella*, and *Lactobacillus* (20, 32). The evolutionary relationships between these three genera and their Apid hosts has received substantial attention (43). Within the host genus *Osmia* there is a large abundance of *Saccharibacter* (family *Acetobacteraceae*), within *Trigona* there is *Bifidobacterium* and *Lactobacillus*, and within the host genus *Xylocopa* there are significant amounts of *Apibacter*, *Bifidiobacterium*, and *Lactobacillus*. These relationships have been described and documented in previous research (20, 44–46). Thus, we conclude that our data set accurately captures the relationships between the insect pollinator hosts sampled and their common gut communities. We uncover

previously undescribed associations, including the genera *Gilliamella, Lactobacillus,* and members of the family *Enterobacteriaceae*, such as *Cedecea*, associating with the fly genus *Eristalis*, the intra-cellular bacterium *Wolbachia* in the butterfly *Lobocla,* members of the family *Niesseriaceae* in *Trigona*, *Thyreus*, and *Vespa*, and members of the family *Orbaceae* in *Amegilla, Trigona, Xylocopa, Thyreus,* and *Megachile*. Additionally, we find the principally phytopathogenic genus *Lonsdalea* in samples from the Hymenopteran genus *Vespa*, which adds to previous isolated observations (47). We use a comparative genomics approach to further analyze the *Eristalis*-associated *Cedecea* and *Gilliamella* and *Vespa*-associated *Lonsdalea* to elucidate their possible functions and the nature of their relationships with the host taxa.

***Alpha and beta diversity of the pollinator microbiomes***

Alpha diversity measures differed across the insect pollinator genera, and host genus explains a significant proportion of the variation in gut microbiome structure. We calculated the richness for each host genus in addition to the Shannon index. We find significant differences ($p<0.001$) in both of these metrics across host genera (Figure 2). *Thyreus* had the highest Shannon index, while *Aphis* and *Halymorpha* both had the lowest index values. The richness from the two Hemipterans but also from *Apis*, *Bombus*, and *Xylocopa* is low compared to many of the other genera, demonstrating that the gut microbial communities of individuals from these genera are dominated by a few abundant microbial genera. The low alpha diversity measures for these species are supported by previous work demonstrating the core conserved gut communities of these genera. Patterns of beta diversity among the genera were assessed using Bray Curtis distances. A PERMANOVA was performed on the distances

between host genera and finds that there is significant partitioning of the distance variance across host genera (p=0.001). The distances between samples were visualized with an ordination plot which explained 19.6% of the variation and clearing shows clustering based on host genera (Figure 3). The pattern of structure based on host genera still holds true even with the removal of the *Apis* and *Bombus* samples (Figure 3 inset).

***Investigating association between microbiota structure and host genetic distance***

We find that gut microbiome structure is significantly determined by host genus as shown above, and our data shows known and new described relationships between certain related insect pollinators and bacterial taxa. For example, the well-established colonization of the bacterial genera *Gilliamella*, *Snodgrassella*, and *Lactobacillus* in Apid bees (20, 43, 48). However, across all the insect pollinator species analyzed we do not find any evidence of a broad-scale relationship between host phylogenetic relatedness and similarity of the gut microbiome structure. A Mantel test was used to analyze the relationship between average pairwise distances between host genera and host microbial community Bray Curtis distances. We see no evidence for a significant relationship between these measures (p=0.4807), suggesting no support for host relatedness determining the overall differences in gut microbiome structure across the divergent insect pollinators analyzed, which would be apparent through a phylogenetic signal. The relationship between host phylogenetic distance and associated gut microbiome structure was also not significant (p=0.51) when comparing within only bee samples, where we have a greater phylogenetic breadth of diversity than in some prior comparisons (21, 43, 48, 49).

***Metagenome Assembled Genomes (MAGs)***

Three high quality MAGs (completeness > 95% and contaminations below 1.5%) were obtained from the metagenomic sampling (Table 1). For wasp metagenomes the initial assembly produced 3,532 contigs with the longest contig length of 253,125bp and N50 of 3311bp before binning. The fly metagenome assembly produced 103,530 contigs with the longest contig length of 198,522bp and an N50 of 7,840bp before binning. After binning, a single bin was retrieved from the wasp assembly with 123 contigs, a N50 of 49,197bp, and a completeness of 96.72% and contamination of 1.378%. This was automatically classified as *Lonsdalea britannica*. From the fly assembly two high quality bins were retrieved. The first bin consisted of 158 contigs with a N50 of 42,155bp and a completeness of 99.79% and contamination of 0.614%, which was automatically classified as a *Cedecea* species. The second fly-associated bin consisted of 80 contigs with a N50 of 22,570bp and a completeness of 95.48% and contamination of 0.047%. This second bin was automatically classified as a *Gilliamella* species.

The wasp-associated *Lonsdalea* genome is 3,749,826bp in size with a GC content of 55.3%. This MAG contained 3,324 predicted open reading frames of which 2,977 were successfully annotated with a COG or KEGG function. Additionally, 51 tRNAs were identified covering 16 amino acids. The fly-associated *Cedecea* genome is 3,664,531 bp in size with a GC content of 53.4%. The genome is predicted to contain 3,457 genes with 2,958 genes annotated with a COG or KEGG function, with 43 identified tRNAs covering 17 amino acids. The fly-associated *Gilliamella* genome is 1,901,067 bp in size with a GC content of 39.1%. The MAG contained 1,738 predicted open reading frames with 1,599 being

successfully annotated with a COG or KEGG function. A total of 40 tRNAs were identified covering 18 amino acids (Table 1).

***Phylogenetics and comparative genomics***

To confirm the taxonomic predictions from the automatic classification and distinguish phylogenetic groupings, gene clusters were identified and phylogenies built including comparison genomes obtained for the NCBI Genome database. For the *Lonsdalea* analysis we utilized 7 genomes from four species within this genus. *Pectobacterium sp.* and *Escherichia coli* were used as outgroups. For *Cedecea*, analysis was carried out with 4 *Cedecea* species genomes with 2 *Klebsiella*. *Serratia marcescens* genomes were used as outgroups. Finally, for the fly associated *Gilliamella sp.* we utilized Orbaceae genomes from 15 *Gilliamella apicola*, 4 *Gilliamella apis*, isolated from honeybee and bumblebee species, 2 *Frishella*, 1 *Candidatus Schmidhempelia*, 1 *Orbus*, and 1 *Zophobihabitans*, with *Pseudomonas aeruginosa* as an outgroup. Each genome obtained was assembled at a contig level and classified as full genomes coverage. Genomes were then assembled into pan genome databases, from which a core gene set was utilized for genomic analysis. This resulted in 43 core gene clusters containing a predicted 473 genes from genomes for *Lonsdalea*, 119 core gene clusters containing a predicted 1,428 genes from the genomes for *Cedecea*, 2 core gene clusters with 52 genes from the *Orbaceae* genomes. From the phylogenetic trees we can see that our proposed *Candidatus Gilliamella eristali* clusters as a sister clade to honeybee and bumblebee *Gilliamella* species (Figure 4). The fly-associated *Cedecea* and wasp-associated *Lonsdalea* cluster within respective clades of these genera, with the wasp-associated

*Lonsdalea* nested within the *L. Britannica* clade. Additionally, these clusterings are supported by ANIb calculations for whole genome similarities.

***Metabolic Reconstructions***

For both *Lonsdalea* and *Cedecea* samples, the metabolic functions of these bins did not significantly differ from the genome functions of closely related individuals, as determined by metabolic reconstruction and manual curation (Tables S8 and S9).

Substantial differences existed between our identified *Candidatus Gilliamella eristali* genome and genomes of other distinct but related *Gilliamella* species. Specifically, the *Candidatus Gilliamella eristali* lacks all genes related to cellulose permease, in addition to genes related to further degradation of pectin components found in other *Gilliamella* species. Further, *Candidatus Gilliamella eristali* lacks genes related to cysteine transport across the cell membrane, sulfate transport, and nickel transport. While marked as complete by the Anvio metabolic reconstruction tool, manual curation finds that this bacterium also lacks an essential gene for glycolysis (6 -phosphofructokinase) and several genes for the pentose-phosphate pathways (L-ribokinase, L-arabinose isomerase, fructuronate reductase, L-gulonate 5-dehydrogenase) in strong contrast to the metabolic pathways of bee-associated *Gilliamella*. Interestingly, *Candidatus Gilliamella eristali* possesses a heme transporter, nitrate reductases, and a formamide conversion enzyme not found in other *Gilliamella* species (Figure 5). These findings are supported by the fact that many of these genes are found as being unique to *Candidatus Gilliamella eristali*, or present in several other *Gilliamella* species but not *Candidatus Gilliamella eristali* from synteny analysis.

**Discussion**

Insect pollinators provide essential ecosystem services, increasing yield of agricultural crops and preserving the biodiversity of wild flowering plants (7, 10). Host insect associations with microbes can influence their health (50) and ultimately determine their ability to provide efficient pollination services (51). Accumulating evidence suggests that gut microbes associated with insects can aid their hosts in digestion, detoxification, and pathogen defense (52, 53). However, such work on associations between insect pollinators and their gut microbes is mostly constrained to a few, well-researched, insect pollinators, with limited taxonomic representation. For example, most previous studies are limited to eusocial corbiculate bees, comprised of honeybees, bumblebees, and stingless bees (20). Here, we provide an exploratory examination of the gut microbial communities of diverse insect pollinators spanning the insect orders of Diptera, Lepidoptera, and Hymenoptera. From this exploratory analysis we find significant differences in the gut community structure of these pollinators' microbiomes, with host genus explaining a large proportion of the variation in structure between samples, despite substantial among individual variation. We find abundant indicator microbes, such as *Snodgrassella* and *Gilliamella* in *Apis* hosts, *Gilliamella* in *Eristalis*, and *Lonsdalea* in *Vespa.* Subsequently analysis of gene repertoires suggest that these strongly associated microbes vary in apparent adaptation to the hosts, including not differing substantially from environmental derived bacteria in some cases. This signifies the dominance of tightly host-associated microbes seen in the Apid bees is not ubiquitous. Although similarities in microbiome structure exist between closely related genera, such as honey bees (*Apis*) and bumble bees (*Bombus*), there is no evidence of a strong relationship between gut microbiome structure and the

phylogenetic relatedness of the hosts on the broad-scale investigated. Surprisingly, we found a bacterium assigned as *Gilliamella* to be highly abundant not only in bees, but also in the pollinator fly *Eristalis.* We used metagenomic approaches to comparatively assess the phylogenetic clustering and genetic repertoires to putatively assess function and the host-microbe relationship.

From the colonization levels of these microbes alone, we see strong divergence between taxonomically distinct insect pollinators in their gut-associated microbes, yet also some similarities, even in pollinators of different orders. The colonization abundances of microbial genera in host species and genera in our analysis supports previous findings for well-studied species (20, 32, 42–47), in addition to adding information about host-associated gut microbes for lesser studied species. These include *Gilliamella* and *Lactobacillus* found to be associated with the fly genus *Eristalis*, which is interesting given that these bacteria have otherwise been associated with distinctly taxonomically unrelated bee host species in this and other studies (20). A further intriguing association is that of the dominance of *Lonsdalea* samples from the Hymenopteran *Vespa bicolor*, the black shield wasp. While acknowledged as an important pollinator (54), diverse feeding relationships in this and related species may explain this association. *Lonsdalea* is a well described phytopathogen, particularly of trees (55), and it could be that its presence comes from feeding on tree sap that has been documented in this insect genus (56). There is other limited documentation of the *Vespa-Lonsdalea* association (47), but it is unclear, however, if the relationship could be one where the insect is acting as an alternative host (57). As the assembled metagenome of the wasp-associated *Lonsdalea* nests within a clade of plant pathogens, it seems

unlikely that it is adapted to the wasp host, but raises the possibility of these functionally diverse insects acting as vectors of plant pathogens.

Alpha diversity measures also support differences in the gut microbiomes across the insect pollinators, from low richness and diversity in the butterfly *Lobocla* to high richness and diversity in the cleptoparasitic bee *Thyreus.* It is well known that the social Apid bees have a relatively small, consistent and equal set of associated and evolutionary specialized gut bacteria (19), which is reflected in the diversity measures. The diversity of the gut bacterial communities of insects can be determined by diet, habitat, and phylogeny (58). All the assayed insect pollinators have greater diversity than the Hemipteran reference outgroups, which principally feed on plant sap. There is also the possibility that diet diversity and ecological and evolutionary interactions with microbes explains some of the differences in diversity between insect pollinators. Relatively low diversity is found in the predominantly nectar feeding Lepidopterans. The cleptoparasitic *Thyreus* cuckoo bee's high richness and diversity could directly or indirectly come about through their interactions with their unrelated hosts and their food stores, with female *Thyreus* laying eggs in brood cells of *Amegilla spp.* hosts before larvae emerge and consume their food provisions. Any associations between microbiome diversity and diet breadth could have an adaptive evolutionary origin, with a more diverse microbiota offering greater metabolic potential, or could result from environmental acquisition of microbes determined by diet (59). Some of the relationships identified here merit further in-depth comparative analysis to strengthen potential associations between diet and life-style of gut microbial community metrics.

Analysis of beta diversity, or dissimilarity of the gut bacterial communities across the insect pollinator genera, supported observations made from the relative abundances of microbes. We find a significant effect of host genera which indicates that these different pollinator species harbor distinct microbial communities. The re-emphasizes that the sharing of the ecological function of pollinators, including broad diet overlap as either partial nectivores and/or pollinivores is not sufficient to result in convergence of general microbiome community features. The distinctness of these insect pollinator host gut bacteria communities may be driven by several potential mechanisms that could act in concert or mutually exclusively on ecological and evolutionary scales. These can include neutral processes, such as the random exposure to environmental microbes, or selective processes, such as physiological filtering (25). Diet and microbiomes, as outlined above, may also be linked (25–27, 34). Phylogeny and host genetic distance may also affect dissimilarity in gut microbial community composition (27, 29, 60). Phylosymbiosis, with closely related host species harboring similar symbiont communities, appears stronger for internal host-associated microbes (29). This pattern could result from vertical transmission leading to long-term coevolution, co-speciation, and co-diversification across host lineages, as proposed within social bees (20, 21), but ecological filtering by phylogenetically determined host traits may be an alternative explanation (28, 29). In fact, ecological filtering has been suggested to contribute to significant phylogenetic structuring of gut microbiomes of butterflies (33, 61). Clear close relationships between the microbiomes of the Apid bees are apparent in our data set, but across the insect pollinator host genera sampled we find no association between pairwise genetic distances of the hosts and the dissimilarity of their gut microbial communities. Therefore, it appears that findings that composition of

the microbial communities of animals can be closely associated with host evolutionary history across wide ranging timescales and diverse systems (60), does not extend to insect pollinators across the different orders investigated here. This could result from a neutral turnover of microbial lineages masking any phylogenetic signal between more distantly related taxa or ancestral switches in ecology, such as diet or habitat, driving selective changes in the microbiota. For example, it is likely that the ancestor of bees was predatory (62), meaning that nectarivory and pollinivory shared with distantly related insect pollinators such as the Dipteran *Eristalis* are not ancestral traits linking these lineages. As a result, host phylogenetic signals of gut microbiome structure may be restricted to finer phylogenetic scales, such as those found within Apid bees (20, 21) and butterflies (33, 61).

Following from our exploratory analysis of microbial community taxonomy and diversity we followed up with metagenomic sequencing to obtain metagenome assembled genomes of dominant microbial community members from our sampled pollinators fly and social wasp. The pollinating fly *Eristalis* was of particular interest because we found them to possess an abundant member of the *Orbaceae,* for which we have subsequently proposed as a novel species *Candidatus Gilliamella eristali*. These flies are important non-bee pollinators (4) that also feed on nectar and pollen as adults. Wasps were selected for further analysis as social living styles have the potential to allow for more long-term host microbe associations which may drive adaptation to the host gut environment, but their more diverse ecological roles, including diet, may also play a role in determining their microbe relationships. Host adapted microbes have previously been characterized by smaller or reduced genome sizes, low

G+C content, and unique transport, adhesion, or virulence genes allowing persistence in the gut environment (63, 64).

*Gilliamella spp.* are one of the core gut microbes of corbiculate bee, including in honey and bumble bees, and this microbe has been shown to have important functions in digestion and detoxification (24, 53). The finding of a closely related bacterium in our pollinator fly samples is intriguing and could indicate convergent membership in phylogenetically divergent but ecologically somewhat similar insect pollinator species. However, the comparative genomic analysis of the fly and bee *Gilliamella* indicates potential functional differences that may be influenced by host physiology or ecology, or the presence of other microbiota members. We found that *Gilliamella* from the bee species but not *Gilliamella* from flies have genes to digest pectin and transport cellulose into the cell. Additionally, the *Gilliamella* from bees possess genes associated with glycolysis, the pentose-phosphate pathway, flagella, sulfate and nickel transport, and cysteine transport that are absent in *Candidatus Gilliamella eristali*. The lack of flagellar and cysteine-based proteins are especially interesting given the critical roles they have been shown to play in biofilm formation (65–67), and the lack of several critical genes in the oxidative portion of the pentose-phosphate pathway along with lack of 6-phosphofructokinase may point towards a more fermentative based role in the fly gut. However, *Candidatus Gilliamella eristali* possesses genes for heme transportation, nitrate reduction, and formamide conversion not found in bee associated *Gilliamella.* This suggests potentially important functional gene loss or gain between the *Gilliamella* isolated from flies and from bees. Gene loss could have occurred in these *Gilliamella* lineages due to evolution with their respective hosts and other co-occurring bacteria. For example, within Apid bees

cross-feeding between *Snodgrassella alvi* and *Gilliamella* has been shown to be a core feature of these communities, with a focus on iron, amino acids, and pyrimidines (48, 68). Thus, in the fly host where a co-symbiont such as *S. alvi* is absent, cysteine and siderophores may not be as readily available in the environment, and adhesion to biofilm may not be as critical for maintenance. The lack of *S. alvi* biofilm may therefore explain the lack of flagellar and cysteine acquisition genes and the presence of several heme transporters.

Metagenome assembled genomes were also obtained for a *Lonsdalea sp.* from the *V. bicolor* wasp samples and a *Cedecea sp.* from the *E. tenax* fly samples. Here we focused on whether these genomes displayed any characteristics of host adaptation. We expected it was more likely to find signal of host adaptation dominant wasp-associated microbes, because social living and overlapping generations of eusocial insects has the potential to foster long-term host microbe associations. However, neither the fly nor wasp associated genomes displayed features that are expected to reflect host adaptation. Genome length, G+C content, and predicted metabolic functionalities were similar to those of closely related bacteria species and the genomes clustered phylogenetically within other described species for both *Londsalea* and *Cedecea* isolates. Thus, while these microbes were found at relatively high abundances within the guts of these insect pollinators, it is unlikely that they are necessarily adapted to these insect hosts. In fact, as discussed previously with the case of *Lonsdalea sp.* in *V. bicolor*, they could represent phytopathogens or other environmental microbes. This suggests that host adaptation of dominant microbes may be the exception rather than the rule in insect pollinator gut communities, as highlighted by previous work (69), with rather horizontally acquired and

environmentally derived communities being prevalent when considering more diverse insect pollinating hosts.

**Conclusion**

Wild insect pollinators from diverse holometabolous insect orders provide key services to ecosystems. Their gut associated microbes may influence their ecological roles and the hosts' health, making understanding features of microbiota structure and function their underlying driving mechanisms of high importance. However, most studies have been taxonomically restricted, focusing on insect pollinator microbiomes within insect species, genera, families, or rarely orders but not across more broad-scale phylogenetic ranges representing diverse insect pollinators. By doing so, we address ecological and evolutionary factors that may influence microbiome structure and function. We find insect pollinators harbor specific microbial communities, differing in bacterial taxonomy, alpha diversity, and beta diversity. Although we uncover previously described relationships of core microbes and related hosts, such as those in the Apid bees, we find no evidence that the compositions of these microbial communities correlate with host evolutionary histories across the broader scale. Thus, we conclude that while some pollinator species may harbor vertically transmitted symbiont communities leading to phylogenetic signals, overall larger timescales the community structure of insect pollinator microbiomes has arisen independent of host phylogenetics. Some degree of convergence between distantly related but somewhat ecologically similar taxa is suggested by the analyses and by the sharing of *Gilliamella spp.* related bacteria between distantly related bee and fly pollinators. However, comparative analysis indicates distinct functionality that could be driven by ecology and evolutionary history, and the extent

of the host-microbe association. This work broadens our understanding of the microbiota of wild insect pollinators. Further, it points towards the potential importance of ecological, physiological, and non-evolutionary filters in determining microbiome structure and function when considering microbiomes on a relatively large phylogenetic scale, which calls for future in-depth comparative analyses investigating these avenues in more depth.

**Materials and Methods**

***Sample collection and processing***

Samples were collected with nets in the Yunnan, Hainan, Sichuan provinces and Beijing of China from May to August 2015. All samples were alive when captured and were then stored at −80 °C. Host species identification was initially carried out by experienced field biologists with further confirmation based on analyzing the cytochrome c oxidase subunit I (COI) gene. COI sequences were first obtained from assembled contigs and then confirmed with Sanger sequencing. They were subsequently compared against the NCBI non-redundant nucleotide database and the BOLD database (http://www.boldsystems.org/) for host species identification and confirmation.

For each sample, the whole gut (including crop, midgut, ileum and rectum) was dissected out aseptically and homogenized. This homogenate was used for DNA extraction using DNeasy Blood & Tissue kit (QIAGEN, GmbH, Germany) according to the manufacturer's instructions. DNA samples were then further purified with a Qiagen QIAquick column and eluted in 30 µL Buffer EB (Qiagen, Hilden, GmbH). The final extracts were quantified using a Qubit dsDNA broad range assay (Invitrogen, Life

Technologies, Grand Island, NY, USA), and the resulting DNA samples were sent to the Shanghai Meiji for PCR amplification and sequencing.

***Amplicon and metagenomic sequencing***

The hypervariable V3-V4 region of the bacterial 16S rRNA gene was amplified with the primers 341F (5'-CCTAYGGGRBGCASCAG-3') and 806R (5'-GGACTACNNGGGTATCTAAT-3'). 20 µl PCR reactions were set up with 4µl 5×FastPfu Buffer, 2µl dNTPs (2.5mM), 0.8µl each primer, 0.4 µl FastPfu Polymerase and template DNA (10 ng). Reactions occurred in a GeneAmp® 9700 (ABI) thermocycler with 95°C for 5min, 27 cycles of denaturation at 95°C for 30s, annealing at 55°C for 30s, and elongation at 72°C for 45s, followed by an additional elongation at 72°C for 10min. A dissociation stage was performed at the end of the run for quality control. PCR products were detected by 2% agarose gel electrophoresis, and again purified using the QIAquick Gel Extraction Kit (QIAGEN). Library pools were constructed with equal amounts of each PCR product by using Truseq Nano DNA LT Sample Prep Kit (Illumina). These were amplified through paired-end sequencing on the Illumina MiSeq platform.

We performed metagenomic sequencing by selecting at random and mixing ten intestinal samples from the same species (*V. bicolor* and *E. tenax*, respectively) or from the genus *Apis* (five samples from each of *A. cerana* and *A. dorsata*). The paired-end 250bp sequencing strategy based on the Illumina MiSeq platform was adopted for metagenomic sequencing. At least 20Gbp of raw data for each mixed sample was obtained.

***Bioinformatic analysis***

Amplicon sequences were processed, analyzed, and filtered using the protocol described in Qiime 2-2022.2 moving pictures tutorial (70, 71). Due for the arguments laid out in McMurdie and Holmes (72) we did not rarefy our dataset and opted to utilize differential abundance analysis, despite this, when data is rarefied our results remain consistent. Adapters and barcodes were removed with *cutadapt* package in QIIME2. Sequences were trimmed, filtered, assembled, and chimeras removed with QIIME2 (version 2022.2) using the DADA2 package, sequences were truncated at 270bp and 200 bp for the forward and reverse sequences respectively as quality decreases(70, 73). The taxonomy of the ASVs was assigned using Naïve-Bayes automatic classification against the SILVA SSU database 138 full sequence database (74) with default parameters using Qiime 2-2022.2. Subsequent analysis was done in R using the phyloseq package, with reads associated with chloroplast, mitochondrial, and Eukaryotes filtered out of the data. We then utilized the identified ASV dataset and corresponding taxonomic assignments to construct the ASV count table with taxonomy data.

For metagenomic sequencing the MetaWrap pipeline (75) was used to identify high quality bins with above 80% completion and less than 10% contamination. First, reads were run through the quality control model, trimming adapters and poor-quality bases with Trim-galore and removing host reads with bmtagger packed into MetaWrap. Paired reads with only one read mapping to the host genome are also removed. Following quality control, the remaining reads were assembled into sample-based assemblies with metaSpades (76) and binned with Metabat2, Maxbin2, and concoct (77–79). The resulting bins were then refined with the MetaWrap refinement module and bin completion, and

contamination estimated with CheckM (80). Bins below 80% completion and above 10% contamination were removed.

***Microbial Community Analysis***

Based on genus abundance of gut bacteria, data was imported into R statistical software (81) using the phyloseq package (82). Phyloseq allows for community level analysis and additional package support for analyzing microbial metagenomic data. The phyloseq allowed to comparison of relative abundances to detect microbial genera that are associated with specific host genera. Shannon alpha diversity and Bray-Curtis beta diversity indices were calculated with phyloseq and the ordinate function was used to visualize the Bray-Curtis distances. PERMANOVA from the package adonis was utilized to determine whether the host genera had a significant effect on the Bray-Curtis distances. The packages seqinr, poppr, ape, and ggTree were used to work with the host COI gene data (83–85). Sequences were aligned with muscle and trimmed to even lengths. The host genetic distance was then calculated, and a phylogenetic tree constructed to visualize the phylogeny and ensure the COI genes were accurately capturing true host genetic relationships. Finally, to test for correlations between host genetic distance and microbiome Bray-Curtis average distances, a pairwise matrix across all host genera was constructed. This matrix excluded the outgroup species from *Aphis* and *Halyomorpha,* as the primary focus of this study is on the insect pollinator species. These matrixes were then used to perform a mantel test of 9,999 permutations using the R package ade (86).

***Metagenomic sequences analysis***

The created metagenomic bins were imported into Anvio (87) and taxonomically identified by utilizing the metawrap bin classification module which automatically searches 22 single-copy core genes and searching against the Genome Taxonomy Database to assign taxonomy to each metagenome assembled genome. Next, open reading frames were predicted with Prodigal and annotated utilizing a DIAMOND BLASTp search against both the NCBI COGs database and KEGG KOfam database (88–93). External genomes for pangenomic comparisons were retrieved from the NCBI database and additional information these genomes can be found in. With the Anvio interactive interface, for each of the three comparisons (putative *Lonsdalea*, *Gilliamella*, *Cedecea*) high-quality core genes were identified with the following search parameters: a max functional homogeneity of 0.9 and a minimum geometric homogeneity of 1.0, and present only once in every genome. These core genes were then utilized to aligned with muscle and a PhyML phylogenetic tree with bootstrapping was constructed on the phylogeny.fr platform (94–96). For the *Orbaceae* tree the interactive tree of life for further processing, annotation and figure creation (97). For the other trees ggTree was used to create supplemental figures (85). The metabolic functions of these metagenomes were then reconstructed with Anvio and complete functional pathways determined utilizing Anvio's predict metabolism function combined with manual curation (87).

**References**


1. Klein AM, Vaissière BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, Tscharntke T. 2007. Importance of pollinators in changing landscapes for world crops. *Roy Proc Biol Sci.* **274**:303–13.

2. Requier F, Pérez-Méndez N, Andersson GKS, Blareau E, Merle I, Garibaldi LA. 2023. Bee and non-bee pollinator importance for local food security. *Trends Ecol Evol.* **38**:196–205.

3. Garibaldi LA, Steffan-Dewenter I, Winfree R, Aizen MA, Bommarco R, Cunningham SA, Kremen C, Carvalheiro LG, Harder LD, Afik O, Bartomeus I, Benjamin F, Boreux V, Cariveau D, Chacoff NP, Dudenhöffer JH, Freitas BM, Ghazoul J, Greenleaf S, Hipólito J, Holzschuh A, Howlett B, Isaacs R, Javorek SK, Kennedy CM, Krewenka KM, Krishnan S, Mandelik Y, Mayfield MM, Motzke I, Munyuli T, Nault BA, Otieno M, Petersen J, Pisanty G, Potts SG, Rader R, Ricketts TH, Rundlöf M, Seymour CL, Schüepp C, Szentgyörgyi H, Taki H, Tscharntke T, Vergara CH, Viana BF, Wanger TC, Westphal C, Williams N, Klein AM. 2013. Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science.* **339**:1608–11.

4. Rader R. 2009. Alternative pollinator taxa are equally efficient but not as effective as the honeybee in a mass flowering crop. *Jour Appl Ecol.* **46**:1080–1087.

5. Page ML, Williams NM. 2023. Honey bee introductions displace native bees and decrease pollination of a native wildflower. *Ecology*. **104**:e3939.

6. Gustafson NW, Couture JJ, Dalgleish HJ. 2023. Herbivory, plant traits and nectar chemistry interact to affect the community of insect visitors and pollination in common milkweed, *Asclepias syriaca*. *Oecologia*. **201**:91–105.

7. Ollerton J. 2017. Pollinator diversity: distribution, ecological function, and conservation. *Ann Rev Ecol Evol Sys*. **48**:353–376.

8. Wagner DL. 2020. Insect Declines in the Anthropocene. *Annu Rev Entomol*. **65**:457–480.

9. Wagner DL, Grames EM, Forister ML, Berenbaum MR, Stopak D. 2021. Insect decline in the Anthropocene: Death by a thousand cuts. *Proc Natl Acad Sci USA*. **118**.

10. Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE. 2010. Global pollinator declines: trends, impacts and drivers. *Trends Ecol Evol*. **25:**345–353.

11. Biesmeijer JC, Roberts SP, Reemer M, Ohlemüller R, Edwards M, Peeters T, Schaffers AP, Potts SG, Kleukers R, Thomas CD, Settele J, Kunin WE. 2006. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science*. **313**:351–4.

12. Vanbergenm AJ. 2013. Initiative TIP. Threats to an ecosystem service: pressures on pollinators. *Front Ecol Environ*. **11**:251–259.

13. Steele MI, Motta EVS, Gattu T, Martinez D, Moran NA. 2021. The gut microbiota protects bees from invasion by a bacterial pathogen. *Microbiol Spectr*. **9**:e0039421.

14. Zheng H, Nishida A, Kwong WK, Koch H, Engel P, Steele MI, Moran NA. 2016. Metabolism of toxic sugars by strains of the bee gut symbiont *Gilliamella apicola*. *mBio*. **7**:e01326-16.

15. Zheng H, Powell JE, Steele MI, Dietrich C, Moran NA. 2017. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc Natl Acad Sci USA*. **114**:4775–4780.

16. Douglas AE. 2015. Multiorganismal insects: diversity and function of resident microorganisms. *Annu Rev Entomol*. **60**:17–34.

17. Onchuru TO, Javier Martinez A, Ingham CS, Kaltenpoth M. 2018. Transmission of mutualistic bacteria in social and gregarious insects. *Curr Opin Insect Sci*. **28**:50–58.

18. Lee JB, Park KE, Lee SA, Jang SH, Eo HJ, Jang HA, Kim CH, Ohbayashi T, Matsuura Y, Kikuchi Y, Futahashi R, Fukatsu T, Lee BL. 2017. Gut symbiotic bacteria stimulate insect growth and egg production by modulating hexamerin and vitellogenin gene expression. *Dev Comp Immunol*. **69**:12–22.

19. Engel P, Moran NA. 2013. The gut microbiota of insects - diversity in structure and function. *FEMS Microbiol Rev*. **37**:699–735.

20. Kwong WK, Medina LA, Koch H, Sing KW, Soh EJY, Ascher JS, Jaffé R, Moran NA. 2017. Dynamic microbiome evolution in social bees. *Sci Adv*. **3**:e1600513.

21. Koch H, Abrol DP, Li J, Schmid-Hempel P. 2013. Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. *Molec Ecol*. **22**:2028–2044.

22. Hammer TJ. 2017. Caterpillars lack a resident gut microbiome. *Proc Natl Acad Sci USA*. **114**:9641–9646.

23. Wong AC, Chaston JM, Douglas AE. 2013. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *ISME J*. **7**:1922–32.

24. Antwis RE, Griffiths SM, Harrison XA, Aranega-Bou P, Arce A, Bettridge AS, Brailsford FL, de Menezes A, Devaynes A, Forbes KM, Fry EL, Goodhead I, Haskell E, Heys C, James C, Johnston SR, Lewis GR, Lewis Z, Macey MC, McCarthy A, McDonald JE, Mejia-Florez NL, O'Brien D, Orland C, Pautasso M, Reid WDK, Robinson HA, Wilson K, Sutherland WJ. 2017. Fifty important research questions in microbial ecology. *FEMS Microbiol Ecol*. **93**.

25. Gaulke CA, Arnold HK, Humphreys IR, Kembel SW, O'Dwyer JP, Sharpton TJ. 2018. Ecophylogenetics clarifies the evolutionary association between mammals and their gut microbiota. *mBio*. **9**.

26. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, Henrissat B, Knight R, Gordon JI. 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*. **332**:970–4.

27. Groussin M, Mazel F, Sanders JG, Smillie CS, Lavergne S, Thuiller W, Alm EJ. 2017. Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. *Nat Communic*. **8**:14319.

28. Lim SJ, Bordenstein SR. 2020. An introduction to phylosymbiosis. *Roy Proc Biol Sci*. **287**:20192900.

29. Mazel F, Davis KM, Loudon A, Kwong WK, Groussin M, Parfrey LW. 2018. Is host filtering the main driver of phylosymbiosis across the tree of life? *mSystems*. **3**:e00097-18.

30. Kwong WK, Moran NA. 2016. Gut microbial communities of social bees. *Nat Rev Microbiol*. **14**:374–84.

31. Koch H, Schmid-Hempel P. 2011. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc Natl Acad Sci USA*. **108**:19288–92.

32. Cerqueira AES, Hammer TJ, Moran NA, Santana WC, Kasuya MCM, da Silva CC. 2021. Extinction of anciently associated gut bacterial symbionts in a clade of stingless bees. *ISME J*. **15**:2813-2816.

33. Hammer TJ, Dickerson JC, McMillan WO, Fierer N. 2020. Butterflies host characteristic and phylogenetically structured adult-stage microbiomes. *Appl Environ Microbiol*. **86**:e02007-20.

34. Figueroa LL, Maccaro JJ, Krichilsky E, Yanega D, McFrederick QS. 2021. Why did the bee eat the chicken? symbiont gain, loss, and retention in the vulture bee microbiome. *mBio*. **12**:e0231721.

35. Cardinal S, Straka J, Danforth BN. 2010. Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. *Proc Natl Acad Sci USA*. **107**:16207–11.

36. Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim JW, Lambkin C, Bertone MA, Cassel BK, Bayless KM, Heimberg AM, Wheeler BM, Peterson KJ, Pape T, Sinclair BJ, Skevington JH, Blagoderov V, Caravas J, Kutty SN, Schmidt-Ott U, Kampmeier GE, Thompson FC, Grimaldi DA, Beckenbach AT, Courtney GW, Friedrich M, Meier R, Yeates DK. 2011. Episodic radiations in the fly tree of life. *Proc Natl Acad Sci USA*. **108**:5690–5.

37. Danforth BN, Cardinal S, Praz C, Almeida EA, Michez D. 2013. The impact of molecular data on our understanding of bee phylogeny and evolution. *Annu Rev Entomol*. **58**:57–78.

38. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutel RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust J, Aberer AJ, Aspöck U, Aspöck H, Bartel D, Blanke A, Berger S, Böhm A, Buckley TR, Calcott B, Chen J, Friedrich F, Fukui M, Fujita M, Greve C, Grobe P, Gu S, Huang Y, Jermiin LS, Kawahara AY, Krogmann L, Kubiak M, Lanfear R, Letsch H, Li Y, Li Z, Li J, Lu H, Machida R, Mashimo Y, Kapli P, McKenna DD, Meng G, Nakagaki Y, Navarrete-Heredia JL, Ott M, Ou Y, Pass G, Podsiadlowski L, Pohl H, von Reumont BM, Schütte K, Sekiya K, Shimizu S, Slipinski A, Stamatakis A, Song W, Su X, Szucsich NU, Tan M, Tan X, Tang M, Tang J, Timelthaler G, Tomizuka S, Trautwein M, Tong X, Uchifune T, Walzl MG, Wiegmann BM, Wilbrandt J, Wipfler B, Wong TK, Wu Q, Wu G, Xie Y, Yang S, Yang Q, Yeates DK, Yoshizawa K, Zhang Q, Zhang R, Zhang W, Zhang Y, Zhao J, Zhou C, Zhou L, Ziesmann T, Zou S, Xu X, Yang H, Wang J, Kjer KM, Zhou X. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*. **346**:763–7.

39. Branstetter MG, Danforth BN, Pitts JP, Faircloth BC, Ward PS, Buffington ML, Gates MW, Kula RR, Brady SG. 2017. Phylogenomic insights into the evolution of stinging wasps and the origins of ants and bees. *Curr Biol*. **27**:1019–1025.

40. Peters RS, Krogmann L, Mayer C, Donath A, Gunkel S, Meusemann K, Kozlov A, Podsiadlowski L, Petersen M, Lanfear R, Diez PA, Heraty J, Kjer KM, Klopfstein S, Meier R, Polidori C, Schmitt T, Liu S, Zhou X, Wappler T, Rust J, Misof B, Niehuis O. 2017. Evolutionary history of the Hymenoptera. *Curr Biol*. **27**:1013–1018.

41. Kawahara AY, Plotkin D, Espeland M, Meusemann K, Toussaint EFA, Donath A, Gimnich F, Frandsen PB, Zwick A, Dos Reis M, Barber JR, Peters RS, Liu S, Zhou X, Mayer C, Podsiadlowski L, Storer C, Yack JE, Misof B, Breinholt JW. 2019. Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proc Natl Acad Sci USA*. **116**:22657–22663.

42. Bansal R, Michel AP, Sabree ZL. 2014. The crypt-dwelling primary bacterial symbiont of the polyphagous pentatomid pest *Halyomorpha halys* (Hemiptera: Pentatomidae). *Environ Entomol*. **43**:617–25.

43. Kwong WK, Moran NA. 2015. Evolution of host specialization in gut microbes: the bee gut as a model. *Gut Microbes*. **6**:214–20.

44. Keller A, Grimmer G, Steffan-Dewenter I. 2013. Diverse microbiota identified in whole intact nest chambers of the red mason bee *Osmia bicornis* (Linnaeus 1758). *PLoS One*. **8**:e78296.

45. Alberoni D, Gaggìa F, Baffoni L, Modesto MM, Biavati B, Di Gioia D. 2019. *Bifidobacterium xylocopae sp. nov*. and *Bifidobacterium aemilianum sp. nov*., from the carpenter bee (*Xylocopa violacea*) digestive tract. *Syst Appl Microbiol*. **42**:205–216.

46. Subta P, Yodsuwan P, Yongsawas R, In-On A, Warrit N, Panha S, Khongphinitbunjong K, Chantawannakul P, Attasopa K, Disayathanoowat T. 2020. Bacterial communities in three parts of intestinal tracts of carpenter bees. *Insects*. **11**:497.

47. Suenami S, Konishi Nobu M, Miyazaki R. 2019. Community analysis of gut microbiota in hornets, the largest eusocial wasps, *Vespa mandarinia* and *V. simillima*. *Sci Rep*. **9**:9830.

48. Kwong WK, Engel P, Koch H, Moran NA. 2014. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc Natl Acad Sci USA*. **111**:11509–14.

49. Sauers LA, Sadd BM. 2019. An interaction between host and microbe genotypes determines colonization success of a key bumble bee gut microbiota member. *Evolution*. 73:2333–2342.

50. Engel P, Kwong Waldan K, McFrederick Q, Anderson Kirk E, Barribeau Seth M, Chandler James A, Cornman RS, Dainat J, de Miranda Joachim R, Doublet V, Emery O, Evans Jay D, Farinelli L, Flenniken Michelle L, Granberg F, Grasis Juris A, Gauthier L, Hayer J, Koch H, Kocher S, Martinson Vincent G, Moran N, Munoz-Torres M, Newton I, Paxton Robert J, Powell E, Sadd Ben M, Schmid-Hempel P, Schmid-Hempel R, Song Se J, Schwarz Ryan S, vanEngelsdorp D, Dainat B. 2016. The Bee Microbiome: Impact on bee health and model for evolution and ecology of host-microbe interactions. *mBio*. **7**:e02164-15.

51. de Vega C, Álvarez-Pérez S, Albaladejo RG, Steenhuisen S-L, Lachance M-A, Johnson SD, Herrera CM. 2021. The role of plant–pollinator interactions in structuring nectar microbial communities. *Journ Ecol*. **109**:3379–3395.

52. Engel P, Martinson VG, Moran NA. 2012. Functional diversity within the simple gut microbiota of the honey bee. *Proc Natl Acad Sci USA*. **109**:11002–7.

53. Zheng H, Perreau J, Powell JE, Han B, Zhang Z, Kwong WK, Tringe SG, Moran NA. 2019. Division of labor in honey bee gut microbiota for plant polysaccharide digestion. *Proc Natl Acad Sci USA*. **116**:25909–25916.

54. Wei W, Wu H, Li X, Wei X, Lu W, Zheng X. 2019. Diversity, daily activity patterns, and pollination effectiveness of the insects visiting *Camellia osmantha*, *C. vietnamensis*, and *C. oleifera* in South China. Insects 10:98.

55. Li A, He W. 2019. Molecular aspects of an emerging poplar canker caused by. *Front Microbiol*. **10**:2496.

56. Sitz RA, Aquino VM, Tisserat NA, Cranshaw WS, Stewart JE. 2019. Insects visiting drippy blight diseased red oak trees are contaminated with the pathogenic bacterium. *Plant Dis*. **103**:1940–1946.

57. Nadarasah G, Stavrinides J. 2011. Insects as alternative hosts for phytopathogenic bacteria. *FEMS Microbiol Rev*. **35**:555–75.

58. Yun JH, Roh SW, Whon TW, Jung MJ, Kim MS, Park DS, Yoon C, Nam YD, Kim YJ, Choi JH, Kim JY, Shin NR, Kim SH, Lee WJ, Bae JW. 2014. Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Appl Environ Microbiol*. **80**:5254–64.

59. Brunetti M, Magoga G, Gionechetti F, De Biase A, Montagna M. 2022. Does diet breadth affect the complexity of the phytophagous insect microbiota? The case study of *Chrysomelidae*. *Environ Microbiol*. **24**:3565–3579.

60. Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR. 2017. Correction: phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoS Biol*. **15**:e1002587.

61. Ravenscraft A, Berry M, Hammer T, Peay K, Boggs C. 2019. Structure and function of the bacterial and fungal gut microbiota of Neotropical butterflies. *Ecol Mono*. **89**:e01346.

62. Sann M, Niehuis O, Peters RS, Mayer C, Kozlov A, Podsiadlowski L, Bank S, Meusemann K, Misof B, Bleidorn C, Ohl M. 2018. Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. *BMC Evol Biol*. **18**:71.

63. Martinson VG, Magoc T, Koch H, Salzberg SL, Moran NA. 2014. Genomic features of a bumble bee symbiont reflect its host environment. *Appl Environ Microbiol*. **80**:3793–803.

64. Siddaramappa S, Challacombe JF, Duncan AJ, Gillaspy AF, Carson M, Gipson J, Orvis J, Zaitshik J, Barnes G, Bruce D, Chertkov O, Detter JC, Han CS, Tapia R, Thompson LS, Dyer DW, Inzana TJ. 2011. Horizontal gene transfer in *Histophilus somni* and its role in the evolution of pathogenic strain 2336, as determined by comparative genomic analyses. *BMC Genomics*. **12**:570.

65. Valentin JDP, Straub H, Pietsch F, Lemare M, Ahrens CH, Schreiber F, Webb JS, van der Mei HC, Ren Q. 2022. Role of the flagellar hook in the structural development and antibiotic tolerance of *Pseudomonas aeruginosa* biofilms. *ISME J*. **16**:1176–1186.

66. Moran Brenner, Sivan Friedman, Adi Haber, Ilya Borovok, Nadejda Sigal, Oded Lewinson, Anat A. Herskovits. 2022. *Listeria monocytogene*s TcyKLMN cystine/cysteine transporter facilitates glutathione synthesis and virulence gene expression. *mBio*. **13**:5.

67. Soutourina O, Poupel O, Coppée J-Y, Danchin A, Msadek T, Martin-Verstraete I. 2009. CymR, the master regulator of cysteine metabolism in *Staphylococcus aureus*, controls host sulphur source utilization and plays a role in biofilm formation. *Molecul Microbio*l. **73**:194–211.

68. Kešnerová L, Mars RAT, Ellegaard KM, Troilo M, Sauer U, Engel P. 2017. Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol*. **15**:e2003467.

69. Hammer TJ, Sanders JG, Fierer N. 2019. Not all animals need a microbiome. *FEMS Microbiol Lett*. **366**:fnz117.

70. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Author Correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. **37**:1091.

71. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. **7**:335–6.

72. McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol*. **10**:e1003531.

73. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. **13**:581–3.

74. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nuc Aci Res*. **41**:D590-6.

75. Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*. **6**:158.

76. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. **27**:824–834.

77. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. **7**:e7359.

78. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. **32**:605–607.

79. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. *Nat Methods*. **11**:1144–1146.

80. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Gen Res*. **25**:1043–1055.

81. R Core Team. 2020. R: A language and environment for statistical computing. R: A language and environment for statistical computing. https://www.R-project.org/.

82. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. **8**:e61217.

83. Charif D, Lobry JR. 2007. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis, p. 207–232. *In* Bastolla, U, Porto, M, Roman, HE, Vendruscolo, M (eds.), Structural approaches to sequence evolution: molecules, networks, populations. Springer Berlin Heidelberg, Berlin, Heidelberg.

84. Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*. **20**:289–290.

85. Yu G. 2020. Using ggtree to visualize data on tree-like structures. *Curr Protoc Bioinformatics*. **69**:e96.

86. Dray S, Dufour AB. The ade4 Package: Implementing the duality diagram for ecologists. *J Stat Soft*. 22:1–20.

87. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen ÖC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjödin A, Scott JJ, Vázquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-omics with anvi'o. *Nat Microbiol.* **6**:3–6.

88. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 11:119.

89. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol*. **215**:403–10.

90. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nuc Acid Res*. **45**:D353–D361.

91. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. *Nuc Aci Res*. **36**:D480-4.

92. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. *Nuc Aci Res*. **42**:D199–D205.

93. Galperin MY, Wolf YI, Makarova KS, Vera Alvarez R, Landsman D, Koonin EV. 2021. COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nuc Aci Res*. **49**:D274-d281.

94. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-F, Guindon S, Lefort V, Lescot M, Claverie J-M, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nuc Aci Res*. **36**:W465–W469.

95. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nuc Aci Res*. **32**:1792–7.

96. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. **59**:307–321.

97. Letunic I, Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*. **23**:127–128.

98. Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, Tang B, Dong L, Ding N, Zhang Q, Bai Z, Dong X, Chen H, Sun M, Zhai S, Sun Y, Yu L, Lan L, Xiao J, Fang X, Lei H, Zhang Z, Zhao W. 2017. GSA: Genome sequence archive. *Gen Prot Bioinform*. **15**:14–18.

99. Members BDC. 2019. Database resources of the BIG data center in 2019. *Nuc Aci Res*. **47**:D8–D14.

**Table 1.** Metagenome assembled genome bins from the bee, fly, and wasp microbiota metagenome sequencing. Completeness and contamination were calculated using the CheckM plugin for MetaWrap. Predicted open reading frames were determined with Prodigal and annotation was done by matching against the KEGG, COG, and tRNA databases with Anvio.

| Bin | Completeness | Contamination | GC | Contigs | N50 | Size | Predict Genes | Annotated Genes | tRNAs | Predicted Taxonomy |
|---|---|---|---|---|---|---|---|---|---|---|
| Fly.Bin.1 | 99.79 | 0.614 | 53% | 158 | 42,155 | 3,664,531 | 3,457 | 2,958 | 43 | *Cedecea* |
| Fly.Bin.2 | 95.48 | 0.047 | 39% | 80 | 22,570 | 1,901,067 | 1,738 | 1,599 | 40 | *Gilliamella apicola* |
| Wasp.Bin.1 | 96.72 | 1.378 | 55% | 123 | 49,197 | 3,749,826 | 3,324 | 2,977 | 51 | *Lonsdalea britannica* |

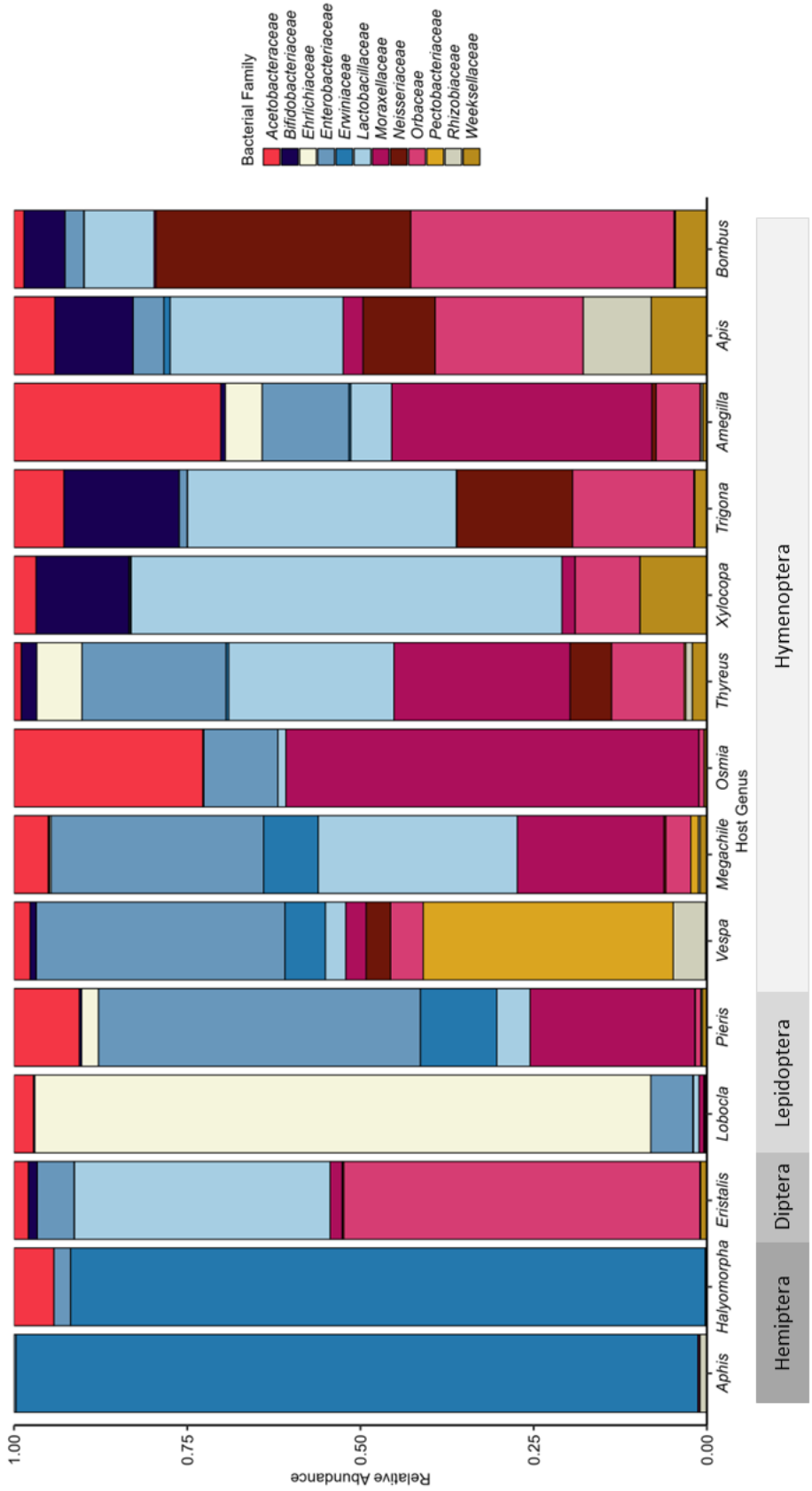


**Figure 1.** Per host genera relative abundance of microbial families with at least 1% abundance in the data set (>296,547 reads).

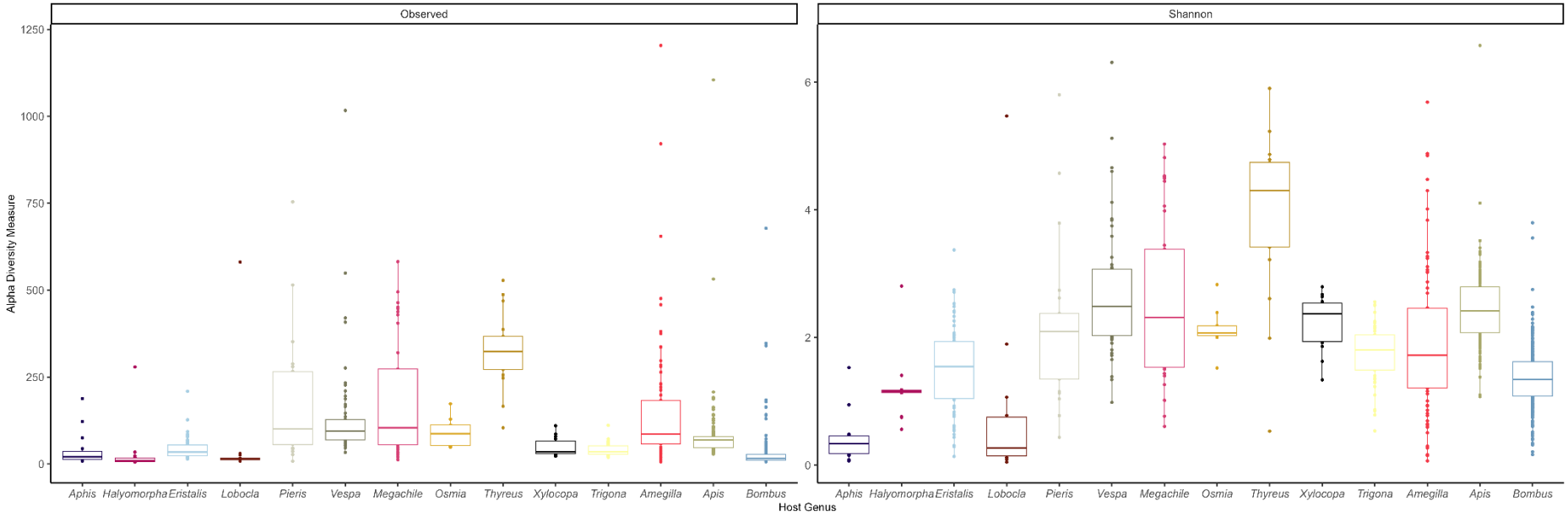


**Figure 2.** Observed species richness and Shannon alpha diversity index of gut communities for each host genus. Bars represent the medians, and the boxes show the interquartile ranges. Whiskers are the upper and lower values, with outliers shown as individual data points beyond these.

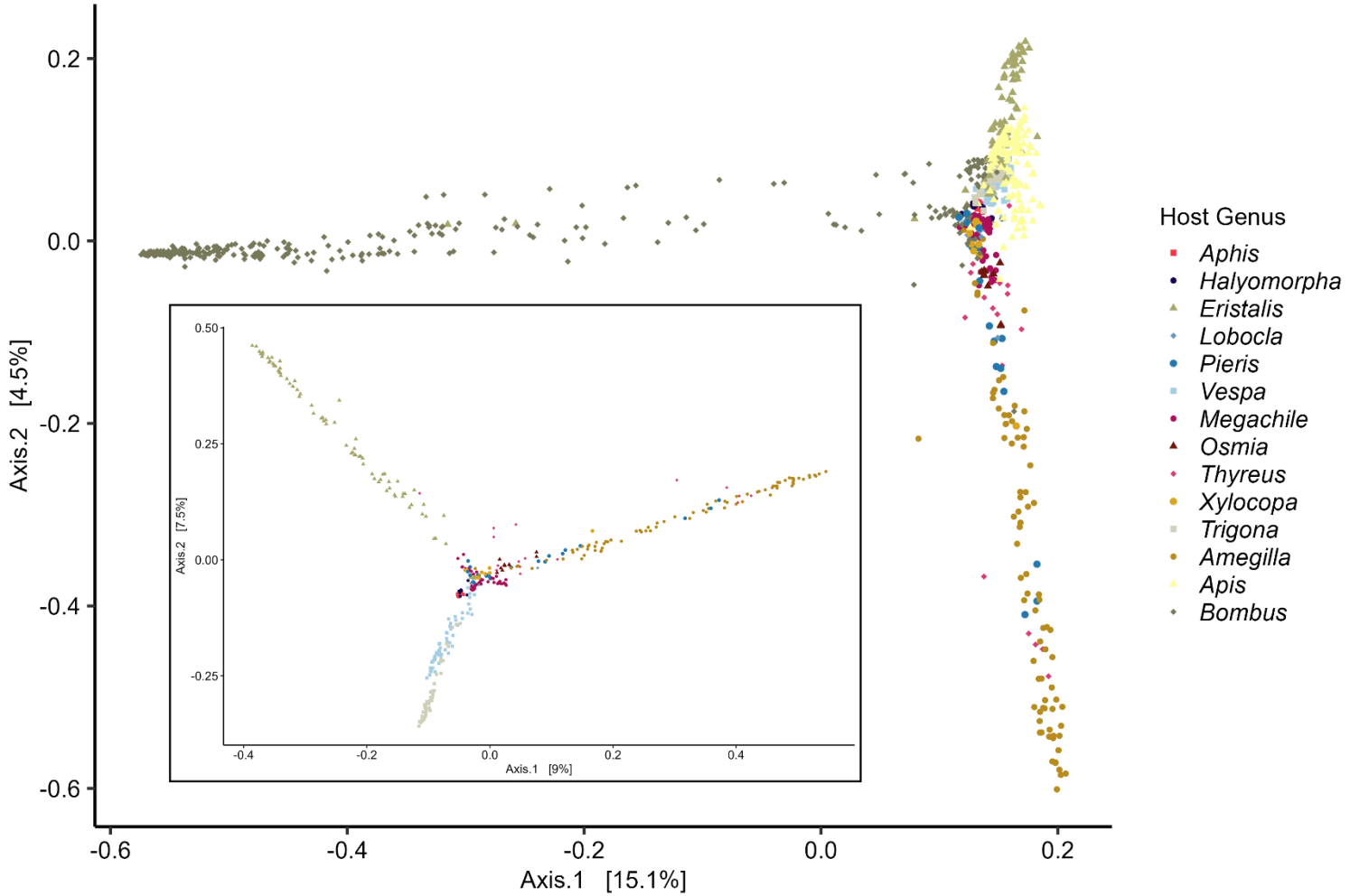


**Figure 3.** Ordination plot of the Bray-Curtis beta diversity distance among individual host microbiomes. Host genera are represented by shape and color combinations. The inset represents the same but with the genera *Apis* and *Bombus* removed.

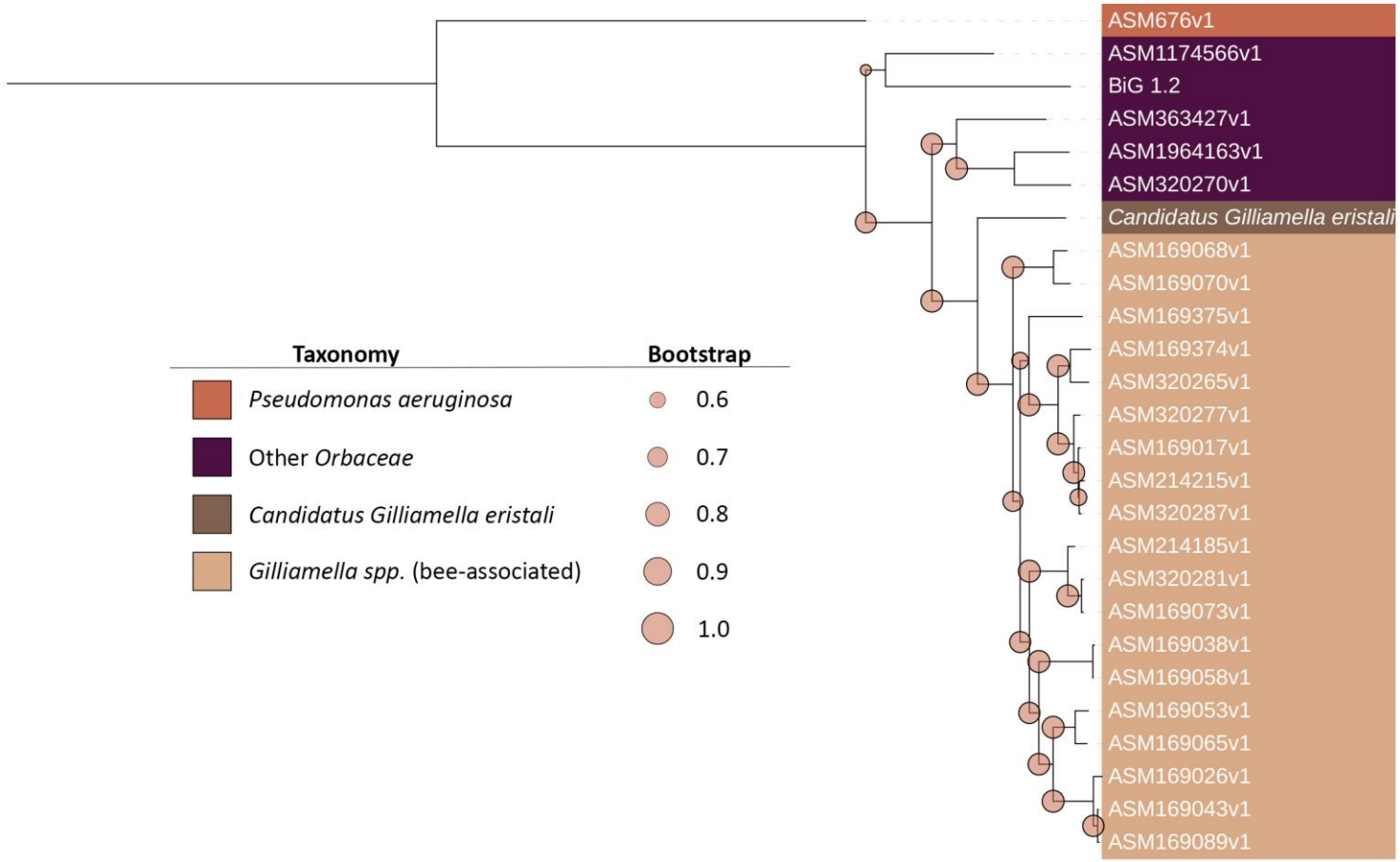


**Figure 4.** *Orbaceae* family phylogeny from selected core genes from the fly-associated metagenome assembled genome assigned as *Gilliamella sp.* and existing genomes. Coloring shows different species while the branch length shows amino acid substitution rates. *Pseudomonas aeruginosa* is used as an outgroup.

**Figure 5.**

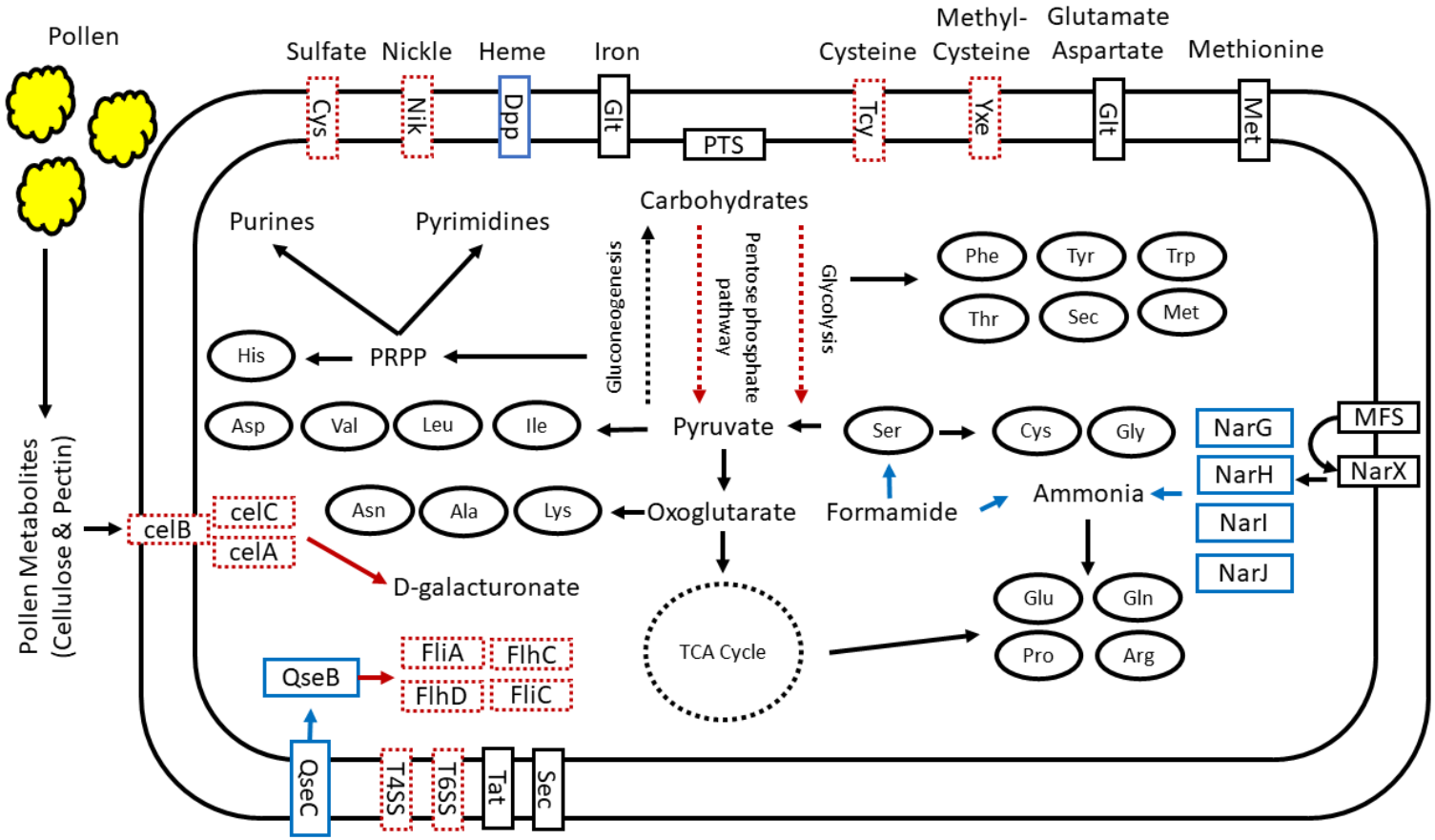


**Figure 5.** Fly-associated *Candidatus Gilliamella eristali* reconstructed metabolic pathways in comparison with other *Gilliamella* species. Solid black lines are genes/pathways shared by *Candidatus G. eristali* and other *Gilliamella* species. Dotted black lines are genes/pathways lacking across *Gilliamella* species. Dotted red lines are genes/pathways found in bee-associated *Gilliamella* but not *Candidatus G. eristali*. Blue solid lines are genes/pathways found in *Candidatus G. eristali* but no in bee-associated *Gilliamella* species.

# CHAPTER III: EXAMINING THE EXTENT OF COLONIZATION SPECIFICITY BETWEEN BUMBLE SPECIES AND A CORE GUT MICROBIOTA MEMBER.

**Abstract**

Hosts harbor complex microbial communities responsible for aiding the host in aspects of development, digestion, and defense. The relationships developed between hosts and their microbiota can lead to evolutionary dynamics that will in turn influence the colonization and functioning of these microbes. Previous work in social bees show that the core gut microbiota member *Snodgrassella alvi* appears to exhibit specificity in colonization of hosts at a genus level, with a honeybee (*Apis*) strain better colonizing their native host than a foreign host genus of bumble bees *(Bombus*), and vice versa. However*,* further work within a single bumble bee species has shown that colonization is governed by genotypic interactions. Here, we expand upon this previous work investigating whether genotypic interactions between *S. alvi* strains and host colonies may mask signals of broader scale host specificity and phylosymbiosis. We inoculated multiple isolated strains of *S. alvi* from honey bees and three bumble bee species (*B. auricomus*, *B. griseocollis* and *B. impatiens*) into germ-free hosts of *Bombus impatiens* and *Bombus griseocollis*. We find no support for genera nor species-level specificity, but further demonstrate that genotypic interactions have a strong effect on the colonization of *S. alvi* in the bumble bee gut. This finding has important implications for studies of host specificity and phylosymbiosis and indicates that consideration is needed for the level at which these studies investigate host-microbe associations.

**Introduction**

The appreciation for the benefits that microbial communities provide to their associated hosts continues to increase as research further expands our knowledge of these communities. Often, these communities are essential for aspects of host health and fitness, including digestion [1–3], development[4,5], and defense[6–8]. Additionally, the detrimental outcomes associated with perturbation of these microbial

communities or states of dysbiosis demonstrates the importance of understanding microbial colonization[9]. Thus, research has shifted to focusing on the benefits these communities provide to understanding the fundamental rules which govern their establishment and persistence.

The colonization of gut microbial communities is linked to several factors including the environment a host resides in[10,11], the life history of the host [11,12], and phylogenetic factors[13–21]. In humans microbiota is heavily influenced by method of birth, social contact, and diet[22,23]. Within bumble bees the microbiota varies as both individual workers and the colony ages[12]. Additionally, environmental, or chemical factors such as temperature or antibiotic exposure can greatly influence the establishment and stability of these important communities[9,24]. Finally, host and microbe phylogeny may play important roles in determining colonization, with hosts and microbes often displaying phylogenetic signals where the degree of microbiota similarity correlates with relatedness between species[8,14,16,18].

Previous work highlights the importance of understanding factors that influence when phylosymbiosis may or may not occur, as it sheds important light on the evolutionary and ecological processes driving microbiota establishment[14–16,25,26]. Several factors may help govern this phylosymbiotic signal, such as selection, drift, dispersal, and mode of transmission. Vertical or social transmission has been linked to host specificity of microbial species[25,27]. Furthermore, such modes of transmission are considered to have the potential to lead to finer scale interactions of genotype matching of host and microbe lineages[28], which have been demonstrated in certain relationships between hosts and microbiota members[8,29]. The level at which specificity occurs and its underlying drivers has important consequences for understanding and detecting patterns of host and associated microbe evolution. Thus, further studies investigating host microbe interactions and their influences on signatures of host phylosymbiosis and specificity at varying scales can advance our understanding of these patterns and their implications in the evolutionary and ecological establishment of host microbiota.

*Snodgrassella alvi* and its social bee hosts provide a viable model system for examining genotypic interactions and host specificity. Apid bees harbor a relatively simple and conserved core gut community, consisting of about 10-15 members[30]. Several of these members, including *S. alvi,* are thought to be coevolved with the bee host and other microbial community members[19,20,31]. The colonization and evolutionary relationships of *S. alvi* with its bumble bee host have previously been investigated. Past research shows that *S. alvi* strains isolated from bumble bees and honeybees tend to cluster into genus-specific clades and that strains appear to better colonize their native hosts genus than they do the foreign host genus[17,19,21]. However, previously, we have also shown that colonization of *S. alvi* is determined by genotypic interactions between the isolated bacterial strain and the host genotypic lineage that is delimited by the colony[29]. Here, we aim to expand on our understanding of the host-specificity of bumble bees and *S. alvi* and address if genotypic interactions may mask signals of host specificity when considering a wider range of microbial and host genotypes. We inoculate multiple strains isolated from honey bees (*Apis*) and different bumble bee species (*Bombus auricomus*, *B. griseocollis*, *B. impatiens*) into workers from distinct genotypic lineages (colonies) of *B. impatiens* and *B. griseocollis.*

**Materials and Methods**

***Standard bee rearing***

Five colonies of *Bombus impatiens* and sixteen colonies of *B. grisecollis* were raised from wild collected queens from the Mackinaw River Study Area (Lexington, IL. U.S.A.) in 2019. Queens were brought into the lab, and their feces of those of a selection of subsequently produced offspring workers were visually examined for the presence of common detectable pathogens utilizing phase contrast microscopy, and maintained under red-light in a climate-controlled room (26±1.5°C) and provided with sugar water (1 g cane sugar, 1 mL boiled tap water, and 0.1% cream of tartar, potassium bitartate, to partially invert the sugars) and pollen (Brushy Mountain Bee Farms, Moravian Falls, NC).

***Snodgrassella alvi* strain culturing, identification, and inoculation preparation**

Strains of *S. alvi* were isolated out of bees that were either field collected or derived from lab-reared colonies in 2017 or 2018. The strains utilized in this study, along with their bee species they were isolated from can be found in Table 2 (**Table 2**). Bees were initially isolated for 24 hours and fed with autoclaved sugar water and gamma-irradiated pollen *ad libitum*. This period was used to clear the gut of any transient microbes, allowing for ease in identifying *S. alvi* isolates on subsequent culturing. After this holding period bees were chilled on ice and their hindguts removed aseptically. Hindguts were placed into 200 µL of ringer saline with a 2.4 metal bead and homogenized in a Bead Ruptor (Omni International) on high for 30 seconds. Following homogenization, samples were serially diluted and plated onto brain heart infusion agar plates (Millipore Sigma [53286]) supplemented with food coloring to aid in visualization. These plates were incubated at 37°C and 10% $CO_2$ for 48 hours[32]. Isolated colonies considered to be *S. alvi* were identified, transferred into brain heart infusion broth, and again grown for 48 hours in the above conditions. After this time 400 mL of the culture was transferred into 400 mL of 50% glycerol and slow frozen (1°C per minute) to -80°C. Additionally, a 1mL aliquot of the bacterial media was frozen to -20°C for subsequent DNA analysis.

16S rRNA sequencing was used to confirm the identity of the *S. alvi* isolates. Cells were pelleted by centrifuging the 1 mL culture (8,000 rpm, 10 minutes, Micro 200R microcentrifuge, Hettich Zentrifugen) and then suspended in gram positive buffer (IBI Scientific gBAC Mini Genomic DNA Kit) with lysozyme. Subsequent DNA extraction followed the protocol provided by the IBI Scientific DNA extraction kit. 16S rRNA sequences were amplified with the 27F and 1492R primers previously used for *S. alvi* identification[32]. PCR was performed following a previously used protocol[29] and amplicons were sent to the University of Illinois Roy J Carter Biotechnology Center, Urbana-Champaign, Illinois. Sequences were manually curated with Sequencher software (Gene Codes Corp, Ann Arbor, MI.) and

BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to identify the closest related 16s sequences from the NCBI Genbank archive of 16S rRNA gene sequences.

Inoculums for the strains were created by culturing portions of the frozen glycerol stock onto brain heart infusion agar (1.5% agar) and growing for 48 hours at the standard cultivation parameters. After this period, isolated colonies of *S. alvi* were identified and subsequently transferred to liquid broth. These broth cultures were grown for 48 hours and then optical density used to estimate cell number per mL. Cultures were then diluted with sterile brain heart infusion agar to contain an estimated $10^7$ *S. alvi* cells per 10 µL and 10 µL of culture mixed with 10 µL of sugar water to form inoculums. Stocks were regrown every 2 days, and from each stock 10µL were taken and serially diluted. These serial dilutions were plated on brain heart infusion agar plates and grown for 48 hours. Colony forming units were counted and used to estimate the viable number of *S. alvi* cells in each inoculum to serve as a covariate in statistical analyses.

***Rearing and inoculation of germ-free bees***

Pupae from five colonies of *B. impatiens* and sixteen colonies of *B. grisecollis* were given a germ-free treatment and assigned to receive either one of twelve *S. alvi* inoculums or to remain un-inoculated. As *B. grisecollis* colonies are relatively small with few workers, we utilized workers, males, and queens to collect as many samples as possible. Additionally, due to the limited size of these colonies it was not always possible to inoculate every strain into each colony. Within *B. impatiens* both workers and males were utilized and all colonies of *B. impatiens* received all the *S. alvi* isolates.

Germ-free bees were created by taking spatially defined cohorts of bumble bee pupae. Last stage pupae in bumble bee colonies are generally spatially defined, making identification of individuals within this stage relatively simple. Pupae were removed from the colonies and placed into sterilized holding containers until a single individual emerged. This allows for age controlling the subsequent treatment, increasing the survival rate of the treated individuals, as in previous studies[29,33]. Pupal clumps

were subsequently submerged in a 3% bleach solution for 90 seconds. This method sterilizes the outer casing, and pupae shed their gut lining allowing sterile adults to emerge into a sterile environment. This technique has been successfully implemented previously to create germ free bees[29,33]. After receiving the bleach treatment, the treated clumps were placed into sterilized containers.

Following the germ-free treatment, emerging adult bees were isolated into individual sterilized holding vials. Individuals were randomly assigned an inoculation treatment and given the associated inoculum after two hours of isolation. Bees were visually monitored to ensure they completely consumed the inoculum and those that did not consume the inoculum within 60 minutes were discarded from the experiment. After consuming the inoculum, bees were placed into new individual, germ-free holding containers and fed autoclaved sugar water and gamma-irradiated pollen for four days. After four days, bees were chilled on ice and had their wings removed and hindguts aseptically dissected. Wings were imaged and then measured with ImageJ software as a proxy for size to be utilized as a covariate in subsequent analysis.

***qPCR quantification of colonization***

Extracted hindguts were used for DNA extraction following the procedure described above. Using the extracted DNA qPCR was performed to determine the number of *Snodgrassella* 16S rRNA gene copies from each bee using species-specific primers described in other work[3]. qPCR was done in 10 µL volume with a QuantStudio Flex 7 with PowerUp SYBR chemistry. Melt curves were generated at the end of each run to confirm single specific products and a synthetic standard (Integrative DNA Technologies gBLOCK) was used to create a standard curve. The standard curve generated from these standards had an efficiency of 89% and an intercept of 49.2, which closely matches the previous parameters for curves generated from these primers. Samples were compared to the standard curve to determine *Snodgrassella* 16S rRNA gene copies and converted to copies per bee. Samples with cycle Ct values greater than 37 were designated as zero copies, due to the unreliability of qPCR values above this

threshold that the germ-free bee samples had Ct levels around 37 (despite previously being shown to not be colonized[29]). This threshold means that our level of detection corresponded to $10^5$ gene copies per bee when converted based on the standard curve. This level of detection is consistent with the research in which the primers were originally described[3].

***Statistical analysis***

Statistical analyses were carried out on the abundance of 16S rRNA gene copies and the presence and absence of *Snodgrassella* colonization. When analyzing abundance models with truncated negative binomial (*nbinom2*) distributions were fit from the R package glmmTMB[34]. When analyzing presence and absence of colonization binomial models were fit using the R package lme4[35]. For all models, factors were tested to find the best model by combination of hypothesis testing and AICc. Additionally, model fits were investigated with the DHARMa package[36] and for the binomial models the dispersion parameter was calculated to ensure there was no overdispersion present. In all initial models wing size and inoculum size were used as covariates, while caste was included as a fixed effect. In some analyses strain and colony were used as random effects unless they were specifically the object of focus in the analysis (e.g. analysis of genotype-level effects). The means and standard errors were retrieved from the models with the emmeans package and graphed with the ggplot2 package[37,38].

We first analyzed our data for evidence of genus-level specificity. We analyzed *B. griseocollis* samples inoculated with strains isolated from *B. griseocollis* (native) or honeybee strains (foreign), and *B. impatiens* samples inoculated with strains isolated from *B. impatiens* (native) or honeybee strains (foreign). Here wing size was dropped as a covariate as it skewed the model fit, likely because of the naturally larger size of *B. griseocollis* individuals compared to *B. impatiens*.

To investigate whether there was any evidence for species level effects we analyzed both *B. impatiens* and *B. griseocollis* inoculated with *Snodgrassella* strains isolated from *B. impatiens, B.*

*griseocollis,* and *B. auricomus.* Bees receiving a strain isolated from conspecific samples were assigned as "native", and "foreign" when receiving a strain from a different host species.

Finally, we investigated the data set for any evidence of genotypic effects and interactions. To do this we utilized only *Bombus impatiens* samples, as only this species had significant replication at the colony level.

**Results**

We find no evidence for genus-level specificity in the presence or absence of colonization and the abundance of colonization (**Table 3**). We also find no significant effects of the covariates included in the model, with the best fitting model excluding any covariates. In *B. impatiens* and *B. griseocollis* strains of S. alvi isolated from those species did not colonize better in their native hosts than honey bee strains (**Figure 6**).

Investigating species-level specificity, there was no effect of the source and recipient host combination on the proportion of colonization (**Table 3**). There was also no effect of the source-recipient combination on colonization abundance, but there was a significant interaction between the host species and whether they received a native or foreign strain (p = 0036). *Bombus impatiens* native strains have slightly higher colonization than foreign strains, but this pattern is reversed for *B. griseocollis* (**Figure 7**). Additionally, there were no significant effects of covariates, however the best fitting model included the covariates caste and wing size.

Examining bumble bee derived strains inoculated into genotypic units (colonies) of *B. impatiens* there is a significant strain by colony interaction on the proportion colonized (p = 0.03) (**Figure 8 and Table 3**) and a significant effect *of S. alvi* strain on colonization abundance (p = 0.026). With the best fitting model included a significant effect of the covariate wing size (p = 0.007).

**Discussion**

Understanding if and at what level we see specificity in colonization of hosts by associated microbiota members offers important insights into the evolutionary dynamics of these systems. We test colonization into the bumble bees *B. impatiens* and *B.* griseocollis of multiple strains of a core gut microbiota member of Apid bees, *S. alvi*, isolated from hosts differing in their relatedness to the recipient, from the same species to a different genus. Overall, we find no evidence for genera level or species level specificity in our study. However, we find further support for the importance of microbe and host genotypes in determining colonization success. This highlights that the scale of specificity is important to consider for future studies on host specificity and phylogenetic signals between hosts and their associated microbiota members.

Our findings that host lineage and microbe strain are important factors determining colonization of *S. alvi* in *B. impatiens* corroborates that colonization specificity in this system appears at the level of genotypes, as evidenced by a prior demonstration of a host-genotype by microbe-genotype pattern[29]. This level of specificity could derive from vertical transmission of microbes within host lineages, with coevolution leading to differentiation of strains and host-microbe genotype matching driving colonization[28]. However, we do not explicitly test such matching here and such patterns of genotypic differences could also arise due to other evolutionary processes outside of coevolution, including genetic drift in host-microbe systems[39].

Despite the affirmation of differences in host and microbe genotypes within species determining colonization of the key gut bacterium *S. alvi* in bumble bees, we do not find evidence for a pattern of specific colonization that is determined by either host species or host genus. This is despite a previous report using a more limited set of bumble bee and honey bee *S. alvi* strains reporting specificity at the level of host genera determining colonization success[19]. Therefore, this does not appear to be a generalizable pattern. The lack of the pattern is interesting, given that honey bee and bumble bee

derived *S. alvi* form phylogenetic clades that are largely determined by their host[17]. However, this pattern of relatedness derived from non-colonization associated or neutral genetic markers and a lack of genus-level specificity are not mutually exclusive. If microbe transmission is largely restricted within host lineages, we expect to find patterns of more related microbes based on neutral markers within more related host clades. If colonization is driven by non-neutral processes within host lineages, that are not necessarily related to time of divergence, we would not necessarily expect patterns of colonization to agree with phylogenetic relatedness[40]. This supports the hypothesis that genotype level interactions that are determined by within host lineage evolution may mask broader-scale patterns of host-microbe phylosymbiosis on the species or genera levels.

The incongruence between our lack of evidence for genus-level specificity and earlier work supporting such a pattern[19] could derive from different potential sources but suggests that while broader-scale patterns of specificity of microbe colonization in the bee host-*S. alvi* association may be present in particular strain subsets, they are not ubiquitous. We use multiple strains of *S. alvi* in field derived colonies, whereas to the best of our knowledge the earlier work used commercial *B. impatiens.* Strains could also differ across locations and or different isolation approaches may bias strains used. For example, *Snodgrassella* from a specific *Bombus*-associated cluster (*Bombus* cluster 5) has been shown to have a relatively broad host range[17] and readily grows on the media utilized in this study[17,41]. Furthermore, as strains in other studies are often not only separated by the source host but also geographic location[17,21], which is not the case here, this could reinforce the patterns in those studies.

The strong influences of host and microbe genotype in determining colonization success of *Snodgrassella* seem to be a particular feature of the bumble bee microbiota system, potentially driven by predominant vertical transmission[25,27] and isolated evolution or coevolution within host lineages. We provide evidence here that specificity of microbe colonization at a fine phylogenetic scale driven by genotypic identity does not necessarily translate to broader scales like species and genera, likely

because it increases variability within those latter units in colonization. In some regard, this pattern is similar to the predicted high variation in infection among hosts closely related to the source host of a pathogen due differences in evolved resistance[42]. Our results also conversely mean that a lack of a pattern of specificity between hosts and microbes at a broader phylogenetic scale does not mean we can't rule out a finer scale of specificity. Additionally, this lack of a scaling up relationship indicates that underlying mechanisms determining patterns of colonization as genetic distance increases from a source host could vary, just as they are suggested to in host-pathogen relationships[41]. These are all important considerations in future investigations of phylosymbiosis and host-microbe specificity in microbiota studies and beyond.

**References**


1. Li, M. *et al.* Symbiotic gut microbes modulate human metabolic phenotypes. *Proc. Natl. Acad. Sci. USA.* **105**, 2117–22 (2008).
2. Zheng, H. *et al.* Metabolism of toxic sugars by strains of the bee gut symbiont *Gilliamella apicola*. *mBio* **7**, (2016).
3. Kešnerová, L. *et al.* Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol.* **15**, e2003467 (2017).
4. Visick, K. L., Foster, J., Doino, J., McFall-Ngai, M. & Ruby, E. G. *Vibrio fischeri* lux genes play an important role in colonization and development of the host light organ. *J. Bacteriol.* **182**, 4578–86 (2000).
5. Kwong, W. K., Mancenido, A. L. & Moran, N. A. Immune system stimulation by the native gut microbiota of honey bees. *R. Soc. Open. Sci.* **4**, 170003 (2017).
6. Koch, H. & Schmid-Hempel, P. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc. Natl. Acad. Sci. USA.* **108**, 19288–92 (2011).
7. Steele, M. I., Motta, E. V. S., Gattu, T., Martinez, D. & Moran, N. A. The gut microbiota protects bees from invasion by a bacterial pathogen. *Microbiol. Spectr.* **9**, e0039421 (2021).
8. Parker, B. J., Hrček, J., McLean, A. H. C. & Godfray, H. C. J. Genotype specificity among hosts, pathogens, and beneficial microbes influences the strength of symbiont-mediated protection. *Evolution.* **71**, 1222–1231 (2017).
9. Raymann, K., Shaffer, Z. & Moran, N. A. Antibiotic exposure perturbs the gut microbiota and elevates mortality in honeybees. *PLoS Biol.* **15**, e2001861 (2017).
10. Duplouy, A., Dotson, B. R., Nishiguchi, M. K. & Cárdenas, C. A. Editorial: symbiosis in a changing environment. *Front. Ecol. Evol.* **9**, (2021).

11. Yun, J. H. *et al.* Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Appl. Environ. Microbiol.* **80**, 5254–64 (2014).

12. Hammer, T. J., Easton-Calabria, A. & Moran, N. A. Microbiome assembly and maintenance across the lifespan of bumble bee workers. *Molec. Ecol.* **32**, 724–740 (2023).

13. Hammer Tobin, J., Dickerson Jacob, C., McMillan, W. O. & Fierer, N. Heliconius butterflies host characteristic and phylogenetically structured adult-stage microbiomes. *Appl. Environ. Microbiol.* **86**, e02007-20 (2020).

14. Lim, S. J. & Bordenstein, S. R. An introduction to phylosymbiosis. *Proc. R. Biol. Sci.* **287**, 20192900 (2020).

15. Kohl, K. D. Ecological and evolutionary mechanisms underlying patterns of phylosymbiosis in host-associated microbial communities. *Philos. Trans. R. Soc. Biol. Sci.* **375**, 20190251 (2020).

16. Pollock, F. J. *et al.* Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat. Commun.* **9**, 4921 (2018).

17. Powell, E., Ratnayeke, N. & Moran, N. A. Strain diversity and host specificity in a specialized gut symbiont of honeybees and bumblebees. *Molec. Ecol.* **5**, 4461–4471 (2016).

18. Mallott, E. K. & Amato, K. R. Host specificity of the gut microbiome. *Nat. Rev. Microbiol.* **19**, 639–653 (2021).

19. Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci. USA.* **111**, 11509–14 (2014).

20. Kwong, W. K. & Moran, N. A. Evolution of host specialization in gut microbes: the bee gut as a model. *Gut Microbes.* **6**, 214–20 (2015).

21. Koch, H., Abrol, D. P., Li, J. & Schmid-Hempel, P. Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. *Molec. Ecol.* **22**, 2028–2044 (2013).

22. Kolodziejczyk, A. A., Zheng, D. & Elinav, E. Diet–microbiota interactions and personalized nutrition. *Nat. Rev. Microbiol.* **17**, 742–753 (2019).

23. Dill-McFarland, K. A. *et al.* Close social relationships correlate with human gut microbiota composition. *Sci. Rep.* **9**, 703 (2019).

24. Hammer, T. J., Le, E. & Moran, N. A. Thermal niches of specialized gut symbionts: the case of social bees. *Proc. Roy. Soc. Bio. Sci.* **88**, 20201480 (2021).

25. Mazel, F., Guisan, A. & Parfrey, L. W. Transmission mode and dispersal traits correlate with host specificity in mammalian gut microbes. *Mol. Ecol.* (2023).

26. Mazel, F. *et al.* Is host filtering the main driver of phylosymbiosis across the tree of life? *mSystems* **3**, (2018).

27. Leeks, A., Dos Santos, M. & West, S. A. Transmission, relatedness, and the evolution of cooperative symbionts. *J. Evol. Biol.* **32**, 1036–1045 (2019).

28. Poisot, T., Bever, J. D., Nemri, A., Thrall, P. H. & Hochberg, M. E. A conceptual framework for the evolution of ecological specialisation. *Ecol. Lett.* **14**, 841–851 (2011).

29. Sauers, L. A. & Sadd, B. M. An interaction between host and microbe genotypes determines colonization success of a key bumble bee gut microbiota member. *Evolution.* **73**, 2333–2342 (2019).

30. Kwong, W. K. & Moran, N. A. Gut microbial communities of social bees. *Nat. Rev. Microbiol.* **14**, 374–84 (2016).

31. Kwong, W. K. *et al.* Dynamic microbiome evolution in social bees. *Sci. Adv.* **3**, e1600513 (2017).

32. Kwong, W. K. & Moran, N. A. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family *Neisseriaceae* of the Betaproteobacteria, and *Gilliamella apicola* gen. nov., sp. nov., a member of *Orbaceae* fam. nov., Orbales ord. nov., a sister taxon to the order 'Enterobacteriales' of the Gammaproteobacteria. *Int. J. Syst. Evol. Microbiol.* **63**, 2008–2018 (2013).

33. Näpflin, K. & Schmid-Hempel, P. Immune response and gut microbial community structure in bumblebees after microbiota transplants. *Proc. Biol. Sci.* **283**, (2016).

34. Team, Rs. RStudio: Integrated development environment for R. (2019).

35. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-Effects models using lme4. *J. Stat. Soft.* **67**, 1–48 (2015).

36. Hartig, F. *DHARMa: Residual diagnostics for hierarchical (multi-level / mixed) regression models*. (2016).

37. Lenth, R. V. *emmeans: estimated marginal means, aka least-squares means*. (2023).

38. Wickham, H. *Ggplot2: elegant graphics for data analysis*. (Springer International Publishing, 2016).

39. McCutcheon, J. P. & Moran, N. A. Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**, 13–26 (2012).

40. Sarton-Lohéac, G. *et al.* Deep divergence and genomic diversification of gut symbionts of neotropical stingless bees. *mBio* **14**, e03538-22 (2023).

41. Cornet, L. *et al.* Phylogenomic analyses of *Snodgrassella* isolates from honeybees and bumblebees reveal taxonomic and functional diversity. *mSystems* **7**, e0150021 (2022).

42. Antonovics, J. *et al.* The origin of specificity by means of natural selection: evolved and nonhost resistance in host-pathogen interactions. *Evolution.* **67**, 1–9 (2013).

**Table 2.** Information on the strains used in this study and the bee species they were isolated from.

| Sample ID | Host Species | Collection Location | Year Isolated |
|---|---|---|---|
| I1 | *B. impatiens* | Lexington, IL | 2018 |
| I2 | *B. impatiens* | Lexington, IL | 2018 |
| I49 | *B. impatiens* | Lexington, IL | 2018 |
| A19 | *B. auricomus* | Lexington, IL | 2018 |
| A102 | *B. auricomus* | Lexington, IL | 2018 |
| AC | *B. auricomus* | Lexington, IL | 2018 |
| G1 | *B. griseocollis* | Lexington, IL | 2018 |
| G3 | *B. griseocollis* | Lexington, IL | 2018 |
| G105 | *B. griseocollis* | Lexington, IL | 2018 |
| H1 | *A. mellifera* | Lexington, IL | 2018 |
| H2 | *A. mellifera* | Lexington, IL | 2018 |
| H3 | *A. mellifera* | Lexington, IL | 2018 |

**Table 3.** Factors and relevant statistics for the final models of *Snodgrassella alvi* colonization. Significant terms ($p<0.05$) are bolded.

| ***Genus-Level Specificity*** | | | |
|---|---|---|---|
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Status (Conspecific/Honey bee) | 2.393 | 1 | 0.122 |
| Species | 3.190 | 1 | 0.074 |
| Status * Species | 0.005 | 1 | 0.943 |
| ***Species-Level Specificity*** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Status (Conspecific/Other bumble bee) | 0.585 | 1 | 0.444 |
| Species | 1.410 | 1 | 0.235 |
| Caste | 2.841 | 2 | 0.242 |
| Wing size | 1.035 | 1 | 0.309 |
| **Status * Species** | **4.386** | **1** | **0.036** |
| ***Genotype x Genotype Analysis*** | | | |
| *Parameter* | *F-statistic* | *Df* | *p-value* |
| **Strain** | **2.285** | **8** | **0.026** |
| Colony | 0 | 4 | 1.00 |
| **Wing size** | **6.965** | **1** | **0.007** |
| **Strain * Colony** | **1.646** | **32** | **0.03** |

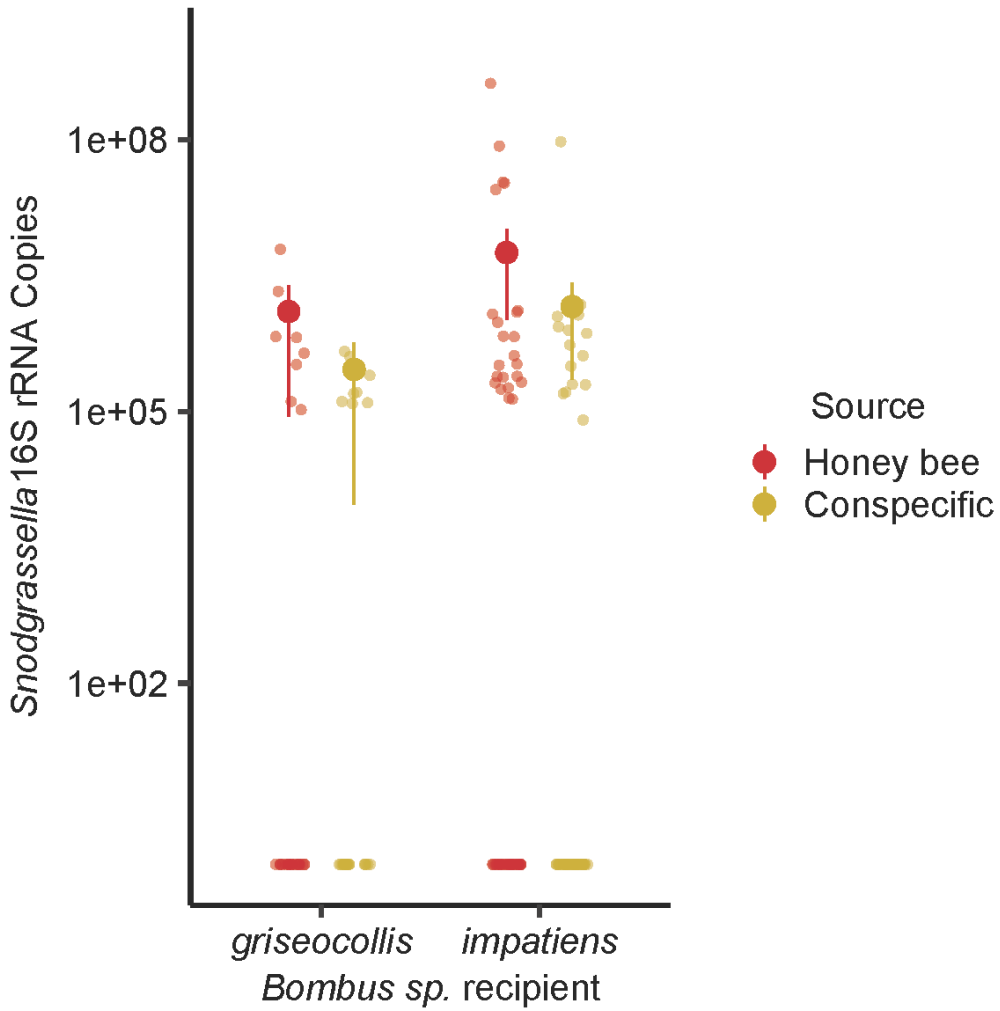


**Figure 6** *Snodgrassella* colonization abundance for *Bombus griseocollis* and *Bombus impatiens* recipients inoculated with honeybee isolated strains of *Snodgrassella* or strains isolated from the conspecific host of the respective recipient. Means and standard errors were estimated from truncated negative binomial models.

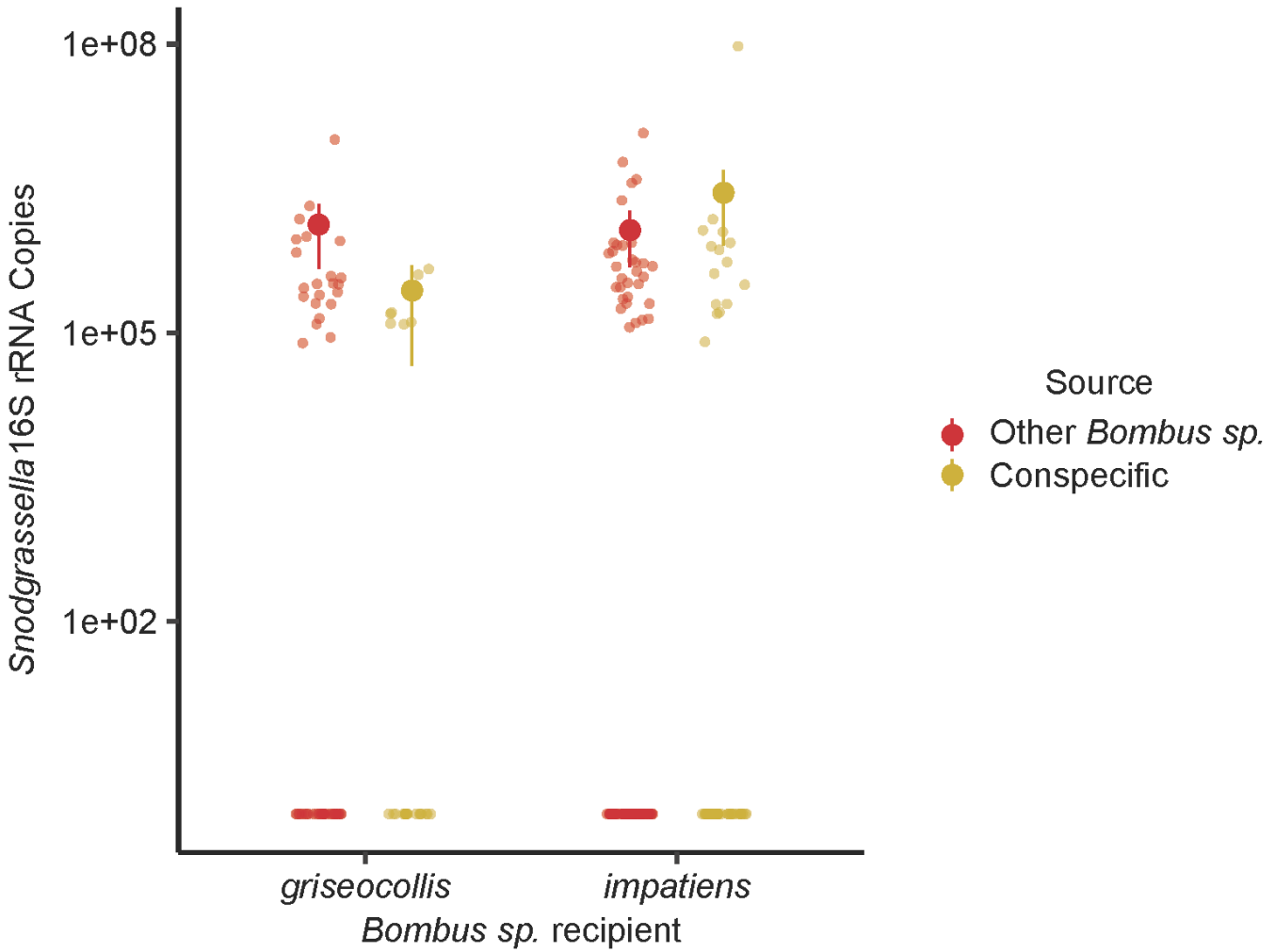


**Figure 7.** Colonization abundance of *Snodgrassella* in *Bombus griseocollis* and *Bombus impatiens* recipients of strains isolated from a conspecific source to the recipient or other bumble bee species. The means and standard errors were estimated from a truncated negative binomial model.

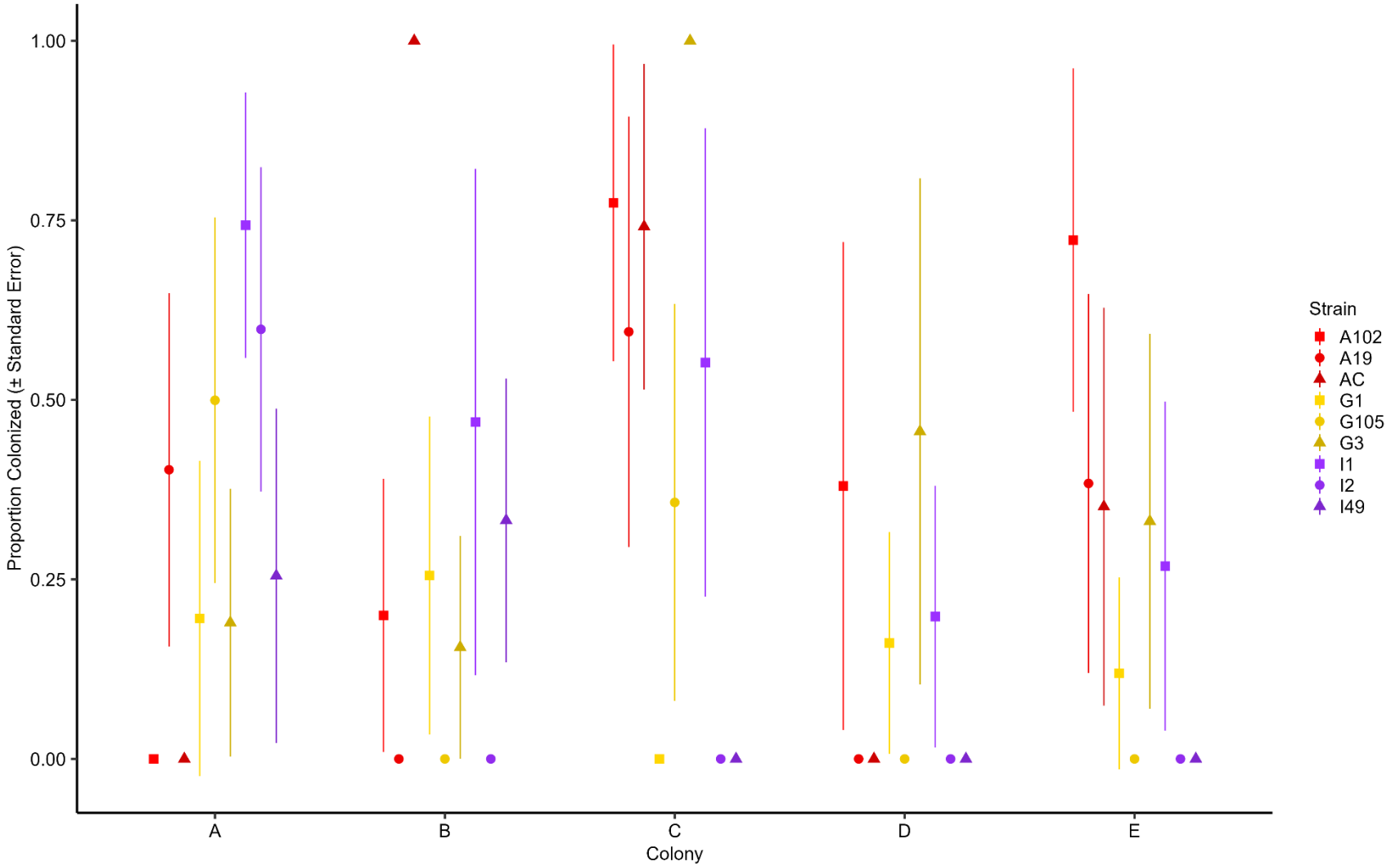


**Figure 8.** The presence or absence of colonization for *Bombus impatiens* samples across colonies inoculated with *Snodgrassella* strains isolated from bumble bee species. Estimated marginal means and standard errors are estimated from a binomial model. When standard errors appear to not be present it is because the errors are extremely small and represent cases where *Snodgrassella* either colonized every sample or colonized no sample

# CHAPTER IV: IMMUNE STIMULATION INDUCED CHANGES TO THE NATIVE GUT MICROBIOTA OF BUMBLE BEES

**This chapter is modified from the submitted article:** Sauers, L., Bassingwaite, T., Sierra-Rivera, B., Hampton, K., Duffield, K., Ramirez, J-L., Sadd, BM. Immune stimulation induced changes to the native gut microbiota of bumble bees. *Proc Royal Soc B*

**Abstract**

Understanding factors influencing the maintenance and membership of beneficial host-associated microbial communities is central to understanding ecological, evolutionary, and health consequences of these communities for their hosts. Host immunity is often implicated as a potential regulator of these microbiota. However, conversely, immunity may play a disruptive role, with immune responses to infection causing collateral damage. Such effects may be more prominent from innate immune responses, with more rapid acting and relatively non-specific components. We investigated how upregulation of antibacterial immunity in the bumble bee *Bombus impatiens* affects the core gut microbiota, testing the hypothesis of immunity induced perturbation of the beneficial microbiota structure. Freshly emerged adult bees received a native microbiota inoculation before being subjected to non-pathogenic immune stimulation treatments. We quantified the microbial community using 16S rRNA amplicon sequencing and targeted qPCR. We find colonization of the core member *Gilliamella* is altered by immune stimulation treatment. Additionally, a positive association in communities between *Gilliamella* and another core bacteria, *Snodgrassella alvi*, is perturbed. These changes are indicative of immune response induced dysbiosis. As such, the potential for collateral perturbation of beneficial microbial communities upon a host innate immune response may contribute to immune costs, shaping the evolutionary optimization of immune investment.

**Introduction**

There continues to be an increasing appreciation of the relevance of host-associated microbial communities, termed microbiota, for host health and fitness. Benefits hosts derive from their microbiota include assistance in digestion [1–3], defense against pathogens [4–6], and promotion of development

[7–10]. Furthermore, an imbalance in these microbial communities or dysbiosis can lead to detrimental effects for hosts [11–13]. As such, understanding factors that influence the maintenance and membership of beneficial microbial communities is central to understanding the ecological, evolutionary, and health consequences of these communities for their host organisms. Selection on hosts that is mediated through the phenotypes produced by the "extended genotype" of the microbiome have the potential to significantly influenced host evolution [14].

The gut microbiota of animal hosts is particularly well-studied and given the benefits that have been uncovered of a native microbiota the factors influencing this microbial community and its variation may be particularly important [15,16]. The gut microbiota structure and function can be highly variable among individuals and over time [17,18]. This variation is influenced by many factors that can be intrinsic or extrinsic, and include mode of birth [19], hygiene [20], exposure to antimicrobials [11,13,21], diet [22–25], co-colonization with enteric parasites [26–28] and host genetics [29–32]. The potential for such modulation of the gut microbiota by intrinsic and extrinsic factors is crucial to understand given the essential health and fitness related roles of the gut microbiota and that deviations from healthy microbial communities often result in detrimental phenotypes for hosts [11,13]. While an apparently healthy gut microbiota can vary, compositions linked to diseases states can also be mediated by the host and/or external environment. Thus, further studies focused on disentangling the membership, abundances, and interactions in these host gut microbial communities are needed.

One crucial intrinsic control of these communities is the host's immune system. Host immunity and the maintenance of microbial communities has been well studied in adaptive immunity [33,34]; however, it is now becoming better understood that innate immunity in vertebrates and invertebrates also represents an important factor [35–38]. There are an increasing number of examples of how host immune regulation maintains host-associated microbes, including in hydra [39,40] and *Drosophila* [41]. However, interplay between host immunity and these beneficial microbial communities is a delicate

balancing act. Hosts must coexist with their beneficial microbes, while being able to mount rapid, non-specific innate anti-pathogen responses [42]. This balance is critical to the mutualistic evolution of hosts and beneficial microbes, and the trajectory of host immune evolution may indeed be constrained due to beneficial symbionts [36]. Teasing apart how the innate immune system can both maintain gut homeostasis but also potentially induce microbiota dysbiosis is essential to advancing our knowledge of these host-microbe relationships and their evolution.

Host immunity, beneficial associated microbes, and parasite and pathogen infections have numerous linkages and potential ecological and evolutionary outcomes [43]. A variety of mechanisms exist for hosts to tolerate or regulate beneficial symbiotic microbes while responding to pathogens [36], such as compartmentalization [44]. However, aberrant immune responses have been implicated in gut microbiota dysbiosis, including in human diseases [45], and unless perfectly regulated and targeted, rapid, non-specific innate immune responses could plausibly disrupt a native, healthy gut microbiota structure. Moreover, it has been shown that a bacterial pathogen co-opts tick host immunity to disrupt the hosts microbiota and enhance infection [46]. However, that a host immune response upon infection may itself be costly due to a perturbation of the status quo of the host beneficial gut microbiota has not been directly investigated. Even if such disruption is short-lived, resulting reduced beneficial functions mediated by the microbiota could represent an under-appreciated cost of immunity in addition to traditionally identified usage costs and others that affect immune system evolution [47,48].

We utilize the gut microbiota of adult Common eastern bumble bees (*Bombus impatiens*) to further examine how host immune stimulation affects the overall gut microbial community structure, microbe associations, and the abundances of specialized core bacterial symbionts. The bumble bee native gut microbiota is relatively well-studied, possessing approximately 11-15 conserved, and often coevolved community members [49–52]. Furthermore, the native microbiota of adult bumble bees or related honey bees has been shown to provide the host various benefits, including digestion,

detoxification, development and pathogen defense [6–8,21,26,28,53–55]. We stimulate the bumble bee antibacterial immune response through heat-killed bacteria injections that represent benign immune elicitors and remove confounding effects of pathogenic infection that could otherwise influence the microbiota through multiple routes [43]. We hypothesize that induction of host immunity can lead to collateral damage through perturbation of the beneficial gut microbiota. In essence, an additional cost of immune system activation on infection is "friendly fire" of the immune system upon the "extended genotype" of the microbiome leading to potential knock-on effects for the host phenotype. Further, we induce the host immune response at two timepoints relative to colonization of the adult microbiota. We hypothesize that the timing of the immune response in relation to native gut microbiota establishment will affect the extent of perturbation, predicting that the microbiota will be more susceptible to infection induced changes early in its colonization.

**Materials and Methods**

***Overall experimental design***

Freshly emerged experimental adult workers were taken from five colonies of *B. impatiens* derived from wild caught queens. Immediately on isolation, bees received a standardized microbiota inoculum from their natal colony. After consuming this inoculum bees were assigned to one of two immune treatment timepoints, of 0 days post adult emergence or 4 days post adult emergence, representing early and late stages of microbiota establishment [49], respectively. On the assigned day, bees received one of the following immune stimulation treatments: i) left naïve, or injected with ii) ringer saline, as wounding only, iii) heat-killed *Escherichia coli* or iv) heat-killed *Staphylococcus epidermidis*. Two days after their respective immune treatments the hindgut microbiota of these bees were analyzed using amplicon sequencing to assess community composition, microbe associations, and abundances of specific core members. However, amplicon sequencing alone is not a good predictor of absolute abundances [56], as such we utilized qPCR targeting total 16S rRNA to determine the overall

abundance of bacteria, and adjusted the amplicon relative abundances by the qPCR determined total load, from here on our referred to as adjusted amplicon or adjusted values. Further, qPCR of the specific abundances of relevant bee associated bacteria, *Snodgrassella alvi*, *Gilliamella spp., Candidatus Schmidhempelia bombi*, and *Lactobacillus spp.* was used to validate the adjustment approach. A separate set of bees received the same immune stimulation treatments at the timepoints above, but were used to verify the antibacterial immune activation.

***Bumble bee maintenance and microbiota inoculation***

Queens were collected with the permission of local landowners and the ParkLands Foundation (http://www.parklandsfoundation.org/) from the Mackinaw River Study Area (Lexington, IL). Founded colonies were held under red light at 26ºC (± 1.5ºC) and provided with sugar water ad libitum and pollen three days per week following standard protocols [57]. Founding queens and a subset of workers were checked for pathogens by fecal screening and confirmed free of detectable gut pathogens. Experimental workers were isolated from the colony within 24 hours of adult emergence. They were starved for 30 minutes in sterile collection vials before being offered a standardized microbiota inoculum from their natal colony. The 15µL inoculum consisted of 1-part freshly collected and pooled feces from at least five random workers (at least 48 hours post adult emergence) from the source colony and 2 parts sugar water. Full consumption of the inoculum was confirmed visually. This procedure has been used in previous work to re-establish gut communities in axenic bees [8]. Early microbiota establishment immune stimulation treatments took place immediately or late establishment treatments 4 days later. When not receiving their immune stimulation treatments bees were held in individual holding boxes with pollen and sugar water *ad libitum* until further processing.

***Immune stimulation treatments of microbiota analysis bees***

Immune stimulation treatments took place either on day 0 post adult emergence, following the microbiota inoculation, or 4 days post adult eclosion and microbiota inoculation. Heat-killed bacteria

immune stimulation treatments were produced by suspending heat-killed cells of gram-negative *E. coli* (tlrl-hkeb2, InvivoGen, San Diego, CA) and *S. epidermidis* (tlrl-hkse, InvivoGen, San Diego, CA) in sterile insect ringer saline at 2 x $10^8$ cells per ml. Bees were placed on ice to immobilize them prior to treatments. Treatments consisted of a 2µL injection of the heat-killed bacteria solution using a sterile pulled glass microcapillary inserted between the first and second abdominal tergites of bees. Wounding only ringer saline injected bees were injected with 2µL of sterile insect saline. Naïve bees were handled in the same manner but not injected. Forty-eight hours following immune stimulation treatments, bees for microbiota analysis were frozen at -80°C.

***Confirmation of immune stimulation treatments***

To confirm the expected immune stimulation and identify any differences between treatments and timepoints a further 100 bees from three of the source colonies were treated as above, but their hemolymph extracted 24 hours post immune stimulation treatment and antibacterial activity measured in zone of inhibition assays against *Arthrobacter globiformis* [58]. Measured zone diameters were converted, based on a standard curve, to units (µg/mL) of the antibiotic tetracycline. Each bee sample was tested in duplicate, with the mean of the duplicates being used in subsequent analyses.

*Microbiota sample processing and amplicon sequencing*

Bees were briefly thawed, surface sterilized in sequentially 70% and 90% ethanol for 1 minute each, before rinsing with ultrapure water and being dried on a Kimwipe. Hindguts were sterilely dissected out and placed individually in sterile microcentrifuge tubes containing 2.4 and 3.2mm stainless steel beads (Scientific Industries, Bohemia, NY), 180µL of Buffer ATL (Qiagen, Hilden, Germany), and 10µL of Buffer Dx (Qiagen, Hilden, Germany). Tissue was disrupted using a TissueLyser II (Qiagen, Hilden, Germany) and total DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's adjusted protocol: Purification of total DNA from insects using the DNeasy Blood & Tissue Kit.

The microbiome was analyzed via 16S rRNA sequencing. Two replicate libraries were generated by following the 16S Metagenomic Sequencing Library Preparation guide from Illumina (Illumina, San Diego, CA), using PCR primers 341f and 806r [59]. The 16S V3-V4 region was amplified by PCR (parameters: 94°C, 3 min, 25 cycles of 95°C, 30 s; 55°C, 30 s; 72°C, 30 s) using 12.5 ng of sample DNA. Amplicons were purified with 0.8 volumes of AMPure beads, washed with ethanol, and eluted with 10 mM Tris. Using the same PCR parameters and clean-up method, unique dual indexes (Integrated DNA Technologies, Coralville, IA) were added to the overhang region of amplicons for library identification. Libraries were quantified with the Qubit 1x dsDNA HS Assay kit and a Qubit 4 Fluorometer (Thermo Fisher Scientific, Waltham, MA) and normalized to 4 nM. Before sequencing, the libraries were pooled, denatured, diluted to a 12 pM final concentration, and spiked with 10% PhiX control library (Illumina, San Diego, CA). Samples were sequenced using the Illumina MiSeq system with a MiSeq v3, 2 × 300 bp, sequencing kit (Illumina, San Diego, CA).

***Targeted qPCR***

qPCR was carried out to quantify total bacterial abundance and the abundance of specific gut bacteria. The targets of each primer were confirmed by running a 10µL PCR reactions (4.3 µL molecular grade water, 2 µL 5x GoTaq Flexi buffer, 0.5 µL of forward and reverse 10µM primer stocks, 1 µL of 2mM dNTP stock, 0.6 µL $MgCl^2$, 0.1 uL of GoTaq, and 1 µL template) ) with initial denaturing of 5 minutes at 95°C, 35 cycles of 95°C for 45 s, 54°C - 60°C (variable based on the primers) for 30 s, and 72°C for 60 s, followed by a final 5-minute extension at 72°C, followed by Sanger sequencing (Eurofins). qPCR was performed using SYBR Green PowerUp (Applied Biosystems, Waltham, MA) chemistry in 20µL reactions for total bacteria quantification and 10µl reactions for bacteria specific primers. Negative controls consisted of 2µL ddH2O and standard curves were made with dilutions of gBlock synthetic DNA (Integrative DNA Technologies, Coralville, IA.) created to match the target sequence. qPCR reactions were performed on the QuantStudio flex 7 platform with a 2-minute initial denaturing step at 50°C and

2 minutes at 95°C followed by 40 cycles of denaturing at 95°C for 15 s, annealing and elongation at 60°C for 1 minute. Although primer specificity has been verified by sequencing, melt curves generated after each run were checked to ensure specificity of amplification. Each DNA sample was run in duplicate, and any duplicates that had a calculated coefficient of variation above 0.20 were rerun.

***Bioinformatics analysis***

The sequences from both amplicon sequencing libraries were concatenated with a BASH script before being analyzed using QIIME2 version 2020.11 [60,61]. First barcode adapters and the primers were removed from sequences using the CutAdapt plugin. Subsequently, sequences were processed, analyzed, and filtered using DADA2 with the following parameters based off visualized quality score plots: forward truncation at 250bp, reverse truncation at 220bp. Taxonomy to amplicon sequencing variants (ASVs) was automatically assigned with a Naïve Bayes automatic classification using the SILVA SSU database 138 full sequence database [62] as a training dataset for the classifier. Downstream analysis was done in R [63].

***Statistical analysis***

All statistical analysis were done in R version 4.2.0. Microbiome data was analyzed using the packages phyloseq, vegan, and NetCoMi [63–66]. First, reads assigned to Eukaryotes, chloroplast, and mitochondria were filtered out of the dataset. Relative abundances generated from phyloseq were converted to absolute abundances based on the qPCR quantification of total 16S rRNA bacterial abundance, by multiplying the relative abundances for each ASV in the sample by the qPCR estimated 16S rRNA load for the sample. To check these absolute abundance estimates correlated with the actual abundances in the samples, we compared the converted values with specific targeted qPCR quantification. We find high agreement between our qPCR values and the adjusted absolute count values calculated from the amplicon sequencing for *Snodgrassella* ($R^2$ = 0.868), *Gilliamella* ($R^2$ = 0.87), *Lactobacillus* ($R^2$ = 0.95), and *Schmidhempelia* ($R^2$ = 0.73). Subsequently, community analysis metrics

were generated by calculating the Bray-Curtis distance between samples. PERMAnova was utilized to determine the effect of treatments on the distance between samples.

Network analysis was carried out with the package NetCoMi. Data was preprocessed by grouping the ASVs by their genus and grouping the individual samples by their assigned treatments. Networks were made by Spearman correlation with a center log-ratio transformation to normalize the data. Zeros were adjusted by adding 1 to the data as with other log-ratio transformations. Correlations were tested for statistical significance by bootstrapping and adjusted for multiple tests using the false discovery rate. Only statistically significant relationships are shown in figures. Network parameters were compared with Fisher's Z Test and permutation testing.

For analyzing the abundance of the total bacterial 16S rRNA and individual targeted bacteria 16S rRNA, generalized linear mixed models were used in glmTMB [67]. Model fixed effects included immune stimulation treatment, timing of treatment (early or late during microbiota establishment), and their interaction, with colony as a random effect. Negative binomial, negative binomial hurdle models, and negative binomial zero-inflated models were fit to the data. The best model was picked by a combination of visualizing the residual vs predicted plots and dispersion testing with the DHARMa package [68], AICc with the AICcmodavg package [69], and hypothesis testing. In all cases the negative binomial distribution was found to be superior.

Antibacterial activity measures to confirm immune activation and identify any differences between treatments was analyzed with a linear mixed effects model with the response variable square-root transformed to adhere to assumptions. Fixed effects included immune stimulation treatment, treatment timing, and their interaction, with colony as a random effect.

Graphs were created with the ggplot2 package [70].

## Results

### *Antibacterial activity following immune stimulation treatment*

There is a significant effect of immune stimulation treatment on antibacterial activity ($X^2$ = 41.635, d.f. = 3, p < 0.001; **Figure 9**), but no significant effect of treatment day ($X^2$ = 1.426, d.f. = 1, p = 0.236) or the interaction of immune stimulation treatment and treatment day ($X^2$ = 0.939, d.f. = 3, p = 0.425). Both *E. coli* and *S. epidermidis* heat-killed bacteria injection treatments significantly increase antibacterial activity above naïve and ringer saline injected treatments (Tukey HSD adjusted pairwise contrasts: p < 0.001), but the two bacteria treatments do not differ from one another (p = 0.112). Ringer saline injections have higher antibacterial activity than naïve bees (p = 0.004), likely a result of the wounding from the injection as expected.

### *Microbiome analysis and diversity indices*

Gut microbiomes from 158 individuals were analyzed using the hypervariable V3-V4 region of the bacterial 16S rRNA gene with Illumina MiSeq amplicon sequencing. In total there were 19,247,933 reads with an average of 101,359.15 (± 46,345.67) reads per sample and 512 identified ASVs. High relative abundance bacterial genera include *Schmidehempelia*, *Gilliamella*, *Lactobacillus*, and *Snodgrassella*, as expected (**Figure 10**). Based on absolute abundance, beta diversity distance is not significantly affected by immune stimulation treatment (p = 0.7472) nor its interaction with treatment day (p = 0.2151). However, there is a significant effect of treatment day (p < 0.001) and colony (p <0.001) on beta diversity distance and the effect of immune stimulation was marginally non-significant when looking at those individuals treated 4 days post adult eclosion (p = 0.054) (**Figure 11**). The total bacterial abundance quantified by qPCR does not differ by immune stimulation treatment ($X^2$ = 1.76, d.f. = 3, p = 0.624), or the interaction between them ($X^2$ = 0.820, d.f. = 3, p = 0.845). However, there is a significant effect of treatment day ($X^2$ = 13.571, d.f. = 1, p = 0.0002) with total bacterial load decreasing from day 0 to day 4 treated bees.

***Quantification of specific microbiota members***

There is a significant effect of immune stimulation treatment on the absolute abundance of *Gilliamella* calculated from the adjusted amplicon data (p=0.0011), with treatments of heat-killed *S. epidermidis* and wounding only, ringer injected having higher abundances relative to *E.coli* injected and naïve bees (**Figure 12** and **Table 4**). In addition, the abundance of *Schmidhempelia* is influenced by treatment day (p=0.007). There are no further effects of immune stimulation treatment, treatment day or their interaction on the other core gut microbes *S. alvi* and *Lactobacillus sp.*, or the non-core, environmental microbe *Pseudomonas* (**Table 4**). The results from species-specific qPCR determined abundances qualitatively agree with the results from these adjusted amplicon values (results not shown).

***Network analysis-based community associations***

Network analysis was performed on communities subset by the immune stimulation treatment (**Figure 13**). There are two main microbe community association clusters in bees in the naïve treatment. The first cluster consists of the core gut symbionts *Gilliamella* and *Snodgrassella*. The second cluster consists of *Schmidhempelia* and *Lactobacillus*, along with many non-core, environmental and potentially pathogenic, bacteria, such as *Enterobacter, Serriatia,* and *Pseudomonas*. Associations within each cluster are predominantly positive, but with negative between the cluster associations. However, immune stimulation of any kind appears to precipitate a large shift in the network of microbe community associations within the bumble bee gut (**Figure 13**). Under the immune stimulation treatments of injection only wounding and the heat-killed bacteria injection, the second cluster of *Schmidhempelia-Lactobacillus*-environmental microbes is retained or even strengthened, but the key positive association between *Gilliamella* and *Snodgrassella* is disrupted.

**Discussion**

Host physiological changes have the potential to select for changes in or disrupt their associated microbiota, either changing membership in the microbial community, overall abundances of microbes or microbe types, or associations between microbes within the communities. Here we examine how stimulation of insect innate immunity against a simulated wounding or pathogenic challenges may alter the bumble bee gut microbiota, which has been implicated as important for host health and fitness. We take an integrated approach, combining insights based on general community structure, abundances of select microbes of potential importance, and specific network associations within the communities. Despite no evidence for significant coarse-level changes in microbiota membership across immune stimulation treatments, we see that the abundance of the core bumble bee gut bacterium *Gilliamella* is significantly affected by the immune stimulation treatment. Moreover, the potential disruptive effect that an induced immune response can have is demonstrated by a perturbation of the network of microbe associations within the bumble bee gut microbiota.

We reveal that stimulation of host immunity, independent of actual infection, differentially affects microbiota members. Our amplicon sequencing results reveal that, in general, membership in the microbiota based on presence or absence is relatively invariant, which may indicate the potential for the microbiota to rebound following acute perturbations to abundances of particular members. We find that our native microbiota inoculums lead to colonization of expected members [49,52], and document temporal changes in *Schmidhempelia* in agreement with earlier work [17]. Importantly, we show that immune stimulation through wounding and *S. epidermidis* gram positive bacteria elicitation leads to an increase in the abundance of *Gilliamella.* This indicates that the microbiota does not respond as a whole to host induced changes, but rather particular members may be affected disturbing the normal balance of the community. Furthermore, these results indicate that the source of immune stimulation matters. Our confirmation of immune activation used a relatively broad assessment of antibacterial activity, but it

is known that different immune challenges can lead to different host gene expression profiles [71]. Such differences in host response could dictate different outcomes for the microbiota. Intriguingly, it has been shown that increased abundances of *Gilliamella* within the host are positively associated with prevalence of infection of the trypanosome *Crithidia bombi* [4]. The increase in *Gilliamella* abundance during *Crithidia* infection could be in part mediated by the host immune response. Although intuitively it may seem beneficial for the host to have increased numbers of *Gilliamella*, which is thought to aid bee hosts in digestion and nutrition [3,51,53,55,72], ultimately it could be the result of disrupted associations with other microbial species and signal a state of microbiota dysbiosis within the host.

The network analysis of community associations provides further support for disrupted associations between microbiota members. In naïve individuals, the two core bacteria *Snodgrassella* and *Gilliamella* show a positive association. These two bacteria have been shown to cross-feed with each other and are proposed to colonize together in the hindgut creating a biofilm on the host epithelium [73]. However, under all other immune stimulation treatments this positive association in the communities is broken up, indicating a potential shift in their interaction. This change is also seen in the wounding treatment, which we see elicits an immune response, albeit at a lower level than the heat-killed bacteria injections. A second, largely positively associating cluster in the naïve group includes the core microbe *Schmidhempelia* and environmentally acquired, potentially pathogenic microbes including *Pseudomonas* and *Enterobacter*. The associations in this cluster remains largely consistent, perhaps even strengthening, following immune stimulation. This cluster has negative associations with the core beneficial bacteria *Snodgrassella* and *Gilliamella*, and these results further indicate the potential for host immune induced disruption of key beneficial associations within the microbiota.

Further studies are needed to elucidate the ultimate effects of immune stimulation induced disturbance of the microbiota of bumble bee hosts. It is currently unclear if the effects seen here are the result of direct effects of immune products acting upon microbiota members or rather the indirect

effect of altered immune physiology that shifts it from a state of regulation of the microbiota that has been documented in other systems [36]. In addition, it is important to relate the changes that are seen in the microbiota to functional consequences for the host. A native microbiota is proposed to provide various benefits to insect hosts [74], including in the gut of bumble bees [75]. The extent to which these benefits are perturbed and if and when a healthy community and functioning is restored will dictate the consequences for host health and the ecological and evolutionary dynamics between hosts, their immunity, and associated microbes [36].

Overall, this work shows that the insect innate immune response can play a role in determining the structure of the host-associated microbial community and may disrupt the natural community causing a state of dysbiosis. This provides evidence for an additional link between the immunity of hosts and their microbiota, in addition the regulatory and other bi-directional effects [36,43,76]. Disruption of the host-associated microbiota upon the immune response to a wounding or pathogenic threat represents yet another potential usage cost of insect immunity in addition to previously documented energetic costs, physiological, and autoreactive costs [47,48,77], especially if effects have lasting functional consequences. As such, this could explain natural variation in immune responses, features of the microbiota such as redundancy [78], and the evolutionary optimization of host immunity [36].

**References**


1. Brune A. 2014 Symbiotic digestion of lignocellulose in termite guts. *Nat Rev Microbiol* **12**, 168–180. (doi:10.1038/nrmicro3182)
2. Cerqueira FM, Photenhauer AL, Pollet RM, Brown HA, Koropatkin NM. 2020 Starch digestion by gut bacteria: crowdsourcing for carbs. *Trends Microbiol* **28**, 95–108. (doi:10.1016/j.tim.2019.09.004)
3. Zheng H, Perreau J, Powell JE, Han B, Zhang Z, Kwong WK, Tringe SG, Moran NA. 2019 Division of labor in honey bee gut microbiota for plant polysaccharide digestion. *Proc Natl Acad Sci USA* **116**, 25909–25916. (doi:10.1073/pnas.1916224116)
4. Cariveau DP, Elijah Powell J, Koch H, Winfree R, Moran NA. 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)
5. Heine D *et al.* 2018 Chemical warfare between leafcutter ant symbionts and a co-evolved pathogen. *Nat Commun* **9**, 2208. (doi:10.1038/s41467-018-04520-1)
6. Steele MI, Motta EVS, Gattu T, Martinez D, Moran NA. 2021 The gut microbiota protects bees from invasion by a bacterial pathogen. *Microbiol Spectr* **9**, e0039421. (doi:10.1128/Spectrum.00394-21)
7. Kwong WK, Mancenido AL, Moran NA. 2017 Immune system stimulation by the native gut microbiota of honey bees. *R Soc Open Sci* **4**, 170003. (doi:10.1098/rsos.170003)
8. Näpflin K, Schmid-Hempel P. 2016 Immune response and gut microbial community structure in bumblebees after microbiota transplants. *Proc Biol Sci* **283**. (doi:10.1098/rspb.2016.0312)
9. Maynard CL, Elson CO, Hatton RD, Weaver CT. 2012 Reciprocal interactions of the intestinal microbiota and immune system. *Nature* **489**, 231–241. (doi:10.1038/nature11551)

10. Palm NW, de Zoete MR, Flavell RA. 2015 Immune-microbiota interactions in health and disease. *Clin Immunol* **159**, 122–127. (doi:10.1016/j.clim.2015.05.014)

11. Gasparrini AJ, Crofts TS, Gibson MK, Tarr PI, Warner BB, Dantas G. 2016 Antibiotic perturbation of the preterm infant gut microbiome and resistome. *Gut Microbes* **7**, 443–449. (doi:10.1080/19490976.2016.1218584)

12. Li RW, Wu S, Baldwin RL VI, Li W, Li C. 2012 Perturbation dynamics of the rumen microbiota in response to exogenous butyrate. *PLoS One* **7**, e29392. (doi:10.1371/journal.pone.0029392)

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

14. Henry LP, Bruijning M, Forsberg SKG, Ayroles JF. 2021 The microbiome extends host evolutionary potential. *Nat Commun* **12**, 5141. (doi:10.1038/s41467-021-25315-x)

15. Moeller AH, Sanders JG. 2020 Roles of the gut microbiota in the adaptive evolution of mammalian species. *Phil Trans Roy Soc Bio Sci* **375**, 20190597. (doi:10.1098/rstb.2019.0597)

16. Macke E, Tasiemski A, Massol F, Callens M, Decaestecker E. 2017 Life history and eco-evolutionary dynamics in light of the gut microbiota. *Oikos* **126**, 508–531. (doi:10.1111/oik.03900)

17. Hammer TJ, Easton-Calabria A, Moran NA. 2023 Microbiome assembly and maintenance across the lifespan of bumble bee workers. *Molec Ecol* **32**, 724–740. (doi:10.1111/mec.16769)

18. Vandeputte D, De Commer L, Tito RY, Kathagen G, Sabino J, Vermeire S, Faust K, Raes J. 2021 Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. *Nat Comm* **12**, 6740. (doi:10.1038/s41467-021-27098-7)

19. Selma-Royo M, Calatayud Arroyo M, García-Mantrana I, Parra-Llorca A, Escuriet R, Martínez-Costa C, Collado MC. 2020 Perinatal environment shapes microbiota colonization and infant growth: impact on host response and intestinal function. *Microbiome* **8**, 167. (doi:10.1186/s40168-020-00940-8)

20. Azad MB *et al.* 2013 Infant gut microbiota and the hygiene hypothesis of allergic disease: impact of household pets and siblings on microbiota composition and diversity. *Allergy, Asthma & Clinical Immunology* **9**, 15. (doi:10.1186/1710-1492-9-15)

21. Li JH, Evans JD, Li WF, Zhao YZ, DeGrandi-Hoffman G, Huang SK, Li ZG, Hamilton M, Chen YP. 2017 New evidence showing that the destruction of gut bacteria by antibiotic treatment could increase the honey bee's vulnerability to *Nosema* infection. *PLoS One* **12**, e0187505. (doi:10.1371/journal.pone.0187505)

22. Tamura K, Brumer H. 2021 Glycan utilization systems in the human gut microbiota: a gold mine for structural discoveries. *Curr Opin Struct Biol* **68**, 26–40. (doi:10.1016/j.sbi.2020.11.001)

23. Tuncil YE, Xiao Y, Porter NT, Reuhs BL, Martens EC, Hamaker BR. 2017 Reciprocal prioritization to dietary glycans by gut bacteria in a competitive environment promotes stable coexistence. *mBio* **8**. (doi:10.1128/mBio.01068-17)

24. Marcobal A, Sonnenburg JL. 2012 Human milk oligosaccharide consumption by intestinal microbiota. *Clin Microbiol Infect* **18**, 12–15. (doi:10.1111/j.1469-0691.2012.03863.x)

25. Lawson MAE, O'Neill IJ, Kujawska M, Gowrinadh Javvadi S, Wijeyesekera A, Flegg Z, Chalklen L, Hall LJ. 2020 Breast milk-derived human milk oligosaccharides promote Bifidobacterium interactions within a single ecosystem. *ISME J* **14**, 635–648. (doi:10.1038/s41396-019-0553-2)

26. Mockler BK, Kwong WK, Moran NA, Koch H. 2018 Microbiome structure influences infection by the parasite *Crithidia bombi* in bumble bees. *Appl Environ Microbiol* **84**. (doi:10.1128/AEM.02335-17)

27. Palmer-Young EC, Raffel TR, McFrederick QS. 2019 pH-mediated inhibition of a bumble bee parasite by an intestinal symbiont. *Parasitology* **146**, 380–388. (doi:10.1017/S0031182018001555)

28. Koch H, Schmid-Hempel P. 2011 Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc Natl Acad Sci USA* **108**, 19288–92. (doi:10.1073/pnas.1110474108)

29. Cornet L, Cleenwerck I, Praet J, Leonard RR, Vereecken NJ, Michez D, Smagghe G, Baurain D, Vandamme P. 2022 Phylogenomic analyses of *Snodgrassella* isolates from honeybees and bumblebees reveal taxonomic and functional diversity. *mSystems* **7**, e0150021. (doi:10.1128/msystems.01500-21)

30. Hammer TJ, Dickerson JC, McMillan WO, Fierer N. 2020 Butterflies host characteristic and phylogenetically structured adult-stage microbiomes. *Appl Environ Microbiol* **86**. (doi:10.1128/AEM.02007-20)

31. Pollock FJ, McMinds R, Smith S, Bourne DG, Willis BL, Medina M, Thurber RV, Zaneveld JR. 2018 Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat Commun* **9**, 4921. (doi:10.1038/s41467-018-07275-x)

32. Sauers LA, Sadd BM. 2019 An interaction between host and microbe genotypes determines colonization success of a key bumble bee gut microbiota member. *Evolution* **73**, 2333–2342. (doi:10.1111/evo.13853)

33. Barroso-Batista J, Demengeot J, Gordo I. 2015 Adaptive immunity increases the pace and predictability of evolutionary change in commensal gut bacteria. *Nat Commun* **6**, 8945. (doi:10.1038/ncomms9945)

34. Rooks MG, Garrett WS. 2016 Gut microbiota, metabolites and host immunity. *Nat Rev Immunol* **16**, 341–352. (doi:10.1038/nri.2016.42)

35. Günther C, Josenhans C, Wehkamp J. 2016 Crosstalk between microbiota, pathogens and the innate immune responses. *Int J Med Microbiol* **306**, 257–265. (doi:10.1016/j.ijmm.2016.03.003)

36. Gerardo NM, Hoang KL, Stoy KS. 2020 Evolution of animal immunity in the light of beneficial symbioses. *Phil Trans Roy Soc Bio Sci* **375**, 20190601. (doi:10.1098/rstb.2019.0601)

37. Thaiss CA, Zmora N, Levy M, Elinav E. 2016 The microbiome and innate immunity. *Nature* **535**, 65–74. (doi:10.1038/nature18847)

38. Tafesh-Edwards G, Eleftherianos I. 2023 The role of *Drosophila* microbiota in gut homeostasis and immunity. *Gut Microbes* **15**, 2208503. (doi:10.1080/19490976.2023.2208503)

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

40. Fraune S, Anton-Erxleben F, Augustin R, Franzenburg S, Knop M, Schröder K, Willoweit-Ohl D, Bosch TC. 2015 Bacteria–bacteria interactions within the microbiota of the ancestral metazoan Hydra contribute to fungal resistance. *ISME J* **9**, 1543–1556. (doi:10.1038/ismej.2014.239)

41. Bosco-Drayon V, Poidevin M, Boneca IG, Narbonne-Reveau K, Royet J, Charroux B. 2012 Peptidoglycan sensing by the receptor PGRP-LE in the *Drosophila* Gut induces immune responses to infectious bacteria and tolerance to microbiota. *Cell Host & Microbe* **12**, 153–165. (doi:10.1016/j.chom.2012.06.002)

42. Milligan-Myhre K, Small CM, Mittge EK, Agarwal M, Currey M, Cresko WA, Guillemin K. 2016 Innate immune responses to gut microbiota differ between oceanic and freshwater threespine stickleback populations. *Dis Model Mech* **9**, 187–198. (doi:10.1242/dmm.021881)

43. Armitage SA, Genersch E, McMahon DP, Rafaluk-Mohr C, Rolff J. 2022 Tripartite interactions: how immunity, microbiota and pathogens interact and affect pathogen virulence evolution. *Curr Opin Insect Sci* **50**, 100871. (doi:10.1016/j.cois.2021.12.011)

44. Ferrarini MG *et al.* 2022 Efficient compartmentalization in insect bacteriomes protects symbiotic bacteria from host immune system. *Microbiome* **10**, 156. (doi:10.1186/s40168-022-01334-8)

45. Zheng D, Liwinski T, Elinav E. 2020 Interaction between microbiota and immunity in health and disease. *Cell Res* **30**, 492–506. (doi:10.1038/s41422-020-0332-7)

46. Abraham NM *et al.* 2017 Pathogen-mediated manipulation of arthropod microbiota to promote infection. *Proc Natl Acad Sci USA* **114**, E781–E790. (doi:10.1073/pnas.1613422114)

47. Sadd BM, Schmid-Hempel P. 2009 Principles of ecological immunology. *Evol Appl* **2**, 113–121. (doi:10.1111/j.1752-4571.2008.00057.x)

48. Lazzaro BP, Tate AT. 2022 Balancing sensitivity, risk, and immunopathology in immune regulation. *Curr Opin Insect Sci* **50**, 100874. (doi:10.1016/j.cois.2022.100874)

49. Kwong WK, Moran NA. 2016 Gut microbial communities of social bees. *Nat Rev Microbiol* **14**, 374–84. (doi:10.1038/nrmicro.2016.43)

50. Kwong WK, Medina LA, Koch H, Sing KW, Soh EJY, Ascher JS, Jaffé R, Moran NA. 2017 Dynamic microbiome evolution in social bees. *Sci Adv* **3**, e1600513. (doi:10.1126/sciadv.1600513)

51. Kwong WK, Moran NA. 2013 Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family *Neisseriaceae* of the Betaproteobacteria, and *Gilliamella apicola* gen. nov., sp. nov., a member of *Orbaceae* fam. nov., Orbales ord. nov., a sister taxon to the order 'Enterobacteriales' of the Gammaproteobacteria. *Int J Syst Evol Microbiol* **63**, 2008–2018. (doi:10.1099/ijs.0.044875-0)

52. Engel P *et al.* 2016 The bee microbiome: impact on bee health and model for evolution and ecology of host-microbe interactions. *mBio* **7**, e02164-15. (doi:10.1128/mBio.02164-15)

53. Kešnerová L, Mars RAT, Ellegaard KM, Troilo M, Sauer U, Engel P. 2017 Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol* **15**, e2003467. (doi:10.1371/journal.pbio.2003467)

54. Engel P, Moran NA. 2013 Functional and evolutionary insights into the simple yet specific gut microbiota of the honey bee from metagenomic analysis. *Gut Microbes* **4**, 60–5. (doi:10.4161/gmic.22517)

55. Zheng H, Nishida A, Kwong WK, Koch H, Engel P, Steele MI, Moran NA. 2016 Metabolism of toxic sugars by strains of the bee gut symbiont *Gilliamella apicola*. *mBio* **7**. (doi:10.1128/mBio.01326-16)

56. Rao C, Coyte KZ, Bainter W, Geha RS, Martin CR, Rakoff-Nahoum S. 2021 Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature* **591**, 633–638. (doi:10.1038/s41586-021-03241-8)

57. Calhoun AC, Harrod AE, Bassingthwaite TA, Sadd BM. 2021 Testing the multiple stressor hypothesis: chlorothalonil exposure alters transmission potential of a bumblebee pathogen but not individual host health. *Proc Biol Sci* **288**, 20202922. (doi:10.1098/rspb.2020.2922)

58. Czerwinski MA, Sadd BM. 2017 Detrimental interactions of neonicotinoid pesticide exposure and bumblebee immunity. *J Exp Zool A Ecol Integr Physiol* **327**, 273–283. (doi:10.1002/jez.2087)

59. Kinoshita Y, Niwa H, Uchida-Fujii E, Nukada T. 2021 Establishment and assessment of an amplicon sequencing method targeting the 16S-ITS-23S rRNA operon for analysis of the equine gut microbiome. *Sci Rep* **11**, 11884. (doi:10.1038/s41598-021-91425-7)

60. Caporaso JG *et al.* 2010 QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335–6. (doi:10.1038/nmeth.f.303)

61. McMurdie PJ, Holmes S. 2014 Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* **10**, e1003531. (doi:10.1371/journal.pcbi.1003531)

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

63. Team Rs. 2019 RStudio: integrated development environment for R.

64. McMurdie PJ, Holmes S. 2013 phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217. (doi:10.1371/journal.pone.0061217)

65. Peschel S, Müller CL, von Mutius E, Boulesteix A-L, Depner M. 2021 NetCoMi: network construction and comparison for microbiome data in R. *Brief Bioinform* **22**. (doi:10.1093/bib/bbaa290)

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

67. Brooks ME, Kristensen K, Benthem KJ van, Magnusson A, Berg CW, Nielsen A, Skaug HJ, Mächler M, Bolker BM. 2017 The R Journal: glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal* **9**, 378–400. (doi:10.32614/RJ-2017-066)

68. Hartig F. 2016 *DHARMa: residual diagnostics for hierarchical (multi-level / mixed) regression models*. See https://github.com/florianhartig/DHARMa.

69. Mazerolle MJ. 2023 *AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c)*. See https://cran.r-project.org/package=AICcmodavg.

70. Wickham H. 2016 *Ggplot2: Elegant graphics for data analysis*. 2nd edn. Cham, Switzerland: Springer International Publishing.

71. Barribeau SM *et al.* 2015 A depauperate immune repertoire precedes evolution of sociality in bees. *Gen Biol* **16**, 83. (doi:10.1186/s13059-015-0628-y)

72. Engel P, Martinson VG, Moran NA. 2012 Functional diversity within the simple gut microbiota of the honey bee. *Proc Natl Acad Sci USA* **109**, 11002–7. (doi:10.1073/pnas.1202970109)

73. Kwong WK, Engel P, Koch H, Moran NA. 2014 Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc Natl Acad Sci USA* **111**, 11509–14. (doi:10.1073/pnas.1405838111)

74. Douglas AE. 2015 Multiorganismal insects: diversity and function of resident microorganisms. *Annu Rev Entomol* **60**, 17–34. (doi:10.1146/annurev-ento-010814-020822)

75. Hammer TJ, Le E, Martin AN, Moran NA. 2021 The gut microbiota of bumblebees. *Insectes Soc* **68**, 287–301. (doi:10.1007/s00040-021-00837-1)

76. Marra A, Hanson MA, Kondo S, Erkosar B, Lemaitre B. 2021 *Drosophila* antimicrobial peptides and lysozymes regulate gut microbiota composition and abundance. *mBio* **12**, e0082421. (doi:10.1128/mBio.00824-21)

77. Sadd BM, Siva-Jothy MT. 2006 Self-harm caused by an insect's innate immunity. *Proc Roy Soc Biol Sci* **273**, 2571–2574. (doi:10.1098/rspb.2006.3574)

78. Moya A, Ferrer M. 2016 Functional redundancy-induced stability of gut microbiota subjected to disturbance. *Trends Microbiol* **24**, 402–413. (doi:10.1016/j.tim.2016.02.002)

**Table 4.** Model results from fitted negative binomial models on adjusted absolute amplicon values. Significant effects are bolded.

| ***Schmidhempelia*** | | | |
|---|---|---|---|
| *Parameter* | $X^2$ | *Df* | *p-value* |
| **Treatment Day** | **30.977** | **1** | **<0.001** |
| Immune Treatment | 0.238 | 3 | 0.971 |
| Treatment Day * Immune Treatment | 2.199 | 3 | 0.532 |
| ***Snodgrassella*** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Treatment Day | 0.151 | 1 | 0.697 |
| Immune Treatment | 4.096 | 3 | 0.251 |
| Treatment Day * Immune Treatment | 1.256 | 3 | 0.740 |
| ***Lactobacillus*** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Treatment Day | 2.218 | 1 | 0.136 |
| Immune Treatment | 5.294 | 3 | 0.152 |
| Treatment Day * Immune Treatment | 6.050 | 3 | 0.109 |
| ***Gilliamella*** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Treatment Day | 1.146 | 1 | 0.284 |
| **Immune Treatment** | **16.062** | **3** | **0.001** |
| Treatment Day * Immune Treatment | 2.364 | 3 | 0.500 |
| ***Pseudomonas*** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Treatment Day | 1.348 | 1 | 0.120 |
| Immune Treatment | 5.833 | 3 | 0.246 |
| Treatment Day * Immune Treatment | 5.681 | 3 | 0.128 |

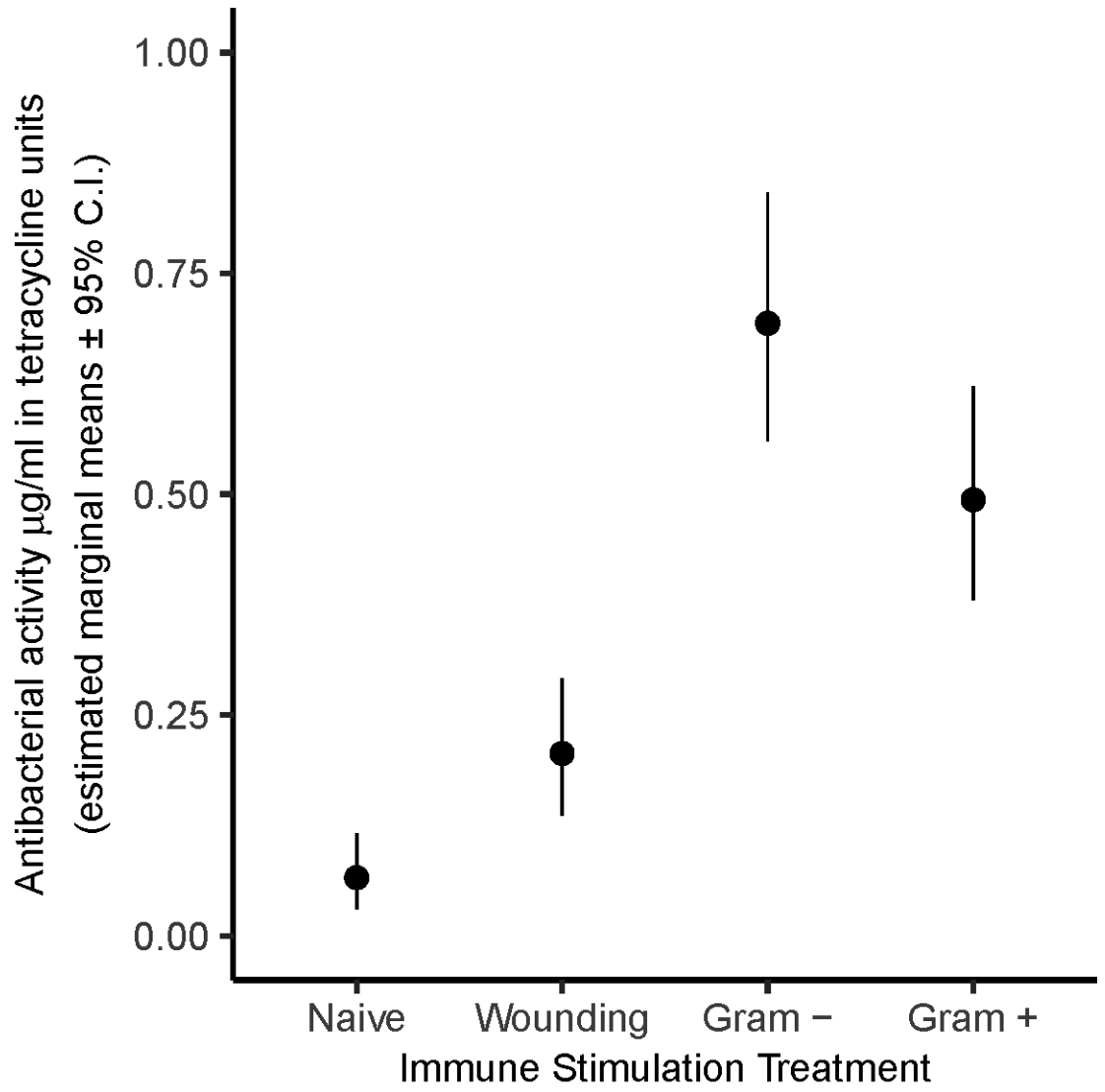


**Figure 9.** The effect of immune stimulation treatment on hemolymph antibacterial activity (standardized to units of the antibiotic tetracycline).

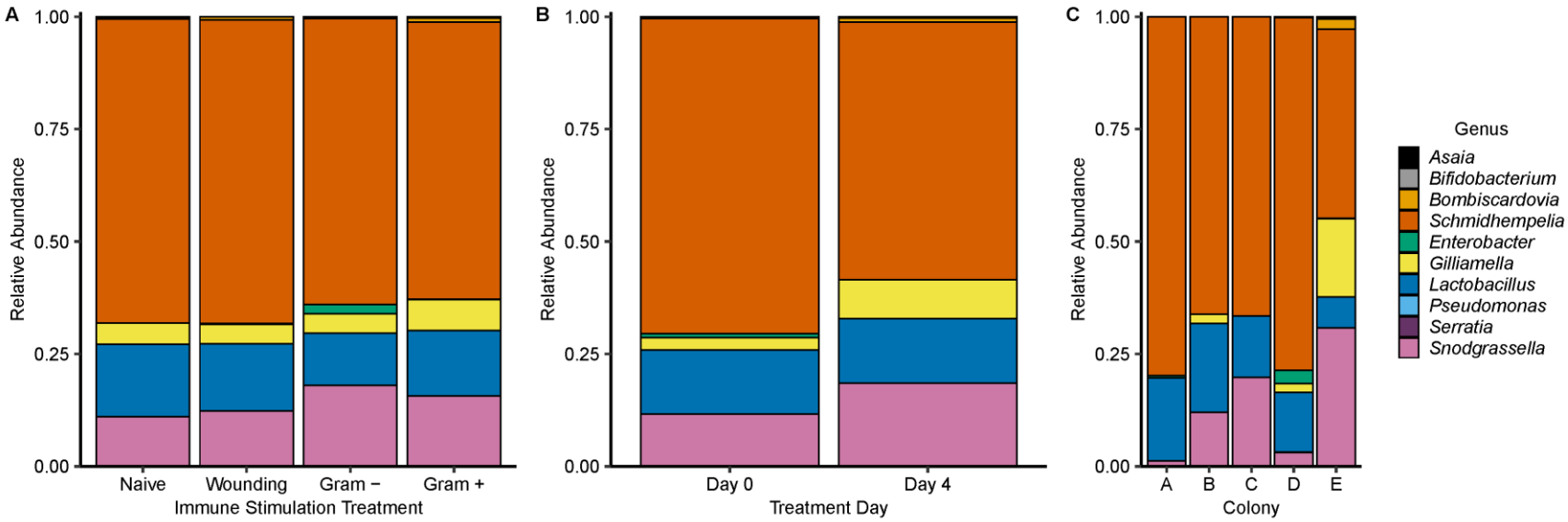


**Figure 10.** The relative abundance of bacterial genera making up at least 1% of the total reads (>161,161 reads) in the amplicon sequencing data set across **A)** immune stimulation treatment, **B)** treatment day, and **C)** bumble bee host colonies.

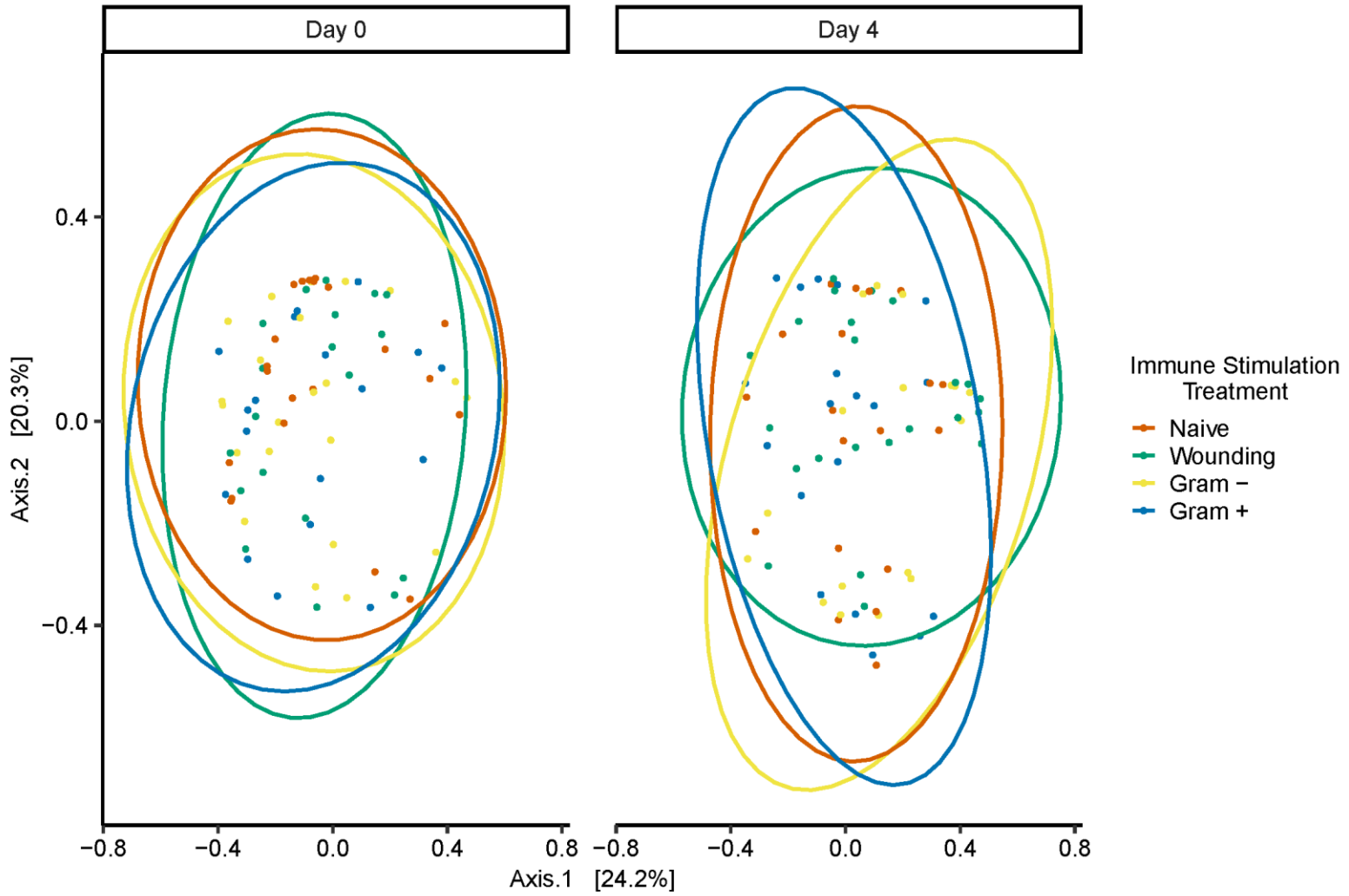


**Figure 11.** Bray-Curtis ordination plots of the composition of bumble bee microbiomes across immune stimulation treatments and treatment days.

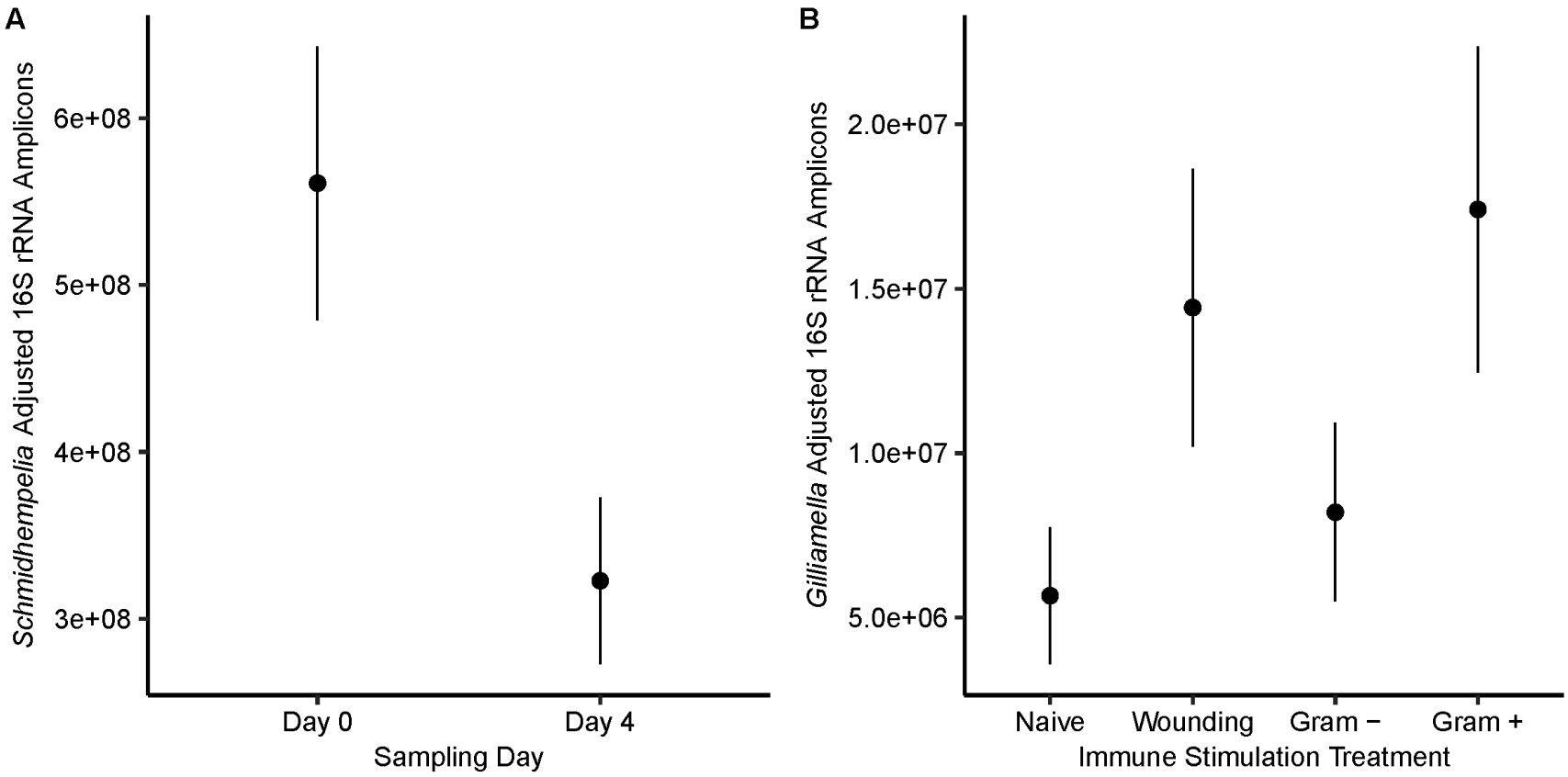


**Figure 12.** The abundances of *Schmidhempelia* **(A)** and *Gilliamella* **(B)** across treatment initiation day and immune stimulation treatments, respectively. Error bars represent the standard error

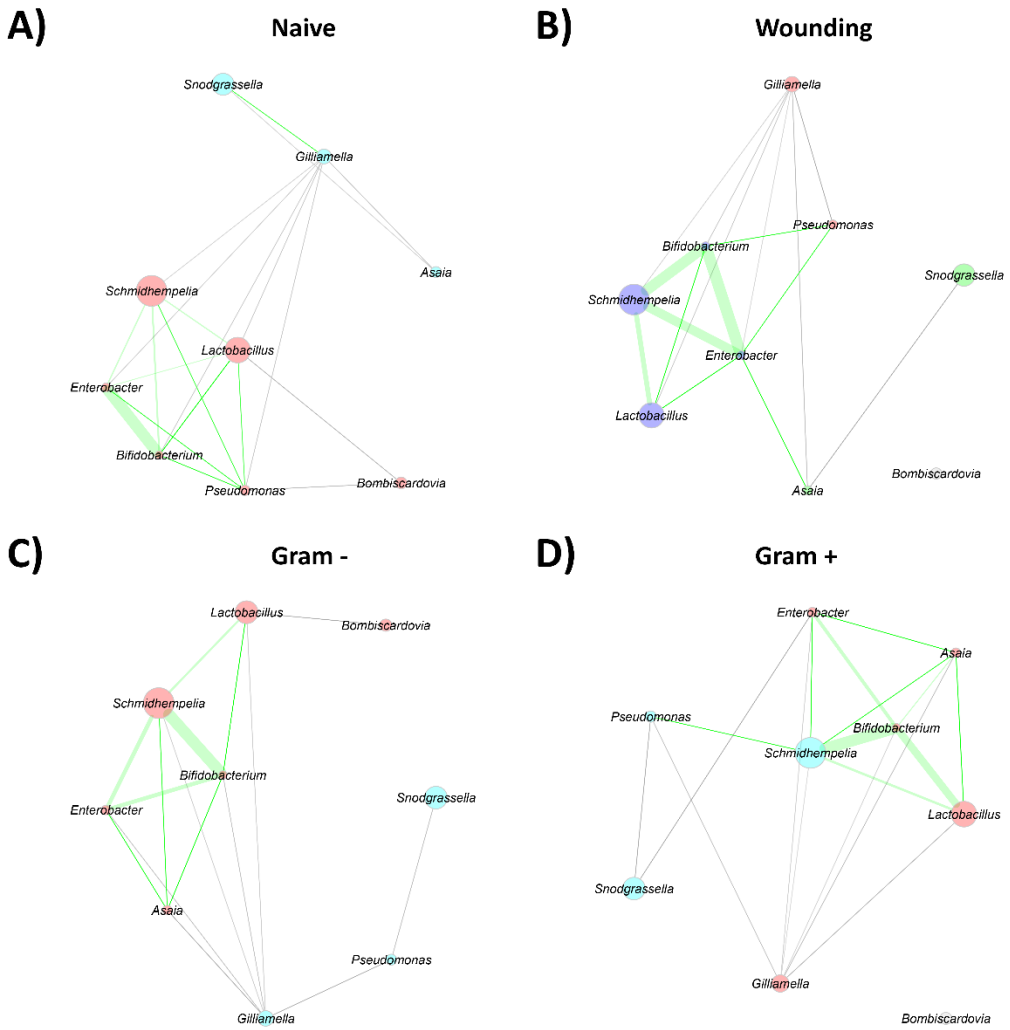


**Figure 13.** Spearman correlation-based networks for immune stimulation treatments of naïve (A), wounding (B), Gram –'ve injected (C), and Gram +'ve injected (D). Node color is related to the cluster the bacterial genus belongs to in the network, black lines indicate negative correlations, green lines indicate positive correlations. Node size indicates normalized abundance and line thickness the strength of the interaction.

# CHAPTER V: GUT MICROBIOME AND HOST METABOLOME CHANGES ASSOCIATED WITH DIFFERENT BUMBLE BEE QUEEN (*BOMBUS LANTSCHOUENSIS*) LIFE STAGES

**Abstract**

Honey bees and bumble bees possess relatively simple but conserved gut bacterial communities, which are proposed to play important roles in nutrition, metabolism, behavior, and physiology. However, while the membership of these gut communities is well studied in honey bee systems, there remain critical gaps in our understanding for bumble bee systems, which differ considerably in certain aspects of their biology from honey bees. For example, while honey bees have an exclusively social life, the life of otherwise social bumble bee queens includes a solitary stage during which queens leave their natal colonies to seek mates, diapause, and subsequently found their own colony. This is a potentially sensitive time during the bumble bee life cycle, with dramatic changes in nutrition, thermal environment, and gene expression. These changes have the potential to strongly influence the membership and functionality of the gut microbial community. Using unmated, mated and post-hibernation ovipositing queens of *Bombus lantschouensis*, we characterize variation in the composition and putative function of the gut microbiota and the host hemolymph metabolome. We find that the microbiota varies across the three queen life stages assessed. In particular, we find that mated queens lack many of the common core microbial community members and contains more environmental microbes. Additionally, we find that the queen metabolomic profiles differ by status, and changes in specific metabolites correlate with both functional and taxonomic changes in the bee gut microbial community.

**Introduction**

Division of reproductive labor is a hallmark of eusocial insects [1–4] and results in castes with differing reproductive ability. In eusocial hymenoptera, including honey bees and bumble bees, the queens of colonies represent the primary source of offspring production and thus are key to colony

fitness [5–7]. Previous work highlights the importance of queen health and condition in determining the ultimate colony fitness and performance for many eusocial species [5–7]. Additionally, many diseases and environmental stressors, such as pesticide exposure, influence queen survival and reproductive success [8–10]. These previous studies highlight the importance in understanding factors contributing to queen health.

One important factor that has gained increasing attention as a determinant of social bee health is their gut microbiota. The gut microbial community of social bees is well documented and studied [11–18]. These communities provide their worker bee hosts with nutrition [19–21], defense from pathogens [22,23], and potentially aid in immunological development [24,25]. Additionally, within worker bees these communities are relatively conserved and specialized [15], consisting of just a few core species [13,16], which are passed on largely through vertical transmission from the queen [26]. Some specific core members have been identified as playing key roles. *Gilliamella* aids the host in digesting pectin from pollen [14,19,20], *Lactobacillus* can ferment many of the sugars present in the bee's gut [19,20], and *Snodgrassella* forms a thick biofilm directly on gut epithelial cells modifying the gut environment for the host and other community members [12]. However, previous work shows that the abundances and composition of the bee gut community and the core gut species can vary over time in workers, and with status and life stage in queens [27]. Further work investigating how the life stage and physiological state of bumble bee queens is related to changes in gut community structure is of critical importance to understanding queen health, because of their vital importance for the colony and unique aspects of their biology compared to worker bumble bees and queens of honey bees.

Compared to social honey bees, bumble bees (*Bombus spp.)* differ in that their queens undergo life stages that differ in their social environment [28]. Queen larvae mature and develop within the social context of their natal colony, but following adult emergence queens leave the colony to mate and begin a solitary phase, which includes a diapause phase represented by overwintering hibernation in

temperate species[29]. This period is characterized by extended cold exposure, depressed metabolic functionality, and changes in gene expression [30–34]. After hibernation, queens found their own colonies as oviposition is initiated, and they re-enter social living as their colony begins to grow with the emergence of adult workers. Due to these aspects of their biology, bumble bee queens experience life stages unique from worker bumble bees that are characterized by not only differences in sociality, but also unique physiological demands and gene expression patterns [34–37]. Such physiological changes, and the absence of social buffering and the experience of considerable environmental change means this period is a potentially sensitive, but also critical period in the bumble bee life cycle.

Here we investigate the bumble bee queen hemolymph metabolome and the composition and putative functioning of their gut microbial communities. This is done across three different queen statuses of unmated, after mating and before hibernation, and ovipositing post-hibernation to determine associated shifts in queen physiology and microbial community during these important transition stages. This approach aids us in further understanding the different physiological conditions and requirements of queens as they move through mating, hibernation, and oviposition and how these may be associated with changes in the gut microbial community, which expands our knowledge of bumble bee queen, and ultimately colony, health.

## Materials and Methods

### *Sample collection*

*Bombus lantschouensis* colonies were reared indoors under dark conditions with a constant temperature of 27±1°C and 55-60% humidity. Colonies were fed sugar water (1:1, vol/vol) and apricot pollen *ad libitum*. Three distinct physiological stages of queens produced by these colonies were collected, unmated queens (UQ), mated queens (MQ) and ovipositing queens (OQ). From these samples individuals were assigned to either metagenomic analysis or metabolome analysis. Unmated queens (UQ) were collected from colonies upon adult eclosion and held until 8-days post eclosion. Mated

queens (MQ) were treated as UQ but mated with males 7-days after queen adult eclosion. Mating occurred in a 4m x 3m x 2m (length x width x height) net cage, with queens and males at a ratio of 1:2. Mated queens were sampled 24 hours after mating. Ovipositing queens (OQ) went through the same procedure but were held until they exhibited a decrease in activity, and subsequently cooled to 4°C for hibernation. Four months later, these queens were taken out of hibernation and returned to the standard conditions above, and following egg laying and the emergence of the first workers in the colony the queens were sampled. All samples for metagenomics (MQ, n=5; UQ, n=5; OQ, n=5) were snap-frozen in liquid nitrogen and then stored at -80°C. Queens assigned for metabolomics (MQ, n=11; UQ, n=15; OQ, n=14) were briefly put into a freezer for cold anesthetization followed by hemolymph extraction as described below

***Queen gut DNA extraction***

Each sample assigned for metagenomic analysis was sterilized with 70% ethanol for 1 minute followed by an additional wash of 90% ethanol solution for 1 minute. Samples were then rinsed using double-distilled water several times. The whole alimentary canal (including crop, midgut, ileum and rectum) was removed and placed into a 1.5 mL centrifuge tube containing 100 µL double-distilled water and ceramic beads (0.1 mm) for homogenization. Extracted guts were homogenized using a tissue lyser (Qiagen, Hilden, Germany) and DNA was extracted using a Wizard Genomic DNA purification kit (Promega; A1120) following the manufacturer's instructions with 30 µL of nuclease-free water used for elution. The DNA quality and concentration were assessed using 2% agarose gel electrophoresis and Qubit fluorometer (Invitrogen, Carlsbad, CA, USA), and the DNA solution was stored at -80°C until sequencing.

***Metagenomic sequencing and data processing***

Metagenome sequencing was carried out using the Illumina Hiseq 2500 platform (paired end; reads length, 250 bp; insert size, 300bp). Metagenomic sequencing analysis was done using the

MetaWrap [38] pipeline and ANVIO software [39]. First, raw reads were modified using the MetaWrap quality control module, which removes adaptor sequences and cleans poor quality reads (Phred score less than 20) and short sequences (less than 20bp) with Trimmomatic. Both human and bumble bee associated reads were then removed with BMTagger. The remaining high-quality reads were assembled into sample-based assemblies with metaSpades [40] and subsequently binned (with completion greater than 80% and less than 10% contamination parameter) using Metabat2, Maxbin2, and concoct [41–43]. These resulting bins were further refined with MetaWrap's unique refinement module, and bins with completion of less than 80% and contamination greater than were 10% removed. Bin taxonomy was identified with MetaWrap by searching for 22 single-copy core genes against the Genome Taxonomy Database. Bins were then dereplicated using dRep 2.0 (with bins with 97% or greater similarity dereplicated) and imported into Anvio along with clean reads and sample data. Using Anvio, open reading frames were predicted with Prodigal and annotated with Diamond BLASTp searching against both the NCBI COGs database and KEGG KOFam database [44–47].

***Queen hemolymph extraction***

Hemolymph (30-50 µL) was collected from between the 3 and 4 abdominal segments using a pulled glass capillary to puncture the membrane and a second capillary used to extract the hemolymph. Hemolymph was transferred into Eppendorf tubes and immediately snap-frozen in liquid nitrogen to prevent melanization. The samples were stored at -80°C until used.

*Gas chromatography-mass spectrometry analysis*

Sample metabolites were extracted by using 100% methanol and BSTFA was used for the derivatization of metabolites. The hemolymph metabolites were analyzed using an Agilent 7890 gas chromatograph system coupled with a Pegasus HT time-of-flight mass spectrometer (GC-TOFMS). The system utilized a DB-5MS capillary column coated with 5% diphenyl cross-linked with 95% dimethylpolysiloxane (30m×250µM inner diameter, 0.25µM film thickness; J&W Scientific, Folsom, CA,

USA). A 1µL aliquot of the analyte was injected in splitless mode. Helium was used as the carrier gas, the front inlet purge flow was 3mL $min^{-1}$, and the gas flow rate through the column was 1mL $min^{-1}$. The initial temperature was kept at 50°C for 1 min, then raised to 310°C at a rate of 20°C $min^{-1}$, then kept for 3min at 310°C. The injection, transfer line, and ion source temperatures were 280, 270, and 220°C, respectively. The energy was -70eV in electron impact mode. The mass spectrometry data was acquired in full-scan mode with the m/z range of 50-500 at a rate of 20 spectra per second after a solvent delay of 6.1min.

Chroma TOF 4.3X software (LECO Corporation) and the LECO-Fiehn Rtx5 database were used for subsequent analysis. The workflow for this analysis consisted of extracting the raw peak, determining the data baselines, filtering and calibration of the baselines, peak alignment, deconvolution analysis, peak identification, and integration of the peak area [48]. The RI (retention time index) method was used in the peak identification, and the RI tolerance was 5000. Metabolic features detected in <50% of samples, or those with a relative standard deviation > 30% or the similarity value <200 were removed from the analysis [49]. This resulted in 139 queen metabolites identified and these metabolite abundances were converted to a mean of zero for future analysis.

The MetaboAnalyst 4.0 platform was used to perform the metabolic pathway analysis (MetPA) based on the identified metabolites using the *Drosophila melanogaster* library[50] based on the pathway-related metabolite sets library. The hemolymph metabolite datasets were compared among the three queen groups.

***Statistical analysis***

MetaboAnalyst 4.0 was used to perform multivariate analysis and statistical analysis of the metabolome data [50]. Principal component analysis (PCA) and partial least square-discriminant analysis (PLS-DA) were used to find the significantly different metabolites, with false discovery rate (FDR) adjusted p-values of <0.05 and Variable Importance in Projection (VIP) >1. Microbiota community

analysis was done using the phyloseq R package. Bray-Curtis distance between the samples was calculated and visualized on an ordination plot. A PERMANOVA was then used to test significant differences between the three queen statuses. Calls for each annotated gene were obtained from Anvio [39] and imported into the Statistical Analysis of Metagenomic Profiles (STAMP) software [51]. Using STAMP, differences in microbiota functionality between queen statuses based on genes annotated with either KOFam or COG functions were analyzed using principal component analysis plots and ANOVA with Storey False Discovery Rate correction for multiple tests.

**Results**

***Profiling of the queen gut microbiota***

A total of 171,398 Mb raw data (mean ± sem, 11,426 ± 161 Mb per sample) was obtained by metagenome sequencing. After applying the quality control filters and removing host contigs, a total of 18.36 million base pairs were retained. Initial metagenome assemblies resulted in 5,694 contigs, which assembled into 17 metagenome bins and dereplicated to 6 high quality bins (**Table 5**). Looking at the bin enrichment across samples (**Figure 14**), *Gilliamella* is enriched in all but one sample, however *Snodgrassella* is absent from several of the mated queens. Additionally, the unmated queens are enriched for *Saccharibacter* while the ovipositing queens are enriched for *Bifidobacterium* and *Apilactobacillus*. The ordination plot and PERMAnova on the Bray-Curtis distance between samples supports microbiota clustering based upon queen status (p-value = 0.001) (**Figure 15**). However, often times varying microbial communities may share convergent function at the genomic level. To analyze whether differences in community structure correlate to microbiota with putatively different functions the annotated genes were analyzed in STAMP. From the ordination plots of annotated genes, we see that there is little difference in the functionality for the unmated and ovipositing queen microbial communities, but significant differences in the predicted functions of mated queens (**Figure 16**).

***Hemolymph metabolome differences across queen status***

To assess metabolome differences among the mated, unmated, and ovipositing groups, non-targeted GC-TOFMS was used to analyze the hemolymph profiles. A total of 139 metabolite features were identified in the hemolymph metabolome. For all metabolite features, PCA and PLS-DA analysis was used to detect differences between mated, unmated, and ovipositing groups (**Figure 17**). From this analysis 40 metabolites were significantly different among the mated, unmated, and ovipositing groups. Eight metabolites were more abundant in the mated group, including 4 amino acids and derivatives (alanine, lysine, glutamine, and N-Methyl-DL-alanine), 1 organic acid (4-Acetamidobutyric acid), 1 amide (lactamide), 1 fatty acid (palmitic acid), and 1 sugar (leucrose). Twenty-four metabolites were higher in the unmated group, including 5 amino acids and derivatives (valine, 3-hydroxynorvaline, N-ethylglycine, asparagine, and N-carbamylglutamate), 6 organic acids (Itaconic acid, glutaconic acid, guanidinosuccinic acid, benzoic acid, glutamic acid, and 3-hydroxypyridine), 2 alcohol (stigmasterol and 2-butyne-1,4-diol), 2 sugars (levoglucosan and isopropyl-beta-D-thiogalactopyranoside), 3 nucleotides (uracil, inosine, and cytidine-monophosphate) and 6 other chemical substances (maleamate, 3-methylcatechol, 5-aminovaleric acid lactam, sulfuric acid, pyrophosphate, and 3-hydroxypyridine). Finally, eight metabolites were significantly enriched in ovipositing group, including 3 amino acids (glycine, ornithine, and citrulline), 4 organic acids (2-ketoadipate, 2,4-diaminobutyric acid, 3-aminoisobutyric acid, and malonic acid), 1 amide (maleimide) (Figure 18).

The 40 significantly different hemolymph metabolites were used for metabolic pathway analysis, resulting in enrichment of 11 pathways, including Arginine biosynthesis, alanine, aspartate and glutamate metabolism, D-glutamine and D-glutamate metabolism, glyoxylate and dicarboxylate metabolism, glutathione metabolism, arginine and proline metabolism, beta-alanine metabolism, pyrimidine metabolism, sulfur metabolism, propanoate metabolism, and glycine, serine and threonine metabolism (False discovery rate (FDR) < 0.01, pathway impact > 0.1) (**Figure 18**).

**Discussion**

Queens are critical to the production and subsequent health of social bumble bee colonies, but transition through potentially sensitive life stages with major abiotic environment, social and physiological changes. Here we investigate how the queen life stage status influences their physiological profile associated metabolome and their microbial community. We find clear evidence that both queen metabolomes and their microbial community composition and putative functioning varies across queen status, with strong divergences occurring in queens post-mating.

Several members of the conserved bumble bee and honey bee gut microbiota have been identified as having potential beneficial effects, but the microbial community structure is significantly affected by the queen's life stage and status. Both unmated and ovipositing queens have similarities and are enriched for the core microbe *Snodgrassella*, but have some differential associations with the microbes *Saccharibacter*, *Bidifiobacterium*, and *Apilactobacillus*. Unmated queens have higher levels of *Saccharibacter*, bacteria from this genus have been implicated in food uptake and digestion of acetic acid in other insect systems. Ovipositing queens are enriched for *Bifidobacterium* and *Apilactobacillus*, of which *Bifidobacterium* species of these genera are commonly considered as part of healthy honey bee core microbiota and have implications in food digestion and more importantly defense against other environmental microbes. Pre-hibernation mated queens were not significantly enriched for any specific microbial species besides *Gilliamella*. The lack of consistent gut microbiota likely signals that these individuals are dominated by more environmentally acquired communities and lack the common community structure, in agreement with previous work [27]. Intriguingly, it appears that bumble bee queen mating leads to significant changes in immune gene expression[52]. As highlighted in chapter 4, changes in immunity could affects the standard microbiota composition. Our current metagenomic data does not enable the separation of *Gilliamella* species from the closely related *Schmidhempelia*. Individually characterizing the abundances of these microbes would be advantageous given that they

are negatively associated and have different associations with environmental and other core community members as demonstrated in chapter 4.

The divergence seen in the microbial community structure across the queen life stage and status is supported by divergent predicted community function. In many microbial community variations in taxonomic community structure does not necessarily result in changing community functionality [53–55]. Thus, looking not only at community membership but also the functional traits and genes of these communities is essential to fully elucidate changes in the community and how these may influence queen health. We see significant divergence in the functionality of the microbial communities in mated queens. This means there is the potential for these changes to drive differences in metabolic functions of the microbial community.

To explore the physiology-associated metabolites and the different metabolic processes in the different queen life stages, we measure the hemolymph metabolomic profile of queens. Most of the differential metabolites among the queen groups belonged to amino acids and organic acids. Amino acids are not only consumed as energy sources by core microbial species, but often converted into other metabolic intermediates required for other essential metabolic pathways. For example, tryptophan metabolism has previous been shown to influence honey bee learning and memory behaviors [56], and is an essential component of benefits bees receive from *Lactobacillus spp*. Additionally, glutamine can aid in pyrimidine biosynthesis and serves as an essential nitrogen donor in the hexosamine pathway[57]. Meanwhile, organic acids have previously been shown to play essential roles in lipid storage, pathogen defense, and even modulating interactions between bee microbiota members [19,20,58].

Our study provides further evidence that queen status and stage influences both the physiological state and the gut microbiota of queens. The change of metabolites within queens may drive changes in host microbiota and could be a direct reflection of changing requirements of queen physiology because of life stage and status. However, many studies have suggested metabolome profile

variation could reflect differences in gut microbiome of hosts[59–61]. Thus, the divergent community structures associated with different queen life stages could drive or be driven by changes in host physiology, with important implications for our understanding of queen health. Our current data does not provide a clear causation for whether the divergence in host metabolome drives changes in microbiota or vice-versa. However, tracking both the metabolome and microbiota across a finer time scale through these different queen stages would enable changes in one of these facets to be tied to subsequent changes in the other providing evidence for causation.

**References**


1. Brahma, A., Mandal, S. & Gadagkar, R. Emergence of cooperation and division of labor in the primitively eusocial wasp Ropalidia marginata. *Proc. Natl. Acad. Sci. USA.* **115**, 756–761 (2018).
2. Hartmann, A. & Heinze, J. Lay eggs, live longer: division of labor and life span in a clonal ant species. *Evolution* **57**, 2424–2429 (2003).
3. Kreider, J. J. *et al.* Resource sharing is sufficient for the emergence of division of labour. *Nat. Commun.* **13**, 7232 (2022).
4. Shimoji, H. & Dobata, S. The build-up of dominance hierarchies in eusocial insects. *Phil. Trans. Roy. Soc. Biol. Sci.* **377**, 20200437 (2022).
5. Wiernasz, D. C. & Cole, B. J. Queen size mediates queen survival and colony fitness in harvester ants. *Evolution* **57**, 2179–2183 (2003).
6. Rangel, J., Keller, J. J. & Tarpy, D. R. The effects of honey bee (*Apis mellifera L.*) queen reproductive potential on colony growth. *Ins. Soc.* **60**, 65–73 (2013).
7. Woyke, J. Correlations between the age at which honeybee brood was grafted, characteristics of the resultant queens, and results of insemination. *J. Api. Res.* **10**, 45–55 (1971).
8. Ravoet, J. *et al.* Comprehensive bee pathogen screening in Belgium reveals *Crithidia mellificae* as a new contributory factor to winter mortality. *PLoS One* **8**, e72443 (2013).
9. Yourth, C. P., Brown, M. J. F. & Schmid-Hempel, P. Effects of natal and novel *Crithidia bombi* (Trypanosomatidae) infections on *Bombus terrestris* hosts. *Ins. Soc.* **55**, 86–90 (2008).
10. Fauser, A., Sandrock, C., Neumann, P. & Sadd, B. M. Neonicotinoids override a parasite exposure impact on hibernation success of a key bumblebee pollinator. *Ecol. Entomol.* **42**, 306–314 (2017).

11. Kwong, W. K. & Moran, N. A. Evolution of host specialization in gut microbes: the bee gut as a model. *Gut Microbes* **6**, 214–20 (2015).

12. Kwong, W. K. & Moran, N. A. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family *Neisseriaceae* of the Betaproteobacteria, and *Gilliamella apicola* gen. nov., sp. nov., a member of *Orbaceae* fam. nov., Orbales ord. nov., a sister taxon to the order 'Enterobacteriales' of the Gammaproteobacteria. *Int. J. Syst. Evol. Microbiol.* **63**, 2008–2018 (2013).

13. Kwong, W. K. & Moran, N. A. Gut microbial communities of social bees. *Nat. Rev. Microbiol.* **14**, 374–84 (2016).

14. Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci. USA.* **111**, 11509–14 (2014).

15. Kwong, W. K. *et al.* Dynamic microbiome evolution in social bees. *Sci. Adv.* **3**, e1600513 (2017).

16. Engel, P. & Moran, N. A. Functional and evolutionary insights into the simple yet specific gut microbiota of the honey bee from metagenomic analysis. *Gut Microbes* **4**, 60–5 (2013).

17. Engel, P. & Moran, N. A. The gut microbiota of insects - diversity in structure and function. *FEMS Microbiol. Rev.* **37**, 699–735 (2013).

18. Engel, P. *et al.* The Bee microbiome: impact on bee health and model for evolution and ecology of host-microbe interactions. *mBio* **7**, e02164-15 (2016).

19. Bonilla-Rosso, G. & Engel, P. Functional roles and metabolic niches in the honey bee gut microbiota. *Curr. Opin. Microbiol.* **43**, 69–76 (2018).

20. Kešnerová, L. *et al.* Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol.* **15**, e2003467 (2017).

21. Zheng, H., Powell, J. E., Steele, M. I., Dietrich, C. & Moran, N. A. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc. Natl. Acad. Sci. USA.* **114**, 4775–4780 (2017).

22. Koch, H. & Schmid-Hempel, P. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc. Natl. Acad. Sci. USA.* **108**, 19288–92 (2011).

23. Steele, M. I., Motta, E. V. S., Gattu, T., Martinez, D. & Moran, N. A. The gut microbiota protects bees from invasion by a bacterial pathogen. *Microbiol. Spectr.* **9**, e0039421 (2021).

24. Kwong, W. K., Mancenido, A. L. & Moran, N. A. Immune system stimulation by the native gut microbiota of honey bees. *R. Soc. Open Sci.* **4**, 170003 (2017).

25. Näpflin, K. & Schmid-Hempel, P. Immune response and gut microbial community structure in bumblebees after microbiota transplants. *Proc. Roy. Soc. Biol. Sci.* **283**, (2016).

26. Powell J. Elijah, Martinson Vincent G., Urban-Mead Katherine, & Moran Nancy A. Routes of acquisition of the gut microbiota of the honey b*ee Apis mellifera*. *Appl. Environ. Microbiol.* **80**, 7378–7387 (2014).

27. Wang, L. *et al.* Dynamic changes of gut microbial communities of bumble bee queens through important life stages. *mSystems* **4**, (2019).

28. Sadd, B. M. *et al.* The genomes of two key bumblebee species with primitive eusocial organization. *Gen. Biol.* **16**, 76 (2015).

29. Goulson, D. Bumblebees: behaviour, ecology, and conservation. in (2003).

30. Amsalem, E., Galbraith, D. A., Cnaani, J., Teal, P. E. A. & Grozinger, C. M. Conservation and modification of genetic and physiological toolkits underpinning diapause in bumble bee queens. *Molec. Ecol.* **24**, 5596–5615 (2015).

31. Beekman, M., van Stratum, P. & Lingeman, R. Diapause survival and post-diapause performance in bumblebee queens (*Bombus terrestris*). *Entomol. Experi. Appl.* **89**, 207–214 (1998).

32. Kim, B.G., Shim, J.K., Kim, D.W., Kwon, Y. J. & Lee, K.Y. Tissue-specific variation of heat shock protein gene expression in relation to diapause in the bumblebee *Bombus terrestris*. *Entomol. Res.* **38**, 10–16 (2008).

33. Jedlička, P., Ernst, U. R., Votavová, A., Hanus, R. & Valterová, I. Gene expression dynamics in major endocrine regulatory pathways along the transition from solitary to social life in a bumblebee, *Bombus terrestris*. *Front. Physiol.* **7**, (2016).

34. Costa, C. P. *et al.* Transcriptome analysis reveals nutrition- and age-related patterns of gene expression in the fat body of pre-overwintering bumble bee queens. *Molec. Ecol.* **29**, 720–737 (2020).

35. Nanfack-Minkeu, F. & Sirot, L. K. Effects of mating on gene expression in female insects: unifying the field. *Insects* **13**, (2022).

36. Jasper, W. C., Brutscher, L. M., Grozinger, C. M. & Niño, E. L. Injection of seminal fluid into the hemocoel of honey bee queens (*Apis mellifera*) can stimulate post-mating changes. *Sci. Rep.* **10**, 11990 (2020).

37. Kocher, S. D., Tarpy, D. R. & Grozinger, C. M. The effects of mating and instrumental insemination on queen honey bee flight behaviour and gene expression. *Ins. Molec. Biol.* **19**, 153–162 (2010).

38. Uritskiy, G. V., DiRuggiero, J. & Taylor, J. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**, 158 (2018).

39. Eren, A. M. *et al.* Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* **6**, 3–6 (2021).

40. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).

41. Wu, Y.-W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605–607 (2016).

42. Kang, D. D. *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).

43. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat. Meth.* **11**, 1144–1146 (2014).

44. Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**, D480-4 (2008).

45. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids R.* **45**, D353–D361 (2017).

46. Kanehisa, M. *et al.* Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* **42**, D199–D205 (2014).

47. Galperin, M. Y. *et al.* COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res,* **49**, D274-d281 (2021).

48. Kind, T. *et al.* FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* **81**, 10038–10048 (2009).

49. Dunn, W. B. *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat, Protoc,* **6**, 1060–1083 (2011).

50. Chong, J. *et al.* MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **46**, W486–W494 (2018).

51. Parks, D. H., Tyson, G. W., Hugenholtz, P. & Beiko, R. G. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124 (2014).

52. Barribeau, S. M. & Schmid-Hempel, P. Sexual healing: mating induces a protective immune response in bumblebees. *J. Evol. Biol.* **30**, 202–209 (2017).

53. Biggs, C. R. *et al.* Does functional redundancy affect ecological stability and resilience? A review and meta-analysis. *Ecosphere* **11**, e03184 (2020).

54. Moya, A. & Ferrer, M. Functional redundancy-induced stability of gut microbiota subjected to disturbance. *Trends Microbiol.* **24**, 402–413 (2016).

55. Louca, S. *et al.* Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* **2**, 936–943 (2018).

56. Zhang, Z. *et al.* Honeybee gut *Lactobacillus* modulates host learning and memory behaviors via regulating tryptophan metabolism. *Nat. Commun.* **13**, 2037 (2022).

57. Yoo, H. C., Yu, Y. C., Sung, Y. & Han, J. M. Glutamine reliance in cell metabolism. *Exper. Molec. Med.* **52**, 1496–1516 (2020).

58. Lee, F. J., Miller, K. I., McKinlay, J. B. & Newton, I. L. G. Differential carbohydrate utilization and organic acid production by honey bee symbionts. *FEMS Microbiol. Ecol.* **94**, (2018).

59. Dekkers, K. F. *et al.* An online atlas of human plasma metabolite signatures of gut microbiome composition. *Nat. Communic.* **13**, 5370 (2022).

60. Rooks, M. G. & Garrett, W. S. Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **16**, 341–352 (2016).

61. Li, M. *et al.* Symbiotic gut microbes modulate human metabolic phenotypes. *Proc. Natl. Acad. Sci. USA.* **105**, 2117–22 (2008).

**Table 5.** Metagenome assembled genome information.

| Bin | Completeness | Contamination | GC | Contigs | N50 | Size | Predicted Genes | Annotated Genes | Predicted Taxonomy |
|---|---|---|---|---|---|---|---|---|---|
| bin_4 | 95.38% | 1.74% | 55.26 | 273 | 12,306 | 2,060,000 | 1,868 | 1,352 | *Bifidobacterium* |
| bin_7 | 98.07% | 1.25% | 36.27 | 22 | 99,124 | 1,360,000 | 1,368 | 1,123 | *Apilactobacillus* |
| bin_9 | 98.80% | 0.25% | 50.33 | 31 | 307,112 | 2,220,000 | 2,071 | 1,568 | *Saccharibacter* |
| bin_11 | 98.30% | 0% | 35.48 | 174 | 27,346 | 2,770,000 | 2,468 | 1,786 | *Gilliamella* |
| bin_15 | 99.57% | 1.28% | 42.95 | 42 | 102,728 | 2,280,000 | 2,085 | 1,683 | *Snodgrassella* |
| bin_16 | 92.14% | 5.76% | 42.75 | 2,284 | 4,203 | 7,680,000 | 7,902 | 5,435 | *Paenibacillus* |

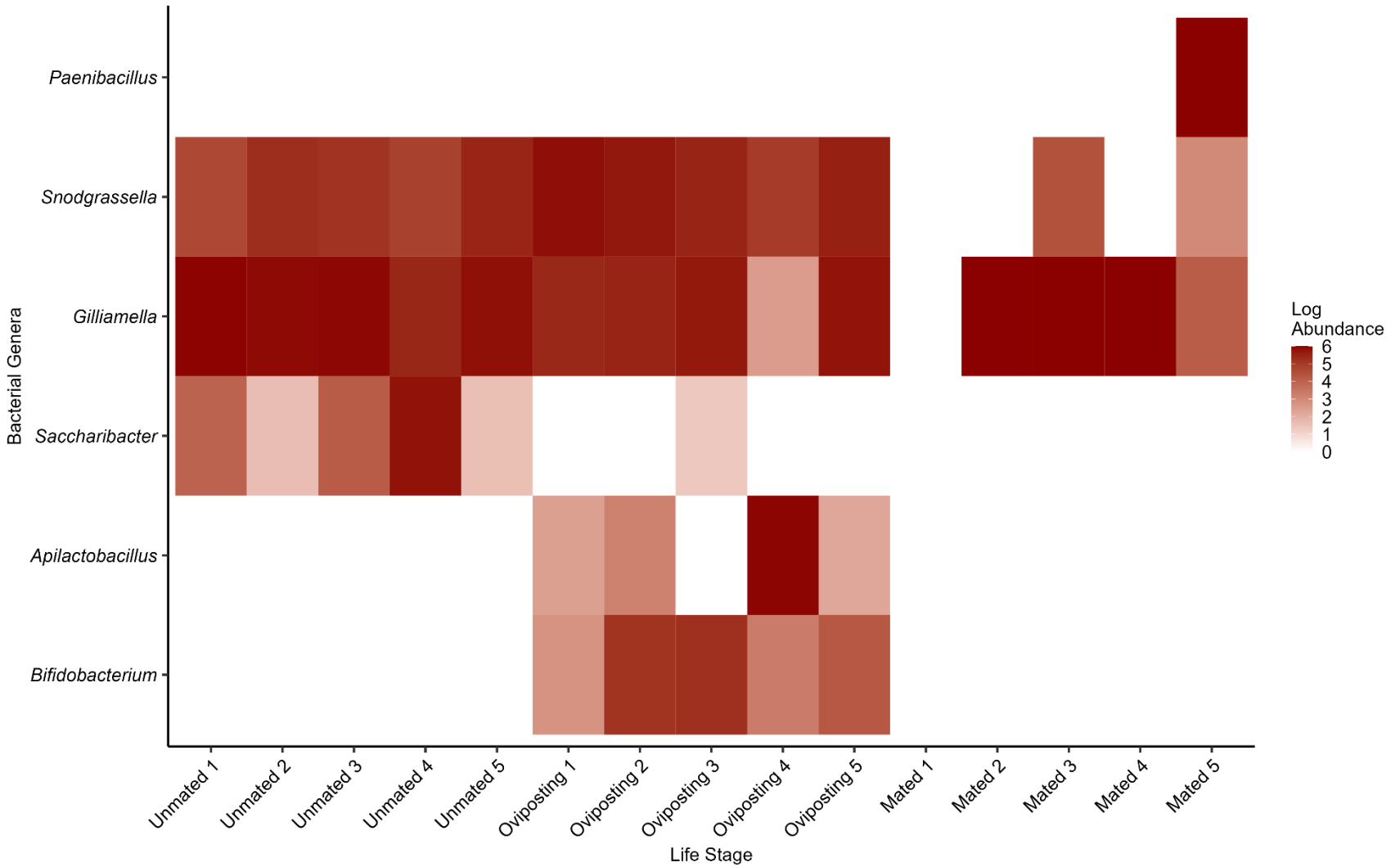


**Figure 14.** Spearman correlation-based networks for immune stimulation treatments of naïve **(A)**, wounding **(B)**, heat-killed E. coli injected **(C)**, and S. epidermidis injected **(D)**. Node color is related to the cluster the bacterial genus belongs to in the network, black lines indicate negative correlations, green lines indicate positive correlations. Node size indicates normalized abundance and line thickness the strength of the interaction.

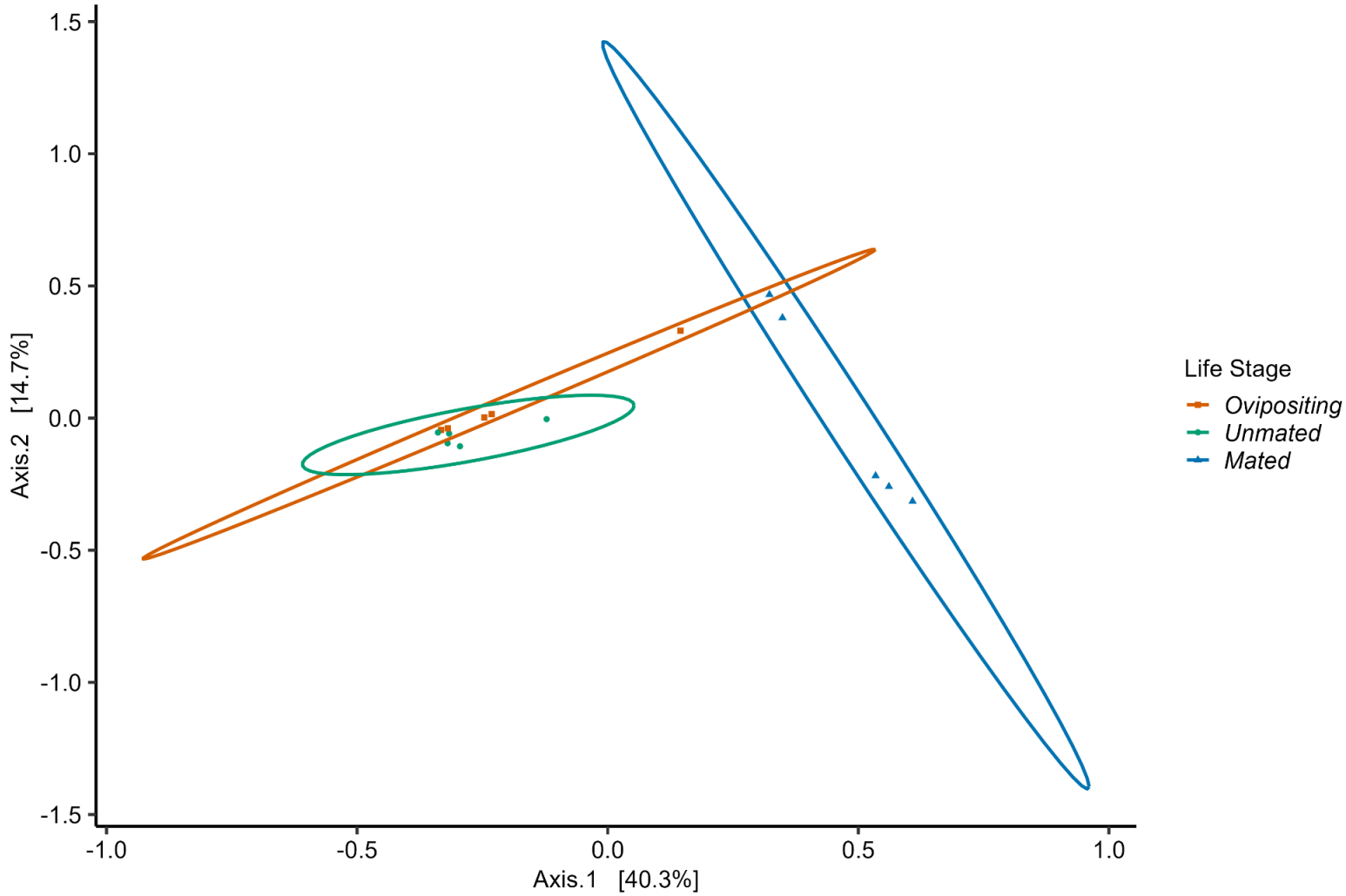


**Figure 15.** Ordination plot of calculated Bray-Curtis distances. Points correspond to individual samples with point shape and color corresponding to queen status. Ellipses demark individual points within each queen status.

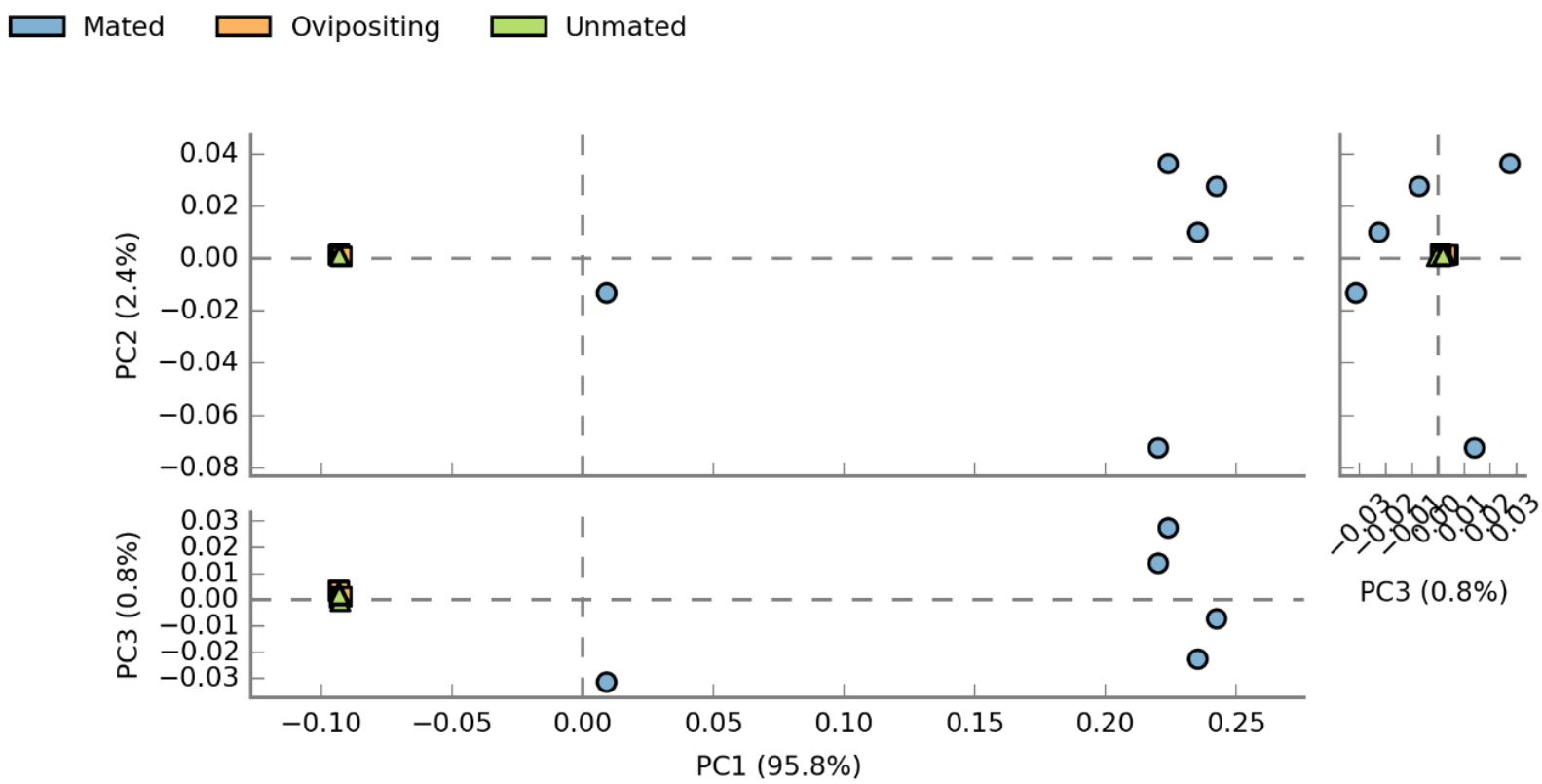


**Figure 16.** Ordination plot created by STAMP from the gene copy abundance of each gene annotated by KEGG or COG annotation. Individual points correspond to samples while the coloration and shape correspond to different queen statuses.

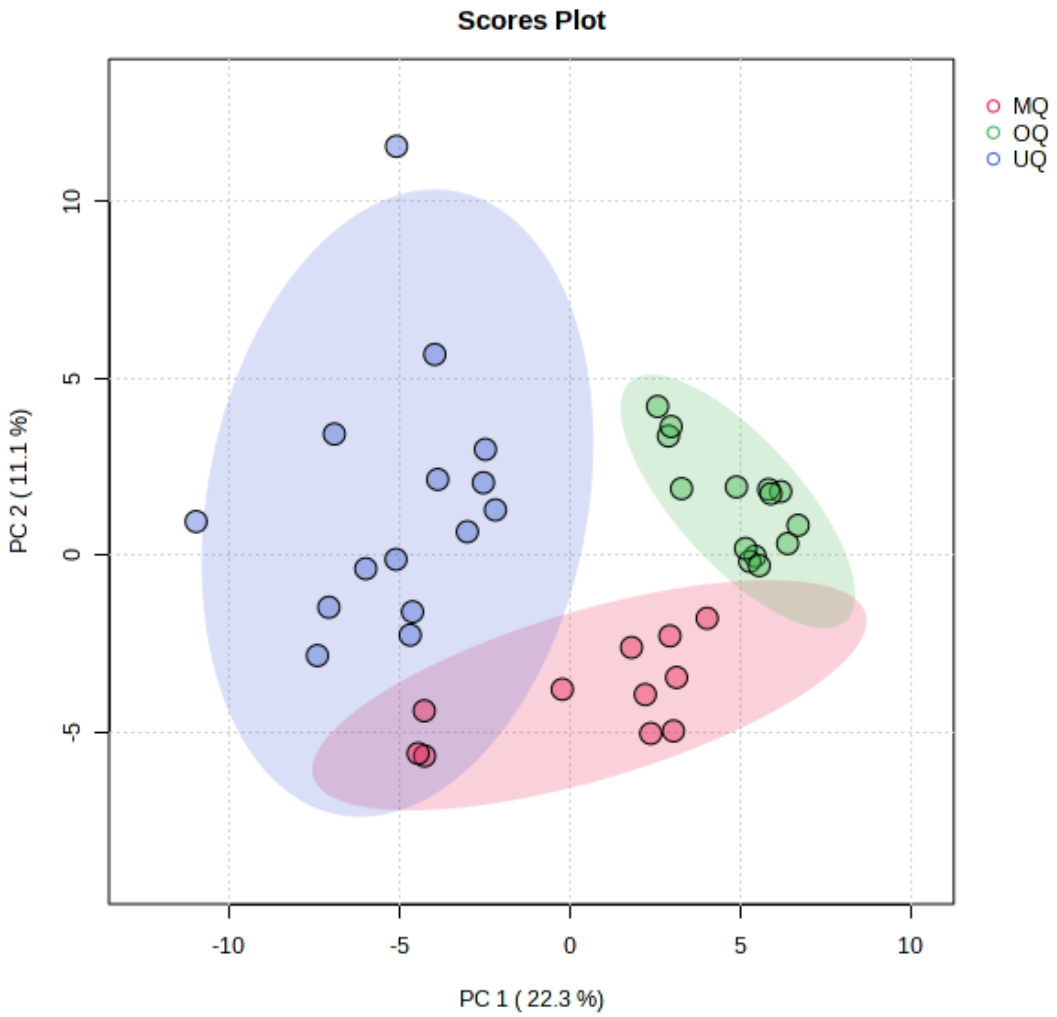


**Figure 17.** Ordination plot of metabolite profiles across queen statuses. Individual points correspond to samples while the coloration and shape correspond to different statuses. The colored regions show the 95% confidence regions.

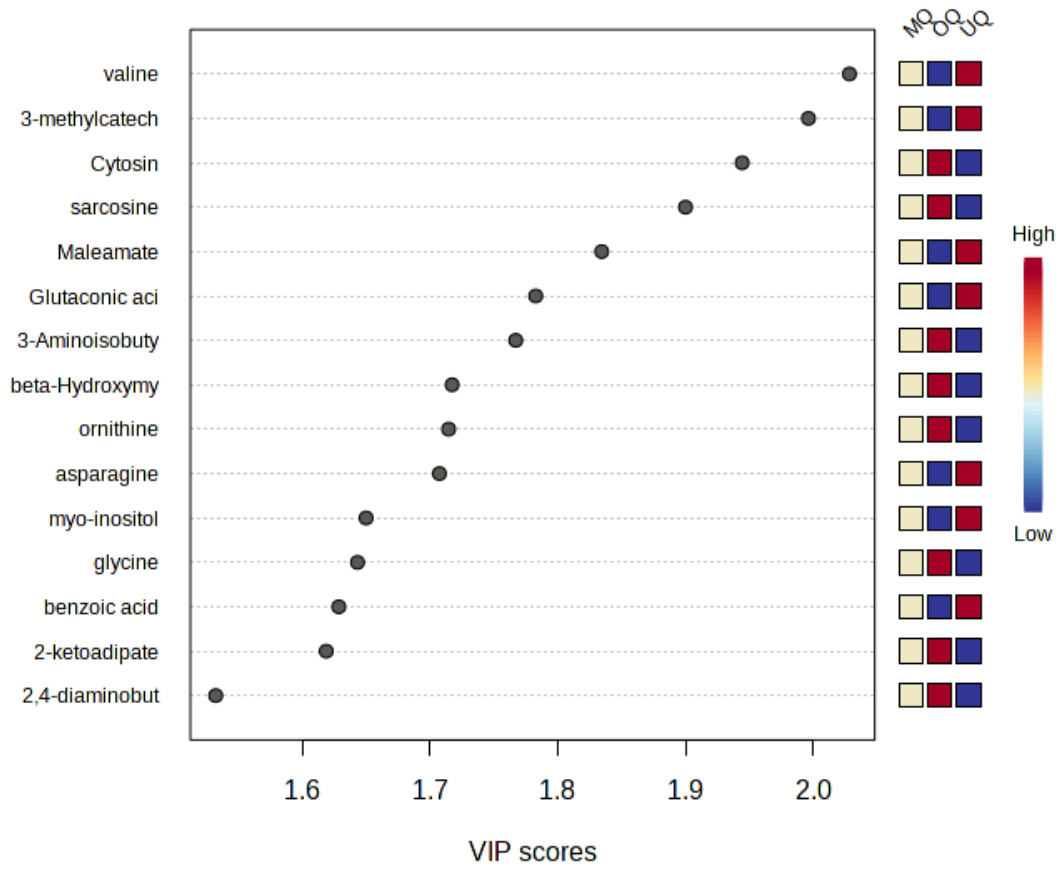


**Figure 18.** Partial least square-discriminant analysis (PLS-DA) identifying the important features among the queen statuses. The metabolites with Variable Importance in Projection (VIP) scores > 1 are listed. The colored boxes on the right show the relative abundance levels of the corresponding metabolite in each group with the mated queen group (MQ) as the reference.

# CHAPTER VI: INVESTIGATING THE ROLE OF INTERMICROBIAL INTERACTIONS IN DETERMINING HOST HEALTH

**Abstract**

Microbial interactions are well documented in many natural ecosystems. However, the role of these interactions in determining the success of microbial communities that are frequently associated with host organisms remains understudied. In these host microbial community systems, we hypothesize that the effects on health of the host may not be accurately predicted from studying component microorganisms in isolation. The gut microbiota of bumble bees (*Bombus spp.*) is made up of a relatively small number of core microbial members, but these members unlikely act alone but rather affect hosts based intrinsic community properties derived from their interactions. To further investigate this hypothesis synthetic communities of combinations of three core bumble bee gut microbes were established in vitro. Following growth of the synthetic communities, culture media was sterile filtered to remove microbial cells and fed to germ-free bees, which were subsequently assayed for several measures relating to host health. Overall, we predict that bee survival in a low nutrition environment and upon pathogen exposure will depend upon its microbial community derived diet, and that the outcomes will be the result of intrinsic properties arising from interactions between the microbes of the microbiota. We find evidence for intrinsic properties in the protein levels of the culture media of the communities. However, we do not find any conclusive evidence for these interactions influencing the measured aspects of bumble bee host health.

**Introduction**

Interacting microbial communities can have important consequences for component community members and the surrounding environment. In addition, the microbiota can have crucial effects on host health and fitness related traits[1–5]. Within these microbial communities metabolic interactions, such as cross-feeding and processing chains, may result in end products that are associated with community

structure[6–11]. These interactions may result from the black queen hypothesis[12], where creation of costly, publicly available goods is selected against[13], or metabolic plasticity where microbes do not fully express their metabolic repertoire unless required[7,14]. The production of specific metabolites depending on interactions with communities are an example of a metabolic intrinsic property[6]. These intrinsic properties are those deriving from a complex community that cannot be predicted based on linearly extrapolation from presence and properties of individual community members. Examples of metabolic intrinsic properties have previously been demonstrated with amylase enzymes for starch[6] and keratin degradation[15]. A result of the presence of metabolic intrinsic properties within host-associated microbiota is that the ultimate benefit a host receives from microbes cannot be accurately predicted from studying the single microbes in isolation.

Pollinating bees are crucial to the health and productivity of natural and managed ecosystems[16–18], but residing within the guts of social bees, such as bumble bees and honey bees are fairly conserved microbial communities predominantly made up of select core members[16–20]. These bee host-associated microbes are stewards of bee health and nutrition[4,21–24]. Unfortunately, despite being a cornerstone in the health of their bee hosts, we lack a complete understanding of how these microbes interact and the impact these interactions have for host health. As such, we also lack necessary information essential to preserving the health of these natural pollinators, many of which such as certain bumble bee species have undergone recent declines[25,26].

One area where research could be more thorough is in investigating how the bumble bee gut microbiota benefits their hosts. Existing research often correlates the bumble bee gut microbial communities with benefits such as weight gain, pathogen defense, and development[4,21–24]. However, these studies have shortcomings that requiring further investigation to disentangle causal relationships between bumble bee microbiota membership and these health-related outcomes. First, no study has directly measured the influence of specific host microbial assemblages on bumble bee nutrient reserves,

a crucial determinate in infection outcomes and hibernation success[27]. Second, many of the previous studies use sequence based amplicon targeting or metagenomics[28,29], which may underestimate important interactions between microbial species. Finally, because of the reliance on sequencing based approaches, few studies experimentally re-establish varying microbial communities, as has been done in honey bees[30]. Experimental re-establishment of microbial communities or their products within otherwise germ-free bee hosts is required to integrate studies on community structure alone into our understandings of the consequences of interactions in these microbial communities for host health.

As an initial investigation into whether we can predict the benefit that bumble bees may receive from the gut microbial communities based on the effects of single members, I cultured in defined nutrient media isolates of the bee associated core microbiota species *Snodgrassella*, *Gilliamella*, and *Lactobacillus* in monoculture, two-member communities, and polyculture. Following community growth, the microbe free media was fed to germ-free bees, with survival tracked during exposure to a low nutrition environment and exposure to a viral pathogen. I hypothesize that the function of microbial communities is dependent upon not only the individual microbial species, but also interactions between community members. Thus, the function of the microbiota is an intrinsic property of the community. This hypothesis generates predictions about how these interactions influence the functional benefits of a host's microbiota. Specifically, I predict that bees fed media from polyculture will have greater nutrient reserves, hence survival under nutrient limitation, and pathogen tolerance than bees fed media from monoculture and that the effects from those receiving monoculture media will underestimate these functional outcomes.

**Materials and Methods**

***Experimental Design***

Germ-free bees from five *Bombus impatiens* colonies were fed one of seven treatments derived from *in vitro* communities of different combinations of core bumble bee gut microbial community

members: i) *Gilliamella bombi* conditioned media, ii) *Lactobacillus* conditioned media, iii) *Snodgrassella alvi* and *G. bombi* conditioned media, iv) *S. alvi* and *Lactobacillus* conditioned media, v) *G. bombi* and *Lactobacillus* conditioned media, vi) *S. alvi*, *G. bombi*, and *Lactobacillus* conditioned media, or vii) a control treatment which consisted of unconditioned, sterile media incubated for the same duration as bacteria inoculated media. Subsequently, health proxies of survival on nutrient limitation[31] and survival after an inoculation of the viral bee pathogen Israelia Acute Paralysis Virus (IAPV)[32] were measured. Media conditioned with *S. alvi* alone could not be used due to a lack of growth of these single microbes in the defined media.

***Bacterial Culturing***

Strains of the common bee gut community members *S.* alvi, *L. bombicola*, and *G. bombi* were isolated from the hindguts of *Bombus impatiens* and preserved in a frozen strain bank, as previously described [33]. The isolated strains were grown on brain heart infusion agar plates for two days after which isolated colonies were used to initiate cultures in brain heart infusion broth. After growing in broth for two days culture optical density was set to 0.1 OD and frozen to -80C with a final glycerol concentration of 25% to create stocks for the strain bank. The experimental media consisted of M9 minimal media as a base, with a mineral supplement, casein hydrolysate, additional trace minerals, vitamin supplements, and heat extracted pollen metabolites as laid out in Kesnerova *et al*[30]. Media was portioned out into 5mL aliquots and assigned microbial treatments added at 1% of total volume from the stocks. Cultures were grown for 72 hours, after which cultures of the same treatment were combined into 50 mL centrifuge tubes, centrifuged at 3,000 rpm for 15 minutes, and then filtered through a sterile filter with a 0.2 mm pore size. The filtered-conditioned media was stored in 2 mL aliquots at -80C until use. Additionally, a subset of the media was used to determine the influence of community composition on the total protein to assess whether differences in the media are a result of the community treatment. Protein concentrations were determined with Pierce-BCA protein assay kit

from ThermoFisher using the manufacturer's well-plate protocol. Due to limitations in stock supplies the full community treatment was not analyzed.

***Germ-free Procedure, Treatment Feeding and Health Measures***

Germ-free bumble bees were created by isolating individuals in the pupal stage. As cohorts of developing larvae are spatially grouped within the colony, identification of cohorts in the pupal stage is relatively simple[33,34]. These pupal stage individuals were removed and held in a sterile holding container until a single worker emerged. This serves as a proxy for age and is important as pupae nearing adulthood are better able to survive the germ-free procedure. Subsequently, the rest of the pupal clump was submerged in 3% bleach solution for 90 seconds and placed into a sterilized holding container. This procedure sterilizes the outer casing, and as bees shed their gut lining during the pupal stage, this leads to emerging adult bees in a germ-free state. Workers emerging from this procedure were assigned treatments and then isolated into sterilized holding containers. Treatments were given by providing 1000 µL of a 1:1 mix of 50% sugar water with conditioned media treatment. After 3 days, consumption was measured, and the sugar water media mixture was replaced. Three days later the consumption was again recorded, and bees in the nutrient limitation treatment were given 25% sugar water. Bees assigned to the IAPV treatment were instead exposed after the first timepoint and subsequently fed 50% sugar water solution. For viral exposures, bees were placed in vials and anesthetized on ice for approximately 15 minutes. Once anesthetized, bees were injected between the first and second abdominal tergite with 2 µL viral inoculum (546 IAPV particles per bee). Bee mortality was recorded every day until all of the samples had died for both the starvation and infection experiments.

***Statistical Analyses***

The media consumption was assessed with a linear mixed effects model with the community treatment as a fixed effect and host colony as a random effect. A cox mixed effects model was used for both the starved and IAPV infected datasets with community treatment as a fixed effect, colony as a

random effect, and consumption as a covariate. Differences in the protein concentration in the media of the treatments were assessed using a generalized linear model with community treatment as a fixed effect and a normal distribution.

**Results**

Microbe community treatment had a significant effect on the amount of protein in the conditioned media ($p = 0.00069$) (**Fig 19** and **Table 6**). Treatments including *Gilliamella* resulted in higher protein levels compared to other treatments. Interestingly, *Lactobacillus* by itself reduced the concentration of protein, but protein levels were restored when other community members were included. We find no significant effects of community structure on the consumption of the conditioned media across either timepoint (**Fig 20** and **Table 6**).

When looking at host survival, we find no effects of community structure on either bee survival during nutrient limitation ($p = 0.924$) or after viral inoculation ($p = 0.647$) (**Fig 21** and **Table 6**). However, we do find significant effects of consumption at both time points for survival on nutrient limitation ($p = 0.001$ and $p = 0.001$ for timepoints 1 and 2 respectively), with higher consumption causing higher mortality in the first time point and higher consumption lowering mortality in the second. For viral infection there is also a significant effect of consumption ($p = 0.014$), with higher consumption at time point 1 lowering mortality (**Table 6**).

**Discussion**

Intrinsic properties arise when the effect of microbial communities cannot be accurately predicted from studying the microbes alone in isolation. In this chapter, I provide an initial test of whether microbial communities within bumble bee guts show any evidence for intrinsic properties and whether there are consequences of intrinsic properties for hosts. We find evidence for intrinsic properties in the protein measurements from the conditioned media. Specifically, showing that *Lactobacillus* alone reduces the protein concentrations of the media, but this reduction is rescued when

*Lactobacillus* is grown in the presence of other community members. This is consistent with the hypothesis that we may not be able to fully understand the influences of these microbes unless studied within a community context [6].

When looking at the influences of potential intrinsic properties from the microbial communities on the measured aspects of host health we find little supporting evidence. Bees did not survive differentally from individuals fed the control media in either the nutrient limitation or the viral exposure experiments. There are several possible explanations for these null results. First, it could be that the media itself contains molecules at levels that are inherently toxic to the bee host, and override any differences. For example, the media contains several trace minerals, such as iron and copper, that may inhibit insect physiology at high enough levels[35,36]. Second, the core microbes may be creating beneficial nutrients for the host, but the process of conditioning media is a static environment, meaning any toxic metabolites will also remain. Within the bee hindgut there is a continuous flow of nutrients in and byproducts out of the digestive system. Thus, the accumulation of toxic metabolites within the conditioned media may offset any benefit the bee receives from the microbes. Finally, previous work has associated up to 80% of metabolic changes between germ free and colonized honey bees to the influences of individual microbes in monoculture, supporting a majority of microbe mediated metabolism can be linked or predicted from individual microbial species[30]. It may be that intrinsic properties could serve a minor role in the bee community concerning metabolism, with benefits from individual metabolic repertoires being more important than these intrinsic properties. However, these results do not rule out intrinsic property effects determining other interactions with the host and the subsequent microbiota-derived host phenotype.

In conclusion, we find some support for intrinsic properties when bee core communities are grown *in vitro*. However, we find no support for the influence of these properties on bee health. Different methodology, such as experimentally recolonizing axenic bees with specific community

structures may aid in further determining the extent of the benefit from microbe interactions. Additionally, metabolic intrinsic properties are not the only important ecological process in bee guts, and experiments further investigating aspects of community structure, such as priority effects would also advance our understanding of the bumble bee gut microbial community.

**References**


1. Brune, A. Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbiol.* **12**, 168–180 (2014).

2. Zheng, H. *et al.* Division of labor in honey bee gut microbiota for plant polysaccharide digestion. *Proc. Natl. Acad. Sci. USA.* **116**, 25909–25916 (2019).

3. Koch, H. & Schmid-Hempel, P. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc. Natl. Acad. Sci. USA.* **108**, 19288–92 (2011).

4. Kwong, W. K., Mancenido, A. L. & Moran, N. A. Immune system stimulation by the native gut microbiota of honey bees. *R. Soc. Open. Sci.* **4**, 170003 (2017).

5. Visick, K. L., Foster, J., Doino, J., McFall-Ngai, M. & Ruby, E. G. *Vibrio fischeri* lux genes play an important role in colonization and development of the host light organ. *J. Bacteriol.* **182**, 4578–86 (2000).

6. Madsen, J. S., Sørensen, S. J. & Burmølle, M. Bacterial social interactions and the emergence of community-intrinsic properties. *Curr. Opin. Microbiol.* **42**, 104–109 (2018).

7. D'Souza, G., Waschina, S., Kaleta, C. & Kost, C. Plasticity and epistasis strongly affect bacterial fitness after losing multiple metabolic genes. *Evolution* **69**, 1244–1254 (2015).

8. Dandekar, A. A., Chugani, S. & Greenberg, E. P. Bacterial quorum sensing and metabolic incentives to cooperate. *Science* **338**, 264–266 (2012).

9. Bonilla-Rosso, G. & Engel, P. Functional roles and metabolic niches in the honey bee gut microbiota. *Curr. Opin. Microbiol.* **43**, 69–76 (2018).

10. Lin, X. B. *et al.* The evolution of ecological facilitation within mixed-species biofilms in the mouse gastrointestinal tract. *ISME J.* **12**, 2770–2784 (2018).

11. Zélé, F., Magalhães, S., Kéfi, S. & Duncan, A. B. Ecology and evolution of facilitation among symbionts. *Nat. Communic.* **9**, 4869 (2018).

12. Morris, J. J. Black Queen evolution: the role of leakiness in structuring microbial communities. *Trends Genet.* **31**, 475–482 (2015).

13. Brockhurst, M. A., Buckling, A., Racey, D. & Gardner, A. Resource supply and the evolution of public-goods cooperation in bacteria. *BMC Biol.* **6**, 20 (2008).

14. Comte, J., Fauteux, L. & Del Giorgio, P. A. Links between metabolic plasticity and functional redundancy in freshwater bacterioplankton communities. *Front. Microbiol.* **4**, 112 (2013).

15. Nasipuri, P. *et al.* Community-intrinsic properties enhance keratin degradation from bacterial consortia. *PLoS One* **15**, e0228108 (2020).

16. Garibaldi, L. A. *et al.* Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science* **339**, 1608–11 (2013).

17. Goulson, D., Lye, G. C. & Darvill, B. Decline and conservation of bumble bees. *Annu. Rev. Entomol.* **53**, 191–208 (2008).

18. Page, M. L. & Williams, N. M. Honey bee introductions displace native bees and decrease pollination of a native wildflower. *Ecology* **104**, e3939 (2023).

19. Kwong, W. K. & Moran, N. A. Gut microbial communities of social bees. *Nat. Rev. Microbiol.* **14**, 374–84 (2016).

20. Engel, P. & Moran, N. A. The gut microbiota of insects - diversity in structure and function. *FEMS Microbiol. Rev.* **37**, 699–735 (2013).

21. Steele, M. I., Motta, E. V. S., Gattu, T., Martinez, D. & Moran, N. A. The Gut microbiota protects bees from invasion by a bacterial pathogen. *Microbiol. Spectr.* **9**, e0039421 (2021).

22. Zheng, H. *et al.* Metabolism of toxic sugars by strains of the bee Gut symbiont *Gilliamella apicola*. *mBio* **7**, (2016).

23. Zheng, H., Powell, J. E., Steele, M. I., Dietrich, C. & Moran, N. A. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc. Natl. Acad. Sci. USA.* **114**, 4775–4780 (2017).

24. Mockler, B. K., Kwong, W. K., Moran, N. A. & Koch, H. Microbiome structure influences infection by the parasite *Crithidia bombi* in bumble bees. *Appl. Environ. Microbiol.* **84**, (2018).

25. Cameron, S. A. & Sadd, B. M. Global trends in bumble bee health. *Annu. Rev. Entomol.* **65**, 209–232 (2020).

26. Potts, S. G. *et al.* Global pollinator declines: trends, impacts and drivers. *Trends Ecol. Evol.* **25**, 345–353 (2010).

27. Beekman, M., van Stratum, P. & Lingeman, R. Diapause survival and post-diapause performance in bumblebee queens (*Bombus terrestris*). *Entomol. Experi. Applic.* **89**, 207–214 (1998).

28. Fischbach, M. A. Microbiome: focus on causation and mechanism. *Cell* **174**, 785–790 (2018).

29. Engel, P. & Moran, N. A. Functional and evolutionary insights into the simple yet specific gut microbiota of the honey bee from metagenomic analysis. *Gut Microbes* **4**, 60–5 (2013).

30. Kešnerová, L. *et al.* Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol.* **15**, e2003467 (2017).

31. Moret, Y. & Schmid-Hempel, P. Survival for Immunity: The Price of Immune System Activation for Bumblebee Workers. *Science* **290**, 1166–1168 (2000).

32. McCormick, E., Cohen, O., Dolezal, A. G. & Sadd, B. M. Consequences of microsporidian prior exposure for virus infection outcomes and bumble bee host health. *Oecologia*.

33. Sauers, L. A. & Sadd, B. M. An interaction between host and microbe genotypes determines colonization success of a key bumble bee gut microbiota member. *Evolution* **73**, 2333–2342 (2019).

34. Näpflin, K. & Schmid-Hempel, P. Immune response and gut microbial community structure in bumblebees after microbiota transplants. *Proc. Roy. Soc. Biol. Sci.* **283**, (2016).

35. Sun, H.-X., Dang, Z., Xia, Q., Tang, W.-C. & Zhang, G.-R. The effect of dietary nickel on the immune responses of *Spodoptera litura Fabricius* larvae. *J. Ins. Physiol.* **57**, 954–961 (2011).

36. Cervera, A., Maymó, A. C., Sendra, M., Martínez-Pardo, R. & Garcerá, M. D. Cadmium effects on development and reproduction of *Oncopeltus fasciatus* (Heteroptera: Lygaeidae). *J. Ins. Physiol.* **50**, 737–749 (2004).

**Table 6.** Model results from the best fitting models for each dataset. Significant effects are bolded.

| **Protein Analysis** | | | |
|---|---|---|---|
| *Parameter* | *F-statistic* | *Df* | *p-value* |
| **Community Structure** | **5.851** | **5** | **0.0007** |
| **Consumption Analysis Time Point 1** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Community Structure | 3.598 | 7 | 0.825 |
| **Consumption Analysis Time Point 2** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| Community Structure | 9.029 | 7 | 0.251 |
| **Survival Analysis** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| **Time Point 1 Consumption** | **10.211** | **1** | **0.001** |
| **Time Point 2 Consumption** | **17.425** | **1** | **0.001** |
| Community Structure | 2.538 | 7 | 0.924 |
| **Viral Analysis** | | | |
| *Parameter* | $X^2$ | *Df* | *p-value* |
| **Time Point 1 Consumption** | **6.023** | **1** | **0.014** |
| Community Structure | 4.220 | 6 | 0.647 |

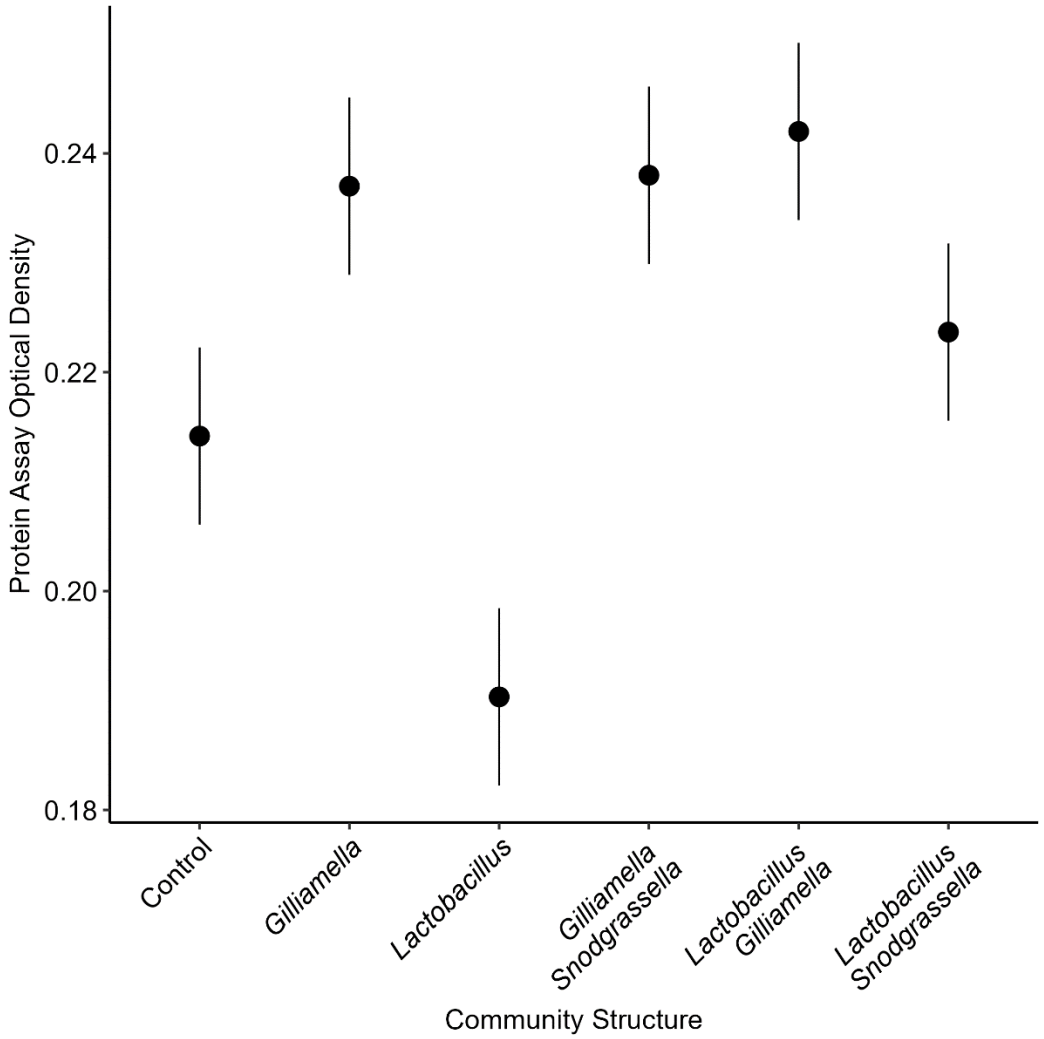


**Figure 19.** Protein levels in media depleted by single microbe and two-microbe community structures. Protein measures were taken as optical densities using the ThermoFisher Pierce BCA Protein Assay kit. Points represent the estimated marginal means and bars represent standard errors for the fitted model with n=6 samples/community structure. The control represents unconditioned, sterile media incubated for the same duration as conditioned media.

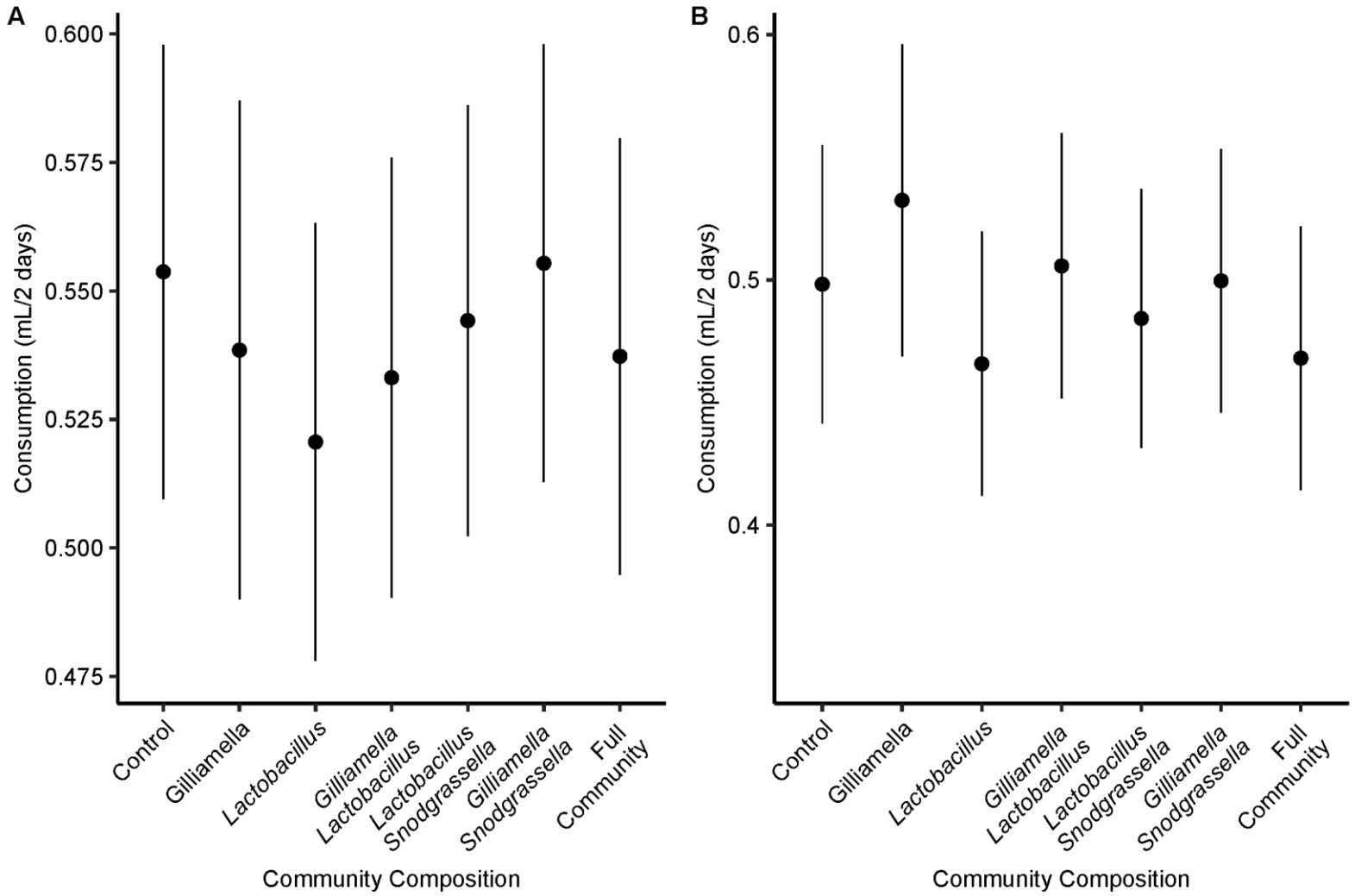


**Figure 20.** Consumption of depleted media treatments for **A)** germ-free bees immediately after emerging and **B)** germ-free bees starting at 3 days after emerging. The points are the estimated marginal means while the error bars are standard errors. Control shows individuals fed a mixture of sugar water with unconditioned, sterile media incubated for the same duration as the conditioned media.

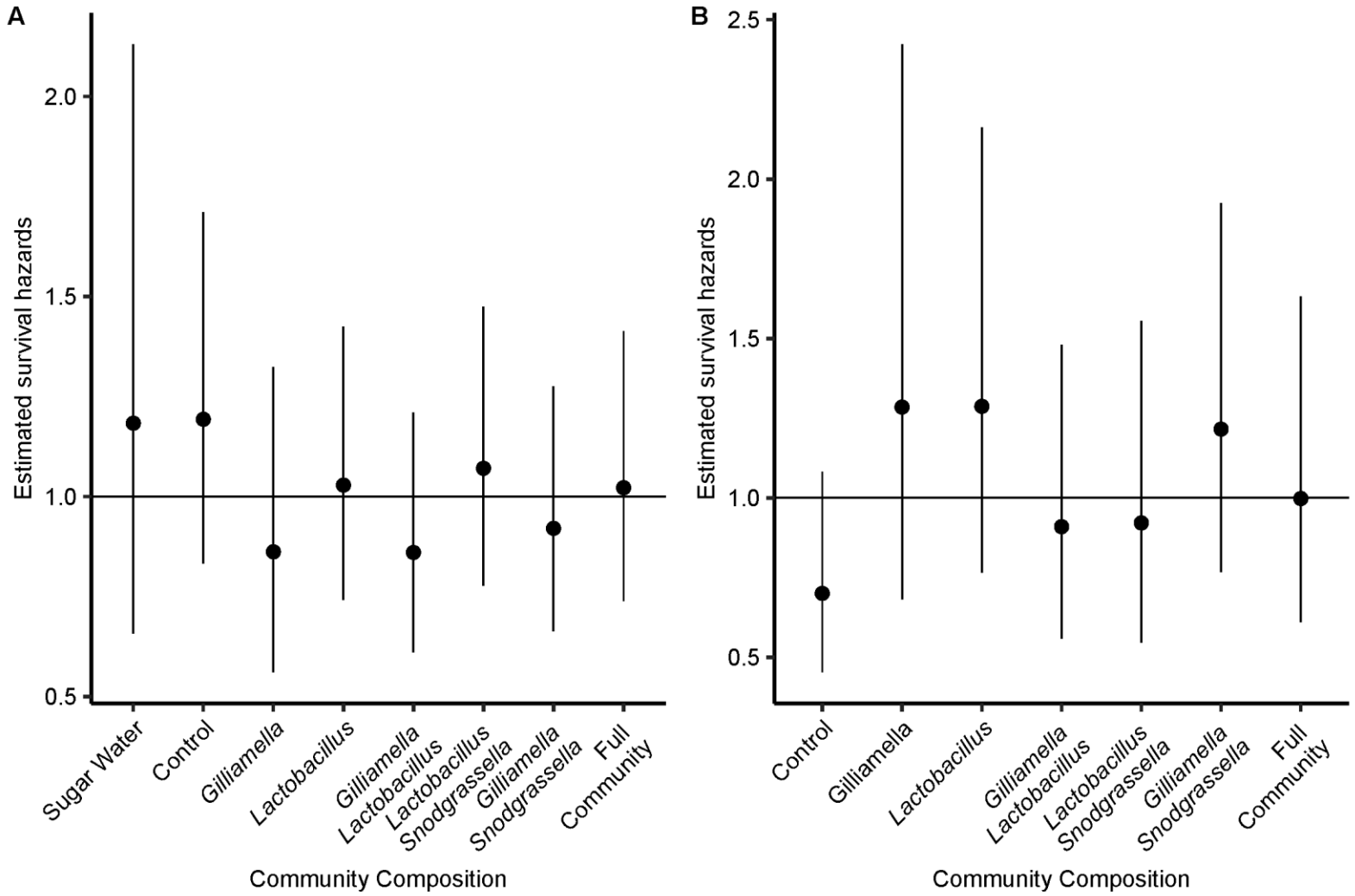


**Figure 21.** The effect of depleted media on bee survival as shown by estimated survival hazards for **A)** starved bees and **B)** IAPV infected bees. Points represent the model-estimated survival hazard and error bars show represent the 95% confidence intervals. Control shows individuals fed a mixture of sugar water with unconditioned, sterile media incubated for the same duration as the conditioned media. Sugar water are bees fed only a sugar water mixture with no media.

# CHAPTER VII: CONCLUSIONS

Overall, I set out to investigate several ecological and evolutionary influences on the composition and functionality of host-associated microbial communities. By utilizing the insect pollinators and particularly the bumble bee as an accessible system to study host-microbiota relationships, I addressed the following questions: (i) how are gut microbial communities structured across diverse insect pollinators, (ii) how do relatedness of hosts based on their genus, species and genotypic lineages, along with microbe genotypes influence gut colonization, (iii) how does stimulation of the host innate immune responses and key changes through the life of a bumble bee queen influence microbiota structure and community associations, and (iv) do interactions between bee associated microbes in communities influence aspects of bee health.

In chapter two, through an international collaboration, I explored the microbiota from diverse insect pollinators. We find that insect genera host unique microbiota, but the structuring of this microbiota is not related to phylogenetic distance among distantly related hosts. Thus, we conclude that the structuring of these communities is more likely driven by host environment, ecology, or diets. Additionally, we discovered microbe in pollinating flies that is related to those previously described from social bee species. We propose a new species within the genera of *Gilliamella,* recommending it be named *Gilliamella eristali*. This microbe is unique from other *Gilliamella* in that it lacks enzymes for digesting pectin and cellulose, transporting cysteine, and flagella assembly. However, this microbe possesses a unique nitrate reduction pathway not described in other species within this clade and genes for ammonia metabolism. Interactions with different hosts but also other members of the microbial community may underlie the evolution of these differences.

In chapter three, I investigated colonization specificity between a core member of the bee gut microbiota across levels of host genera, species, and genotypes. We find further support that host and microbe genotypes determine colonization success. Critically, we also show that across the multiple

strains studied here there is no evidence for broader phylosymbiotic signals or host specificity. Thus, it is important to consider how finer scale interactions between genotypes may influence not only the evolution of microbes and their hosts, but also our understanding of broader patterns of colonization.

In chapter four, I examined how the bumble bee innate immune response targeting pathogens may perturb and disrupt a healthy bee gut microbial community. We find evidence that suggests following stimulation of the immune system there is dysbiosis in the gut microbiota. This could add to traditionally recognized use costs of immune system activation, and critically could influence immune system evolution as and additional selective pressures against ever increasing immune responses.

For chapter five, I investigate how the transition of bumble bee queens through key, solitary phases of the bumble bee colony life cycle effects the their microbiota and metabolomic profiles that may be key to their health during these proposed sensitive periods. Bees from three different queen statuses, either unmated, mated, or ovipositing post-hibernation, were analyzed. We find that mated queens are depauperate in core microbial species and that each of these life stages have unique physiological profiles determined by metabolomics. The mating induced changes documented here could relate back to the immune stimulation induced effects demonstrated in Chapter four, as mating is known to lead to altered immune gene expression.

In chapter six, I sought to address whether we could predict the benefit hosts receive from their microbial communities by studying the members of these communities in isolation or if outcomes are determined by community interaction and cannot be predicted by effects of individual microbes. We evidence for intrinsic properties in composition of growth media of the microbial communities *in vitro*, but there is no conclusive evidence for the role of interactions between microbial species in determining the health of their bee hosts.

Ultimately, the work outlined in this dissertation and my ongoing work at Illinois State University and beyond advances our understanding of the ecological and evolutionary factors that influence the colonization and functioning of host-associated microbial communities.