8-2020

Understanding drivers of inter-population variation in the gut microbiomes of sister species of dung beetles

Claire C. Winfrey
University of Tennessee, Knoxville, cwinfre4@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Claire C. Winfrey entitled "Understanding drivers of inter-population variation in the gut microbiomes of sister species of dung beetles." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Ecology and Evolutionary Biology.

Kimberly S. Sheldon, Major Professor

We have read this thesis and recommend its acceptance:

James A. Fordyce, Benjamin M. Fitzpatrick, Sarah L. Lebeis

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Understanding drivers of inter-population variation in the gut microbiomes of sister species of dung beetles

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Claire Camille Winfrey
August 2020

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my family, Tom, Camilla, Evan, and Dane Winfrey, who have always lovingly encouraged me to follow my dreams. I also acknowledge their generous help during the fieldwork component of this research; I couldn't have done it all without them. I am also very grateful for Nate Brady, who has been a constant source of support throughout my time in Knoxville.

Thanks to my advisor, Dr. Kimberly Sheldon, for always encouraging me to explore my research interests, and for being there for me throughout the ups and downs of my master's project. Dr. Sheldon's patience and mentorship have helped me succeed at UTK, and I will use the skills I learned from her the rest of my scientific career.

I am grateful to my committee members for their support. In particular, I wish to extend a special thanks to my committee member Dr. Jim Fordyce. Dr. Fordyce has always made time to help me, be it answering my (many) statistical questions, providing tips in the lab, or giving me career advice. In addition, I thank my other committee members Drs. Ben Fitzpatrick and Sarah Lebeis, who provided valuable input for the planning and writing stages of this work, and who generously let me use their labs for various aspects of my project.

I would also like to acknowledge others in the Division of Biology at UTK. I appreciate Dr. Brian O'Meara for his dedication to making sure everyone feels welcomed and supported in our department. I thank Dr. Gladys Alexandre and Dr. Sekeenia Haynes for giving me such a wonderful support system in the Program for Equity and Excellence in Research (PEER). Dr. Ben Fitzpatrick and Dr. Mike Blum were fantastic instructors of Core Evolution and Core Ecology. As I wrote this thesis, I frequently returned to my class notes on bedrock papers in ecology and evolutionary biology. Dr. Gary McCracken gave me full access to his well-equipped lab, where I must have spent half of my time on campus. I am also grateful for the help of the staff in EEB's office, especially Marva Anderson. And I thank Judy and Esther, who keep Hesler immaculately clean and never fail to bring a smile to my face. I am also grateful to Veronica Brown and Dr. Robert Murdoch, who gave me many hours of their time to answer my questions about lab and bioinformatics techniques.

This research would not have been possible without the generous financial and research support that I have received. I acknowledge the support of the National Science Foundation's Graduate Research Fellowship, the PEER program funded by the National Institutes of Health, and the Tennessee Fellowship for Graduate Excellence. Not only did these sources support me personally, but they allowed me to fund fieldwork and lab supplies. In addition, I am thankful for funding from the Coleopterists Society, Sigma Xi Grants-in-Aid of Research, and the Department of Ecology & Evolutionary Biology. I thank the Texas Parks and Wildlife Department and the University of Oklahoma Biological Station for generously providing me with accommodations during my fieldwork and advice for finding suitable *Phanaeus* habitat. Finally, this work would not have been possible without the ranchers in Oklahoma and Texas that granted me access to their properties to find *Phanaeus.* Not only did I find the beetles that I needed, but these

ranchers taught me so much about the importance of insects and the soil for sustainable ranching, sparking many additional ecological questions I hope to answer in the next chapter of my academic career.

ABSTRACT

Despite its importance in host physiology, how the microbiome varies within and among populations of hosts is not well understood. However, differential abiotic and biotic selection pressures across a species' range likely lead to variation in the microbiome. In addition, symbiotic microbiota may differ more between closely-related species in sympatry than in allopatry if selection favors the reduction of interspecific competition. We investigated variation in the maternally-transmitted, beneficial gut microbiomes of *Phanaeus vindex* and *P. difformis*, sister species of dung beetle that compete for the same resources in sympatry and occur across a wide range of climatic conditions that may affect their gut microbiota. We sampled and sequenced bacterial/archaeal 16S rDNA from guts of *P. difformis* and *P. vindex* collected across 17 sympatric and allopatric sites, exploring how climatic data, soil microbial diversity, distance between sites, and sympatry or allopatry predicted the observed patterns of gut microbial variation. Gut microbial communities were best predicted by spatial relationships among sampling locations, the abundance of cattle in the sampling area, and temperature and precipitation. Contrary to our hypotheses, we did not find that the gut microbial communities of *P. vindex* and *P. difformis* differed more in sympatry than in allopatry, nor that *P. vindex* exhibits greater turnover in the gut microbiome among populations. However, we found that the gut microbiome communities of *P. vindex* and *P. difformis* both shift between allopatry and sympatry, and that the gut microbiome of *P. vindex* likely experiences a greater shift. While more research is needed, it is possible that differences in their gut microbiomes allow *P. vindex* and *P. difformis* to more effectively partition their niches in sympatry. Our work argues for further exploration of the gut microbiome's potential role in niche partitioning and local adaptation.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

The unicellular organisms found in and on plants and animals, termed the microbiome, are a crucial part of the functioning of many host species. The microbiome plays diverse roles in digestion, metabolism, and immune function (Engel & Moran, 2013), determines the thermal limits of species (Dunbar, Wilson, Ferguson, & Moran, 2007; Kikuchi et al., 2016), and may prevent hybridization (Brucker & Bordenstein, 2013). Despite the link between a host's microbiome and its function, little is known about variation in the microbiome within and among populations of non-human hosts. However, individuals of the same species live in heterogeneous environments, and differential biotic and abiotic pressures may lead to variation in the microbiome across a species' geographic range. To date, most research efforts into variation in the gut microbiome among natural populations have examined hosts from a limited number of locations (e.g. Coon, Brown, & Strand, 2016; Tiede et al., 2017; Parker, Newton, & Moczek, 2020; but see Fietz et al., 2018; Hosokawa, et al., 2016; Wang, Kapun, Waidele, Kuenzel, Bergland, & Staubach, 2020), which are unlikely to be representative of the breadth of environmental conditions across the host's range that may be important in structuring gut microbial communities. Thus, we still lack an understanding of the patterns and processes leading to variation in the gut microbiome of species.

One factor that could lead to variation in the gut microbiome is competition among closely-related taxa. Where ecologically-similar species occur in sympatry, natural selection should favor the species' use of different resources or microhabitats to diffuse competition, which facilitates coexistence (Brown & Wilson, 1956; Hutchinson, 1959; MacArthur, 1958; Schoener, 1974). In contrast, ecologically-similar species occurring in allopatry are predicted to have greater niche overlap than in sympatry because of reduced competition (Brown & Wilson, 1956; Schoener, 1974). If niche partitioning underlies variation in aspects of the niche that are known to affect gut microbial communities, such as temperature or diet, then a host's gut microbiome may be different in sympatry and allopatry. Furthermore, species with similar niches may have more similar gut microbiota where their niches are more similar, in allopatry, than where their niches show greater divergence, where they co-occur. To our knowledge, no studies have investigated how the presence or absence of ecologically-similar species may shape the gut microbiome across the host's range.

Alternatively, different abiotic conditions across the range of the host may also drive variation in the gut microbiome by acting directly on microbes, or by influencing host phenotype that in turn exerts pressure on microbes. For example, aspects of host diet that may vary across geography, such as salinity (Hallali et al., 2018; Wilck et al., 2017), fiber content (Friedman et al., 2017), and prey diversity (Tiede, Scherber, Mutschler, McMahon, & Gratton, 2017), affect gut community composition. Temperature variation across the range of the host may also drive variation in the gut microbiome (Sepulveda & Moeller, 2020). Gut-specialized microorganisms often have narrower thermal tolerances than their hosts (Corbin, Heyworth, Ferrari, & Hurst, 2017; Kikuchi et al., 2016; Zhang, et al., 2019), and thus the community of symbionts may turnover across a hosts' range (Hosokawa et al., 2016). Specifically,

temperatures may alter competitive dynamics within the microbiome, changing the relative abundances of microbial community members (Palmer-Young, Raffel, & McFrederick, 2018).

Variation in gut microorganisms may also be due to deterministic and stochastic processes that positively correlate with distance. The gut microbiome of animals is under a degree of host genetic control (e.g. Benson et al., 2010); thus, as host populations become more genetically distinct with increasing distance, so too might their gut microbiota (Fietz et al., 2018). In addition, the biotic and abiotic conditions that select for the free-living microorganisms present in the environment may explain much of the variation in the horizontally-transmitted gut microbial communities, as has been found in species of flies and mosquitoes (Chandler, Lang, Bhatnagar, Eisen, & Kopp, 2011; Coon, Brown, & Strand, 2016). However, even microorganisms which are passed from mother to offspring may vary over space because these microorganisms were at some point acquired from the environment (Bright & Bulgheresi, 2010), as has been observed in the bacterial endosymbionts of pea aphids (Gauthier et al., 2015) and the gut microbiome of stinkbugs (Hosokawa et al., 2016). In addition, external microbes are subject to stochastic distance-decay processes (Lacap, Lau, & Pointing, 2011; Martiny et al., 2006), such as dispersal limitation, which account for the differences in the gut microbiomes among populations of various hosts, including primates, carnivores, artiodactyls, and rodents (Moeller et al., 2013; Moeller et al., 2017). Finally, ecological drift may also lead to increasing gut microbial dissimilarity between host populations over time (Bordenstein & Theis, 2015; Shafquat, Joice, Simmons, & Huttenhower, 2014).

We used two closely-related species of dung beetles in the genus *Phanaeus* to investigate the patterns and processes structuring the composition of the gut microbiome. Dung beetles develop from egg to adult completely within a ball of dung, or brood ball, that females form below dung pats (Price & May, 2009). Females directly transmit their gut microbiome to their offspring via a bacterial pedestal within the brood ball (Estes et al., 2013). Once established in the larval gut, gut bacteria buffer the organisms against heat and desiccation stresses (Schwab, Riggs, Newton, & Moczek, 2016) and are predicted to fix nitrogen and facilitate the breakdown of cellulose (Shukla, Sanders, Byrne, & Pierce, 2016). This accelerates larval growth and development rates (Schwab, Riggs, Newton, & Moczek, 2016), which are fitness proxies in insects (Kingsolver & Huey, 2008). Therefore, the gut microbiome should be under strong selection.

Phanaeus vindex (MacLeay) and *P. difformis* (LeConte) are sister species of dung beetle (Blume & Aga, 1978) that are morphologically near-identical and share a diet of large mammal dung (Edmonds, 1994; Dickey, 2006). *P. vindex* ranges throughout most of the United States east of the Rocky Mountains, and additionally some populations can be found in desert parts of New Mexico and Arizona. It is noticeably absent in southern and central Texas (Blume & Aga, 1978; Dickey, 2006). *P. difformis* is found in Texas, Oklahoma, Kansas, and parts of adjacent American and Mexican states (Blume & Aga, 1978; Dickey, 2006). Prior studies suggest that *P. difformis* prefers sandy soils, whereas *P. vindex* is found on various soil types, including clay. Because competition for brood ball burial space is thought to be a primary force shaping sympatric dung beetle community structure and breeding and feeding behavior

(Simmons & Risdill-Smith, 2011; Price & May, 2009), different edaphic preferences may play a part in facilitating sympatric co-existence (Blume & Aga, 1978; Edmonds, 1994; Dickey, 2006).

Because competition in sympatry may select for closely-related species to have more distinct niches in sympatry than in allopatry, we hypothesized that the gut microbiota of *P. vindex* and *P. difformis* would be more dissimilar in sympatry than in allopatry. In addition, we hypothesized that the gut microbiome would be more varied (exhibit greater beta diversity) among populations of the broadly distributed, edaphic-generalist *P. vindex*, compared to populations of the narrowly-distributed, sand-specialist *P. difformis* due to the variation in abiotic conditions across the species' ranges.

2. MATERIALS AND METHODS

2.1 Sample collection

We collected *Phanaeus* dung beetles from 5 allopatric *P. difformis* populations, 6 allopatric *P. vindex* locations, and 6 sympatric *P. vindex* and *P. difformis* populations across Texas, Oklahoma, and Kansas (United States) during May and June 2019 in a manner that minimized temporal autocorrelation. For this study, we considered a site allopatric or sympatric based on the species of beetles we encountered in the field. We identified beetles as *P. vindex* or *P. difformis* using a 30x magnifying hand lens. Specifically, we looked for the presence or absence of one or more continuous mid-longitudinal costae on the first or second interstriae of the elytra as well as broad and flat or narrow striae to differentiate species (Edmonds, 1994; Dickey, 2006). We defined a population as all beetles collected within a maximum of 0.5 km^2, and populations were at least 6 km apart. In total, we had 17 sampling locations (Fig. 1, all tables and figures are in appendices at the end of this document). All sampling was done with proper permission and permitting.

We used pitfall traps baited with human or pig dung to collect up to twelve females of each locally present *Phanaeus* species at each sampling location (Table S1). We stored beetles individually in plastic containers, fed them autoclaved cow dung for four days to control for differences in diet consumed prior to capture, and then euthanized them by submersion in 96% molecular grade ethanol (Fisher BioReagents). We sourced all the dung we fed to the beetles from a single organic cattle pasture that was collected on the same morning, homogenized, sterilized by autoclaving twice, and then stored frozen in small batches until use. We cleaned the plastic containers that we stored live beetles in with a 3% bleach solution between uses.

To examine background soil bacterial and archaeal diversity, we took two soil samples per population, each from one meter away in a random direction from a successful beetle trap. We used flame-sterilized tools to take approximately 0.75 g of soil from a depth of 10 cm. We kept soil samples in 96% molecular grade ethanol to match beetle storage conditions. High percentages of ethanol (i.e. 95%-100%) are effective at preserving microbial communities in insect guts and soils, among other sample types (Harry, Gambier, & Garnier-Sillam, 2000; Hammer, Dickerson, & Fierer, 2015; Estes et al., 2013). We stored all beetle and soil samples at 4 °C or on ice for brief periods during transport until further processing.

2.2 DNA extractions, 16S library preparation, and MiSeq sequencing

We used the DNeasy PowerSoil Kit (Qiagen, Venlo, Netherlands) to extract DNA from dissected guts and soil samples in a random order to minimize batch effects. Prior to dissection, we weighed each beetle. We dissected out the whole gut of each beetle using flame-sterilized tools on individual pieces of autoclaved aluminum foil. For soil samples, we dried the samples for approximately 35 minutes in a sterilized laminar flow hood on individual sheets of autoclaved aluminum foil. We followed Qiagen's standard Quick-Start protocol (June 2016), except that we eluted with 50 µL of solution C6 at step 19

and then waited 5 minutes before centrifuging. To characterize the background DNA present in the reagents, we sequenced an extraction kit blank. Although gel electrophoresis post-PCR indicated a very weak band for the extraction blank, NanoDrop spectrophotometry showed that the DNA concentration was too low to be reliably quantified via NanoDrop (<10 ng/ µL). We stored extracted DNA at -20 ºC until amplicon library preparation.

Preparation of the 16S rRNA gene amplicon library was performed by the University of Tennessee, Knoxville Genomics Core in Knoxville, TN, USA. Our protocol followed the 16S Metagenomic Sequencing Library Preparation Workflow published by Illumina Corporation (San Diego, CA, USA), except for three key modifications to the amplicon PCR. First, to amplify the V4 region of the 16S rRNA gene, we used the primers 515 forward and 806 reverse (Caporaso et al., 2011), modified slightly with the addition of two degenerate base pairs (Apprill, McNally, Parsons, & Weber, 2015). The primer sequences of 515F and 806R were GTGYCAGCMGCCGCGGTAA and GGACTACNVGGGTWTCTAAT, respectively. Second, PCR was performed in triplicate, and we reduced each reagent in the amplicon PCR by half and added 0.5 µL of 10 mg/µL bovine serum albumin (Sigma-Aldrich, St. Louis, MO, USA) for a total reaction volume of 13 µL. Finally, the amplicon PCR was carried out in 35 cycles. We ran the amplicon PCR products from each of the triplicate reactions on a 2% agarose gel to verify the amplicon length, then pooled to a volume of 25 µL before performing the index PCR as specified in the Illumina protocol. After completing the index PCR, we pooled the samples to approximately equimolar concentrations using the results from a NanoDrop and confirmed amplicon length and quantity on a Bioanalyzer (Agilent, Santa Clara, CA, United States). We conducted PCR in a UV-sterilized laminar flow hood and included 3 PCR negative controls to verify that minimal DNA contaminants were present in the samples. We were unable to detect PCR negative controls on the agarose gels post PCR, and the DNA concentrations were less than 10 ug/µL when measured via NanoDrop spectrophotometry. To minimize the potential for contamination, we used aerosol barrier pipette tips for all lab work steps.

Illumina MiSeq sequencing was conducted at the University of Tennessee, Knoxville Genomics Core facility using a MiSeq Reagent Kit v3 cycle flow cell (500 base pairs) to obtain two 250 bp paired-end reads. We used a loading concentration of 4 pM, and to increase base pair diversity on the flow cell, we spiked in 20% PhiX control DNA (Illumina). We sequenced 125 *P. vindex* samples, 125 *P. difformis* samples, 34 soil samples, 1 extraction blank, and 3 negative PCR controls for a total of 288 samples on the flow cell.

2.3 Bioinformatics

We conducted all bioinformatic steps using plugins in QIIME2 version 2020.2.0 (Bolyen et al., 2019), all of which were version 2020.2.0 unless otherwise noted. First, we used cutadapt to remove the 515F and 806R primers from our raw reads, allowing an error rate of 0.2 (Martin, 2011). To trim, denoise, dereplicate, and merge our reads, we used DADA2 (Callahan et al., 2016), which resulted in amplicon

sequence variants (ASVs) (Yilmaz et al., 2014). We employed DADA2's default settings and trimmed forward reads at 185 bp and reverse reads at 170 bp. We used scikit-learn version 0.22.1 to train a naïve Bayes taxonomic classifier on the SILVA 128 99% OTUs reference database for our primer region of the 16S gene, and then we used scikit on the trained classifier to assign taxonomy to our sequences. We employed the SATé-enabled phylogenetic placement (SEPP) method (Mirarab, Nguyen, & Warnow, 2012) to place ASVs onto the existing SILVA 128 99% OTUs phylogeny, using the qiime2 plugin fragment-insertion (Janssen et al., 2018; Eddy, 2011; Matsen, Hoffman, Gallagher, & Stamatakis, 2012; Matsen, Kodner, &Armbrust, 2010). SEPP uses hidden Markov models (Finn, Clements, & Eddy, 2011; Eddy, 2011) trained on the reference tree to cluster sequence fragments more accurately and precisely than do traditional de novo tree-building methods with short sequences such as the 16S gene (Janssen et al., 2018), and is now the method that qiime2 recommends. We excluded reads that were not present at least two times, ASVs identified as unassigned at the phylum taxonomic level, and chloroplast and mitochondria sequences.

The remainder of the bioinformatics and the statistical analyses were conducted in R (version 4.0.0, R Core Team, 2020). We exported the ASV tables, taxonomy, and the phylogenetic tree from QIIME2, combining them with a metadata table (see next section for details) in R to create a phyloseq object (phyloseq package, version 1.32.0, McMurdie & Holmes, 2013). We obtained the mean ASV abundance for each sample by rarefying 1,000 times to 3,507 and 26,336 reads for gut and soil samples respectively, using rrarefy in the vegan package available in R (version 2.5-6, Oksanen et al., 2013). Using the rarefied gut communities, we constructed a quantitative Jaccard (Ružička) dissimilarity matrix using vegan's function vegdist and a matrix of weighted UniFrac distances using phyloseq. All analyses were repeated for weighted UniFrac and Jaccard dissimilarities.

2.4 Metadata sourcing

To understand if sympatry and allopatry explain variation in the gut microbiomes of *P. vindex* and *P. difformis*, we had to account for environmental and geographic/spatial factors that may also cause populations of *Phanaeus* to differ in their gut microbiomes. We utilized the GIS map portal of the National Oceanic and Atmospheric Administration (NOAA) to compute average monthly temperature, monthly temperature variability, and average monthly precipitation using data from the years 2009-2018, or the most recent available year prior to 2009 if data from a year between 2009-2018 was unavailable. NOAA measurement stations were no more than 34 km from each sampling location and were on average 20.7 km away. We included an ordinal variable called "Cattle Presence Rank" as an estimate of how many cattle were present at each sampling location. Cattle Presence Rank had three levels: no cows in the beetle collection area or in an area immediately adjacent ("low"), cattle in adjacent area but not present in sampling area ("medium"), or cattle present in sampling area ("high"). To account for the microbial diversity in the beetle's soil habitat that could influence the beetle gut microbiome, we computed the average exponential Shannon entropy index for the soil samples taken from each sampling area.

Averaging the Shannon indices by sampling site was necessary because one soil sample out of the 34 we sequenced failed. See Table S1 for metadata used in this study.

Ecological processes that occur at a variety of spatial scales are important for shaping species interactions (Borcard, Legendre, Avois-Jacquet, & Tuomisto, 2004). Thus, to tease out the effects of spatial autocorrelation on variation in the gut microbiome, we used distance-based Moran's eigenvector map (dbMEM) models (Borcard & Legendre, 2002; Borcard, Gillet, & Legendre, 2018; Borcard, Legendre, Avois-Jacquet, & Tuomisto, 2004; Dray, Legendre, & Peres-Neto, 2006). The dbMEM approach decomposes spatial relationships among sites, resulting in eigenvectors that describe spatial patterns across the entire range of scales detectable given the coordinates of the samples. We used the function dbmem in the package adespatial (version 0.3.8, Dray et al., 2019) to create positive and negative dbMEM eigenvalues in a series of three steps. First, we constructed a pairwise Euclidean distance matrix is among all sites using geodetic, Cartesian coordinates (in our case, taken from the center of each of our sampling locations). Next, we truncated distances in the matrix based on a threshold value equal to the maximum distance between a pair of neighboring sites. We replaced all pairwise distances larger than the threshold, as well as the matrix diagonals, with the threshold value multiplied by four. Thirdly, we computed a PCoA of the truncated distance matrix, the resulting eigenfunctions describing spatial structure at a range of scales. Prior to testing the explanatory power of the constructed dbMEMs, we removed the effects of linear distance among sites by regressing Jaccard and weighted UniFrac dissimilarity matrices on the X-Y coordinates and then saving the model residuals. We implemented distance-based redundancy analyses (db-RDAs, Legendre & Anderson, 1999) and permutational anovas on dbMEMs with positive and negative eigenvalues separately using the detrended dissimilarity matrices as the response variables. If dbMEM variables were significant, we determined which dbMEMs to include in downstream analyses with the help of forward model selection using vegan function ordiR2step. Finally, we applied the Šidák (1967) correction to P-values generated by ordiR2step to account for running multiple tests on positive and negative dbMEMs (Blanchet, Legendre, & Borcard, 2008; Borcard, Gillet, & Legendre, 2018).

We also performed forward model selection separately on environmental variables (i.e. average temperature, temperature variation, average precipitation, cattle presence, and soil habitat microbial biodiversity) using vegan function ordiR2step, which uses Blanchet's double-stopping criterion (Table S2). Following these model selections, we created two separate R objects from the combined geographic variables (significant MEMs and X,Y coordinates because we detected a linear trend in our data) and the significant environmental variables. Performing multiple rounds of model selection before hypothesis testing was necessary to avoid saturating our models with our many initial candidate predictors. In addition, it allowed us to understand how environmental, geographic, and biotic variables may be jointly influencing patterns of gut microbial variation via variation partitioning.

2.5 Examining the drivers of variation in the gut microbiome

To understand the explanatory power of our geographic variables, environmental variables, range overlap, and *Phanaeus* species in relation to variation in the gut microbiome, we performed variation partitioning using the vegan function varpart. Variation partitioning runs independent redundancy analyses on the response variables and each matrix of explanatory variables, and then computes the adjusted R^2 of each fraction via subtraction (Borcard, Legendre, & Drapeau, 1992; Peres-Neto, Legendre, Dray, & Borcard, 2006). Therefore, this technique allowed us to understand the unique and overlapping contributions of all of the types of variables in our models, including those that were correlated.

To test our hypothesis that the gut microbiomes of *P. vindex* and *P. difformis* would be more similar in allopatry than in sympatry, we implemented db-RDAs using the Jaccard and the weighted UniFrac dissimilarity matrices as the response variables. Prior to implementing db-RDAs, we performed global model selection using ordiR2step on the environmental and geographic variables used in variation partitioning, "patry" (sympatry or allopatry), *Phanaeus* species, and beetle mass (Table S3). Based on these model selections, we included X,Y coordinates (i.e. longitude and latitude), five negative dbMEMs, average Shannon entropy of soil samples, cattle presence, *Phanaeus* species, range overlap, and the interaction between *Phanaeus* species and range overlap as predictors in our quantitative Jaccard db-RDA. Despite forward selection identifying average precipitation as a significant variable, we did not use it in our Jaccard model because it had a variance inflation factor (VIF) of over 20, presumably because of a high correlation with longitude (Pearson correlation coefficient = 0.901). For our model using the weighted UniFrac dissimilarity as the response variable, we included the Y coordinates (i.e. latitude), a single significant negative dbMEM, average precipitation, ranked cattle presence, *Phanaeus* species, range overlap, and the interaction between *Phanaeus* species and range overlap as predictors. Other than the interaction term and its variables, all variables had a VIF of 5 or lower. Testing the significance of the interaction between beetle species and range overlap allowed us to explicitly consider if the gut microbiomes of *P. difformis* and *P. vindex* differentially depend on the presence of the other *Phanaeus* species. We investigated the significance of our models' constraints using the vegan function anova.cca with 99,999 permutations. Finally, to separate out the proportions of the variation explained by each predictor, we again performed variation partitioning using the same predictors that were in the db-RDAs.

2.6 Testing differences in beta diversity among populations of *P. vindex* and *P. difformis*

We hypothesized that the gut microbiome of the broadly-distributed, cosmopolitan *P. vindex* would exhibit greater beta diversity than that of the relatively narrowly-distributed, specialist *P. difformis*. Beta diversity describes the variation in species composition among sampling units for a geographic area of interest (Whittaker, 1960). One way of quantifying beta diversity is as the average distance between sampling units and their group centroid in multivariate space, i.e. multivariate dispersion (Anderson, Ellingsen, & McArdle, 2006). However, to obtain beta diversities representing variation in the gut microbiome among populations of *P. vindex* and *P. difformis*, we had to account for uneven sample size,

spatial relationships among our sample sites, and other random factors including cattle presence and soil biodiversity that could influence the gut microbiome from one site to the next. To do this, we first created average ASV tables for each population of *P. vindex* and *P. difformis,* and then obtained pairwise weighted UniFrac and quantitative Jaccard dissimilarities among them. Next, in the same way as described above, we identified dbMEMs and random environmental variables that were important for our models using forward selection. Finally, we regressed Jaccard and weighted UniFrac dissimilarities on significant spatial and environmental predictors. To calculate the multivariate dispersions of *P. vindex* and *P. difformis,* we input the residuals of these models into the function betadisper in vegan. We tested the null hypothesis of no difference between the beta diversities of *P. vindex* and *P. difformis* gut communities using function permutest (99,999 permutations).

2.7 Indicator species analyses

We performed indicator species analyses to identify which gut microbes differed in *P. vindex* and *P. difformis* in sympatry and in allopatry. Specifically, we identified ASVs that were associated with samples of *P. vindex* in sympatry, *P. difformis* in sympatry, *P. vindex* in allopatry, *P. difformis* in allopatry, and combinations of these groups using indicator species analyses (De Cáceres & Legendre, 2009; De Caceres, Legendre, & Moretti, 2010). To do this, we estimated the group equalized, point-biserial correlation coefficients, r_g, between ASVs and groups described above and the significance of these associations using 99,999 permutations with the function multipatt in version 1.7.9 of the package indicspecies (De Cáceres & Legendre, 2009). The point-biserial correlation coefficient takes into account the abundance of each ASV in the group, as well as whether or not it occurs in the other groups under consideration. We then calculated the relative abundances of each indicator species in each sample grouping.

3. RESULTS

3.1 Characterization of *Phanaeus* gut microbial communities

Our MiSeq sequencing run yielded 11,130,657 raw sequences. For all samples, the median number of sequences per sample was 37,458, and for beetle gut and soil samples, the median number of sequences per sample was 33,720.5 and 62,157, respectively. After trimming, merging, and chimera removal with DADA2, 6, 442,702 sequences remained representing 33,488 unique ASVs. Of these, we retained 23,641 bacterial or archaeal ASVs that were assigned to at least the phylum level and that appeared at least twice in our dataset. This included an average of 18,310 sequences per gut sample and an average of 40,699 sequences per soil sample. Rarefaction curves indicated that a sampling depth of 3,500 and 26,336 for gut and soil samples, respectively, was more than adequate to capture the full microbial richness of the communities (Fig. S4). After rarefying, we retained 110 *P. difformis*, 89 *P. vindex*, and 33 soil samples. Because we took the average communities of 1,000 rarefactions, we retained all 23,641 original ASVs, which included 855 ASVs in *P. vindex* guts, 907 ASVs in *P.* difformis guts, and 22,365 ASVs and soil samples.

P. vindex and *P. difformis* gut communities were dominated by the same bacterial phyla and families; however, the relative abundances of the top phyla (Fig. 2) and families (Fig. 3) differed between *P. vindex* and *P. difformis*. The most abundant phyla included Firmicutes (*P. vindex:* 38.91%, *P. difformis:* 42.72%), Proteobacteria (*P. vindex*: 36.76 %, *P. difformis:* 31.44%), Bacteroidetes (*P. vindex*: 15.6% , *P. difformis*: 19.31%), and Actinobacteria (*P. vindex*: 5.17 %, *P. difformis:* 7.12 %). In addition, 1.29% of the ASVs found in *P. vindex* belonged to the phylum Fusobacteria. No other phylum characterized more than 2% of the ASVs in either *Phanaeus* species. The five most abundant bacterial families were Enterococcaceae (*P. vindex*: 23.45% , *P. difformis*: 27.82%), Moraxellaceae (*P. vindex*: 17.86% , *P. difformis*: 10.03%), Porphyromonadaceae (*P. vindex*: 10.82%, *P. difformis*: 16.14%), Enterobacteriaceae (*P. vindex*: 14.34%, *P. difformis*: 13.22%, and Planococcaceae (*P. vindex*: 8.66%, *P. difformis*: 9.17%). All remaining families characterized fewer than 5% of the ASVs we found in *P. vindex* and in *P. difformis*. Although our primers were designed to detect archaea and bacterial 16S rDNA, archaea represented only 0.0032% and 0.0027% of reads in *P. difformis* and *P. vindex,* respectively.

We performed indicator species analyses to identify ASVs that have different affinities for *P. vindex* and *P. difformis* in sympatry and in allopatry, both species in sympatry, and both species in allopatry (Table S5). The identified indicator species comprised a high percentage of total abundance of ASVs. We found 21 indicator species representing 10.25% of allopatric *P. difformis* sequences, 11 indicator species representing 20.61% of sympatric *P. difformis* sequences, 20 indicator species comprising 18.79% of the allopatric *P. vindex* sequences, and 15 species comprising 13.96% of sympatric *P. vindex* sequences. However, only four indicator species identified in each of these four groups comprised more than 1% of that group's total relative abundance. Notable indicator taxa that were both highly correlated and highly abundant included an unidentified ASV in Planococcaceae for allopatric *P. difformis* (r_g=0.267, abundance = 2.35%), two *Vagococcus spp.* in sympatric *P. difformis* samples (r_g =0.363 and 0.202; abundances =13.30% and 3.46%), *Acinetobacter sp*. (r_g=0.278, abundance =7.161%) and an unknown ASV within Enterobacteriaceae (r_g=0.318, abundance = 4.854%) for allopatric *P. vindex*, and an ASV within the genus *Glutamicibacter* (r_g=0.311, abundance =1.22%) for sympatric *P. vindex* samples. Different ASVs within the genus *Dysgonomonas* were identified as indicator taxa for all four groups of *P. vindex* and *P. difformis* samples. Interestingly, an indicator taxon for sympatric *P. difformis* was an ASV within the family Rhodobacteraceae that is a known endosymbiont of the dung beetle *Onthophagus taurus* (r_g=0.196, abundance = 0.013%). This ASV was not found in any other group of samples. In addition, we found up to three indicator ASVs for all combinations of two of the groups described above except for the combination of *P. vindex* and *P. difformis* sympatric samples.

3.2 Contributions of range overlap, *Phanaeus* species, and environmental and spatial variables to gut microbiome structure

We used variation partitioning to explore how the environmental, geographic, and biotic variables ("patry" and *Phanaeus* species identity) contributed to the overall explanation of the gut microbiome data

(Fig. 4). For our analysis on the Jaccard dissimilarity data, the geographic variables (significant dbMEMs and X-Y coordinates), environmental variables, *Phanaeus* species, and sympatry or allopatry individually explained 2.76%, 1.14%, 1.04%, and .25% of the variation in the gut microbiome data respectively, together explaining 6.42% of the variation. In total, our model of the weighted UniFrac distances among gut microbiome samples accounted for 6.24% of the observed variation. The unique contributions to gut microbiome variation of geographic variables, environmental variables, *Phanaeus* species identity, and range overlap were 1.13%, 1.81%, .99%, and 1.1%, respectively. Positive dbMEM eigenvalues did not explain a significant amount of the variation in our detrended dissimilarity matrices (Jaccard: $P = 0.6997$, weighted UniFrac: $P = 0.1257$); thus, they were not included in analyses. More information on model selection is given in Table S2.

Permutational anovas revealed that different variables explained weighted UniFrac and Jaccard dissimilarities among gut samples. Surprisingly, geographic predictors ($P < 0.01$), average precipitation ($P < 0.001$), and cattle presence ($P < 0.05$) were the only significant explanatory variables in our weighted UniFrac db-RDA model (Fig. 5). However, the interaction term between *Phanaeus* species and range overlap was significant ($P < 0.01$) in our model with the Jaccard dissimilarity matrix as the response variable, as were geographic variables ($P < 0.0001$) and level of cattle present in the sampling area ($P < 0.001$) (Fig. 6). Variation partitioning analyses based on the predictors used in the quantitative Jaccard model indicated that geographic, environmental, species identity, and range overlap respectively accounted for 2.84%, .82%, .96%, and 0.4% of variation, and that together, the unique fractions explained 5.68% of the variation in our dataset. In addition, despite our attempt to reduce collinearity, 0.3% of the variation was explained jointly by geography and the environmental variables. For the predictors used in the weighted UniFrac model, we found that the model explained 6.07% of the total variation, and that 2.41%, 1.82%, 1.07%, and 0.62% were explained by geography, environmental variables, *Phanaeus* species identity, and range overlap, respectively.

To understand the significant interaction between *Phanaeus* species and range overlap in our Jaccard model, we ran this model another three times on Jaccard dissimilarity matrices separated by *P. vindex* and *P. difformis* and on a Jaccard dissimilarity matrix based on only sympatric samples. For these models, we did not include the dbMEMs as predictors because they were constructed using all coordinates, including those from sites where only one *Phanaeus* species was collected. Interestingly, we found that range overlap explained a significant amount of the variation in both the *P. vindex* and the *P. difformis* models (*P. vindex* model: $P < 0.01$; *P. difformis* model: $P < 0.05$), while *Phanaeus* species identity was significant in the sympatry model: $P < 0.0001$). All models were also explained by geographic variables (*P. vindex*: $P < 0.001$; *P. difformis*: $P < 0.001$; sympatry model: $P < 0.0001$) and the ranked prevalence of cattle in the area (*P. vindex* model: $P < 0.01$; *P. difformis* and sympatry models: $P < 0.001$). However, the averaged Shannon index of the soil samples was only significant in the sympatry model ($P < 0.01$).

3.4 Comparisons of the beta diversity of *P. vindex* and *P. difformis* populations

Prior to comparing beta diversities, we removed the effects of random environmental variables (i.e. average soil Shannon diversity and the presence of cattle) and geographic distance (no positive or negative dbMEMs were significant) that could cause random differences among populations of *P. vindex* and *P. difformis*. We did not detect a difference in the beta diversities of *P. vindex* and *P. difformis* for Jaccard (P= 0.3895) or weighed UniFrac models (P = 0.1888) (Fig. 7).

4. DISCUSSION

4.1 Summary of results

In this study, we sampled *Phanaeus vindex* and *P. difformis* dung beetles across their allopatric and sympatric ranges to understand the factors influencing variation in the gut microbiome. We did not find that the gut microbiomes of *P. vindex* and *P.difformis* were more similar in allopatry than in sympatry as we expected. However, our db-RDAs based on Jaccard dissimilarities indicated that the gut microbiomes of *P. vindex* and *P. difformis* respond differently to sympatric and allopatric conditions, consistent with the expectations of character displacement. In addition, we hypothesized that because *P. vindex* is found in more diverse habitats and has a larger geographic range than *P. difformis*, its gut microbiome would vary more among populations. While it appears that *P. vindex* has greater beta diversity than does *P. difformis* in our analysis based on weighted UniFrac dissimilarities (Fig. 7), the results were not statistically significant (P = 0.1888), suggesting that we may have needed more populations of both *Phanaeus* species. Overall, geographic distance among populations, environmental variables, and local dung sources were the most important for shaping the gut microbiomes of *P. difformis* and *P. vindex*. Finally, taxa identified as highly abundant or as indicator species echo those found by other dung beetle researchers, suggesting candidate taxa for the core vertically-transmitted gut microbiome of Scarabaeinae dung beetles. We discuss our findings and their implications below.

4.2 Evidence of character displacement in sympatry

Surprisingly, our db-RDAs and subsequent anovas on the full gut microbiome dataset suggested that *Phanaeus* species identity and range overlap were poor predictors of overall variation in the gut microbiome (Figs. 5 and 6). However, in our models based on quantitative Jaccard dissimilarities, we found a significant interaction effect between *Phanaeus* species and range overlap. This indicates that the gut microbiome of one of the *Phanaeus* species shifts more from allopatry to sympatry than does the gut microbiome of its sister *Phanaeus* species. When we performed additional analyses on separated *P. vindex* and *P. difformis* Jaccard dissimilarities to disentangle this interaction effect, we found that the gut microbiomes of both *Phanaeus* species differed based on range overlap. While we cannot validate it statistically, it appears that this trend is stronger among *P. vindex* samples (Fig. S6) than among *P. difformis* samples (Fig. S7), despite the fact that many of our allopatric *P. vindex* locations were geographically very close to where we encountered sympatric *Phanaeus* communities (Fig. 1, Table S1).

Sympatric *Phanaeus* compete fiercely for dung and space to bury brood balls (Price & May, 2009), and thus, ecological theory suggests that their co-existence is predicated on a change in a trait that facilitates niche partitioning (e.g. Brown & Wilson, 1956). We speculate that shifts in the gut microbiome in sympatry, possibly particularly that of *P. vindex*, may be a form of character displacement which reduces competition between *P. vindex* and *P. difformis* in sympatry and allows for co-existence. There are several ways in which a distinct gut microbiome may aid niche partitioning. One possibility is that *P. vindex* and *P. difformis* bury their brood balls at different depths, a strategy found among sympatric *Onthophagus* dung beetles that reduces competition for space below the dung pat to bury brood balls (Macagno, Moczek, & Pizzo, 2016). *P. difformis* usually has an additional point on its front tibiae compared to *P. vindex* (Dickey, 2006; Edmonds, 1994), potentially making it more adept at displacing soil and burying brood balls deeper. Brood balls buried nearer the surface of the soil experience more extreme temperatures and more temperature variation (Snell-Rood, Burger, Hutton, and Moczek, 2016), and the dung beetle gut microbiome increases fitness under thermal stressors (Schwab, Riggs, Newton, & Moczek, 2016). Thus, if *P. vindex* buries its brood balls at shallower depths in sympatry than in allopatry, members of its gut microbiome may be adapted to a broader range of temperatures and thus be better equipped to help developing *P. vindex* larvae digest cellulose and fix nitrogen, likely functions of the dung beetle gut microbiome (Shukla, Sanders, Byrne, & Pierce, 2016). An alternative means of niche partitioning raised in the literature is that *P. vindex*, an edaphic generalist throughout most of its range, is displaced by *P. difformis*, a sand-specialist, on sand where they co-occur (Blume & Aga 1976; 1978). This possibility does not seem likely because we consistently caught both *P. vindex* and *P. difformis* in high abundance in the same trap in sandy collection sites (Table S1). In accordance with the literature however, we observed that *P. difformis* specializes on sandy soils and that *P. vindex* is absent from southern Texas (Fig. 1, Table S1).

We cannot say with certainty why the interaction between species and range overlap was significant in our Jaccard analysis but not our weighted UniFrac analysis. UniFrac calculates the distance between two samples as the proportion of unshared tree branch lengths out of all tree branch lengths (Lozupone & Knight, 2005). If the change detected by Jaccard dissimilarities barely changed the overall phylogenetic signal, it likely would not change the weighted UniFrac distance calculation in a significant way. However, a lack of phylogenetic signal does not preclude the importance of the change in the gut microbiome, because physiologically important traits of gut microbes can differ among bacterial species, or even from strain to strain (e.g. Arnold, Simpson, Roach, Kwintkiewicz, & Azcarate-Peril, 2018). Future research should investigate differences in brood ball burial depths where *P. vindex* and *P. difformis* occur alone and co-occur, coupled with metagenomic and metatranscriptomic approaches to understand if differences in members of the gut microbiome correspond to differences in functionality.

4.3 Local and broad scale trends shape gut microbial communities

Both the weighted UniFrac and the quantitative Jaccard models highlighted that geographic distance among *Phanaeus* communities, a few negative dbMEMs, and cattle presence drove patterns of gut microbial variation (Figs. 5 and 6). However, our first variation partitioning analyses and model selections suggested that precipitation and temperature were also important predictors for gut microbial variation (Table S2, Table S3), but that they were colinear with geographic variables such as latitude and longitude (Fig. 4). Thus, we captured most of the contribution of longitude by including precipitation in our UniFrac model, whereas the effects of temperature variation were likely accounted for by including latitude in both db-RDAs.

Many manipulative lab studies on insects have shown that temperature induces changes in the composition or function of the gut microbiome. For example, insects in general tend to have increases in the relative abundance of Proteobacteria in their guts under heat stress (Sepulveda & Moeller, 2020), and elevated temperatures can cause a loss of beneficial endosymbionts in aphids (Russell & Moran, 2006) and in stinkbugs (Kikuchi et al., 2016). This research is certainly useful for understanding the function of the gut microbiome and its potential response to climate change. However, these studies, which usually source individuals from one population, tell us little about how temperature may interact with other selection pressures and host genetics across a species' range to influence the gut microbiome. Other than two studies on the gut microbiota of the common fruit fly *Drosophila melanogaster* (Walters et al., 2020; Wang, Kapun, Waidele, Kuenzel, Bergland, & Staubach, 2020), ours is the only study to our knowledge that examines the gut microbiome of an insect across a temperature gradient representing much of its geographic range. We know of no other studies that examine how precipitation variation correlates with the gut microbiome. However, because the dung beetle gut microbiome leads to greater fitness outcomes under desiccation and temperature stressors (Schwab, Riggs, Newton, & Moczek, 2016), it follows that selection would favor different gut microbial members under different precipitation and temperature regimes. Alternatively, changes in the gut microbiome may not reflect selection pressures on the host, but instead be the result of selection acting on free-living, environmental microbes encountered in the diet (Rosa, Minard, Lindholm, & Saastamoinen, 2019).

The significance of negative dbMEMs and cattle abundance in our models suggests that local, even site-specific, factors are at work to shape the gut microbiome. In our system, negative spatial autocorrelation (i.e. significant negative dbMEMs) meant that dissimilar gut microbiome samples tended to cluster close to one another in space. Negative spatial autocorrelation is often the signature of biotic interactions (Borcard, Gillet, & Legendre, 2018), and we found that among sympatric samples, the gut microbiome differs based on *Phanaeus* species identity (Fig. S8). Thus, this trend was likely due to structuring within sympatric sampling locations and among allopatric and sympatric *P. vindex* populations that were in close proximity to one another.

In contrast to negative dbMEMs, not a single positive dbMEM was significant, indicating that we did not detect positive spatial autocorrelation in our data. There are a few possible reasons for this. First,

the differences in gut microbial communities with decreased distance among nearby sites may be caused by temperature, precipitation, or other environmental variables that scale with distance that we did not measure. In addition, the amount of cattle at a sampling site, which varied randomly across the landscape, was a strong predictor of gut microbiome function, suggesting factors related to diet are impacting the gut microbiome. As an example, the gut microbiome of beetles in areas with abundant cattle may be a result of priority effects, where the first microbes that beetles encounter in their diet (i.e. a brood ball made from cattle dung) are able to colonize the gut first with limited competition from other microbes, as has been found in the gut microbiomes of honeybees (Ellegaard & Engel, 2019). Another possibility is that microbes specialized to facilitate the digestion of cattle dung are selected for in the dung beetle gut, and they may even be passed on from one generation to the next. An exciting area for future research might be transgenerational studies to tease apart the degrees of vertical versus horizontal transmission characterizing ASVs associated with diet.

4.4 Important taxa of the gut microbiome

Our study confirms that abundant microbes in the gut of *Phanaeus* dung beetles are similar to those found in other genera of dung beetles. Some of the most abundant bacterial families identified here, including Enterobacteriaceae, Comamonadaceae, Moraxellaceae, and Planococcaceae (Fig. 3), are also among the most prevalent in *Onthophagus taurus* (Estes et al., 2013; Hammer et al., 2016) and *Aphodius fossor* (Hammer et al., 2016), and likely perform roles that aid *Phanaeus* in digesting its nutrient-poor dung diet. For example, some members of Enterobacteriaceae perform cellulose digestion and nitrogen fixation in fruit flies (Behar, Yuval, & Jurkevitch, 2005), while others assist bark beetles in uric acid recycling (Morales-Jiménez et al., 2013). In addition, our indicator species analyses revealed that different ASVs within the genus *Dysgonomonas* associate with samples taken from the sympatric and allopatric ranges of both *Phanaeus* species, whereas ASVs within *Acinetobacter* were associated with 3 different groups of *Phanaeus* by patry (Table S5). *Dysgonomonas* has also been found on multiple continents in multiple species of dung beetles (Parker, Newton, & Moczek, 2020), and breaks down lignocellulose in the guts of termites (Sun, Yang, Zhang, Shen, & Ni, 2015). *Acinetobacter* may aid in lipid and ester digester (Kok, Christoffels, Vosman, & Hellingwerf, 1993). Together our results and those from previous researchers suggest strongly that *Acintobacter spp.* and *Dysgonomnas spp.* are likely vertically-transmitted, core members of the dung beetle microbiome.

4.5 Conclusion

This study offers a first look into how biotic interactions and climatic factors across a species range may interact to shape its gut microbiome. We found that the gut microbiomes of *Phanaeus vindex* and *P. difformis* vary across their ranges in patterns predicted by the dung available in the local area, climatic factors such as precipitation and temperature, and due to biotic interactions. Range overlap and species identity were surprisingly poor predictors of variation among *Phanaeus* beetles in our full, pooled

dataset. However, the gut microbiomes of *P. vindex* and *P. difformis* consistently exhibited turnover from allopatric populations to sympatric populations and sympatric communities of *Phanaeus* harbor species-specific gut microbiomes. Furthermore, it appears that gut microbiome of *P. vindex* may shift more between allopatric and sympatric communities than that of *P. difformis*, possibly indicating that the presence of *P. difformis* is driving character displacement of the gut microbiome of *P. vindex* where both *Phanaeus* species co-occur. However, more research is required to assess any physiological changes that may accompany the sympatric shifts in the gut microbiome and their effects on the fitness of *Phanaeus*. Overall, our work emphasizes the need to consider biotic interactions and interpopulation variation in the gut microbiome.

REFERENCES

1. Anderson, M. J., Ellingsen, K. E., & McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. *Ecology letters*, *9*(6), 683-693.
2. Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129-137.
3. Arnold, J. W., Simpson, J. B., Roach, J., Kwintkiewicz, J., & Azcarate-Peril, M. A. (2018). Intra-species genomic and physiological variability impact stress resistance in strains of probiotic potential. *Frontiers in microbiology*, *9*, 242.
4. Behar, A., Yuval, B., & Jurkevitch, E. (2005). Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. *Molecular ecology*, *14*(9), 2637-2643.
5. Benson, A. K., Kelly, S. A., Legge, R., Ma, F., Low, S. J., Kim, J., … & Kachman, S. D. (2010). Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences*, *107*(44), 18933-18938.
6. Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Forward selection of explanatory variables. *Ecology*, *89*(9), 2623-2632.
7. Blume, R. R., & Aga, A. (1976). Phanaeus difformis LeConte (Coleoptera: Scarabaeidae): Clarification of published descriptions, notes on biology, and distribution in Texas. *The Coleopterists' Bulletin*, 199-205.
8. Blume, R. R., & Aga, A. (1978). Observations on ecological and phylogenetic relationships of *Phanaeus difformis* LeConte and *Phanaeus vindex* MacLeay (Coleoptera: Scarabaeidae) in North America. *Southwest Entomology*, 3(2), 113-120.
9. Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., … & Bai, Y. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology*, *37*(8), 852-857.
10. Borcard, D., & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of 18rick18or matrices. *Ecological modelling*, *153*(1-2), 51-68.
11. Borcard, D., Gillet, F., & Legendre, P. (2018). *Numerical ecology with R*. Springer.
12. Borcard, D., Legendre, P., & Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology*, *73*(3), 1045-1055.
13. Borcard, D., Legendre, P., Avois-Jacquet, C., & Tuomisto, H. (2004). Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, *85*(7), 1826-1832.
14. Bordenstein, S. R., & Theis, K. R. (2015). Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PloS biology*, *13*(8), e1002226.
15. Bright, M., & Bulgheresi, S. (2010). A complex journey: transmission of microbial symbionts. *Nature Reviews Microbiology*, *8*(3), 218.
16. Brown, W. L., & Wilson, E. O. (1956). Character displacement. *Systematic zoology*, *5*(2), 49-64.
17. Brucker, R. M., & Bordenstein, S. R. (2013). The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus *Nasonia*. *Science*, *341*(6146), 667-669.
18. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nature method*s, 13(7), 581.
19. Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N. & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the national academy of sciences*, *108*, 4516-4522.
20. Chandler, J. A., Lang, J. M., Bhatnagar, S., Eisen, J. A., & Kopp, A. (2011). Bacterial communities of diverse *Drosophila* species: ecological context of a host–microbe model system. *PloS genetics*, *7*(9), e1002272.
21. Coon, K. L., Brown, M. R., & Strand, M. R. (2016). Mosquitoes host communities of bacteria that are essential for development but vary greatly between local habitats. *Molecular ecology*, *25*(22), 5806-5826.
22. Corbin, C., Heyworth, E. R., Ferrari, J., & Hurst, G. D. (2017). Heritable symbionts in a world of varying temperature. *Heredity*, *118*(1), 10.
23. De Cáceres, M. D., & Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology*, *90*(12), 3566-3574.

24. De Cáceres, M., Legendre, P., & Moretti, M. (2010). Improving indicator species analysis by combining groups of sites. *Oikos*, *119*(10), 1674-1684.
25. Dickey, A. M. (2006). *Population genetics of* Phanaeus vindex *and* P. difformis *and congruence with morphology across a geographic zone of species overlap* (unpublished MS thesis). The University of Texas at Arlington.
26. Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N. & Wagner, H.H. (2019). Adespatial: Multivariate Multiscale Spatial Analysis. R package version 0.3-8. https://CRAN.R-project.org/package=adespatial.
27. Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of 19rick19or matrices (PCNM). *Ecological modelling*, *196*(3-4), 483-493.
28. Dunbar, H. E., Wilson, A. C., Ferguson, N. R., & Moran, N. A. (2007). Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS biology*, *5*(5).
29. Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput Biol*, *7*(10), e1002195.
30. Edmonds, W. D. (1994). Revision of *Phanaeus* Macleay, A New World Genus of Scarabaeinae Dung Beetles (Coleoptera: Scarabaeidae, Scarabaeinae). *Natural History Museum of Los Angeles County Contributions in Science*, 443, 1-105.
31. Ellegaard, K. M., & Engel, P. (2019). Genomic diversity landscape of the honey bee gut microbiota. *Nature communications*, 10(1), 1-13.
32. Engel, P., & Moran, N. A. (2013). The gut microbiota of insects–diversity in structure and function. *FEMS microbiology reviews*, 37(5), 699-735.
33. Estes, A. M., Hearn, D. J., Snell-Rood, E. C., Feindler, M., Feeser, K., Abebe, T., Dunning Hotopp, J.C., & Moczek, A. P. (2013). Brood ball-mediated transmission of microbiome members in the dung beetle, *Onthophagus taurus* (Coleoptera: Scarabaeidae). *PLoS One*, 8(11), e79061.
34. Fietz, K., Hintze, C. O. R., Skovrind, M., Nielsen, T. K., Limborg, M. T., Krag, M. A., Palsbøll, P.J., Hansen, L.H., Møller, P.R., & Gilbert, M. T. P. (2018). Mind the gut: genomic insights to population divergence and gut microbial composition of two marine keystone species. *Microbiome*, *6*(1), 1-16.
35. Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic acids research*, *39*(suppl_2), W29-W37.
36. Friedman, N., Shriker, E., Gold, B., Durman, T., Zarecki, R., Ruppin, E., & Mizrahi, I. (2017). Diet-induced changes of redox potential underlie compositional shifts in the rumen archaeal community. *Environmental microbiology*, *19*(1), 174-184.
37. Gauthier, J. P., Outreman, Y., Mieuzet, L., & Simon, J. C. (2015). Bacterial communities associated with host-adapted populations of pea aphids revealed by deep sequencing of 16S ribosomal DNA. *PloS one*, *10*(3), e0120664.
38. Hallali, E., Kokou, F., Chourasia, T. K., Nitzan, T., Con, P., Harpaz, S., … & Cnaani, A. (2018). Dietary salt levels affect digestibility, intestinal gene expression, and the microbiome, in Nile tilapia (*Oreochromis niloticus*). *PloS one*, *13*(8), e0202351.
39. Hammer, T. J., Dickerson, J. C., & Fierer, N. (2015). Evidence-based recommendations on storing and handling specimens for analyses of insect microbiota. *PeerJ*, *3*, e1190.
40. Hammer, T. J., Fierer, N., Hardwick, B., Simojoki, A., Slade, E., Taponen, J., Viljanen, H., & Roslin, T. (2016). Treating cattle with antibiotics affects greenhouse gas emissions, and microbiota in dung and dung beetles. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1831), 20160150.
41. Harry, M., Gambier, B., & Garnier-Sillam, E. (2000). Soil conservation for DNA preservation for bacterial molecular studies. *European Journal of Soil Biology*, *36*(1), 51-55.
42. Hosokawa, T., Ishii, Y., Nikoh, N., Fujie, M., Satoh, N., & Fukatsu, T. (2016). Obligate bacterial mutualists evolving from environmental bacteria in natural insect populations. *Nature Microbiology*, *1*(1), 15011.
43. Hutchinson, G. E. (1959). Homage to Santa Rosalia or why are there so many kinds of animals?. *The American Naturalist*, *93*(870), 145-159.
44. Janssen, S., McDonald, D., Gonzalez, A., Navas-Molina, J. A., Jiang, L., Xu, Z. Z., Winkler, K., Kado, D.M., Orwoll, E., Manary, M., Mirarab, S., & Knight, R (2018). Phylogenetic placement of exact amplicon sequences improves associations with clinical information. *Msystems*, *3*(3).

45. Kikuchi, Y., Tada, A., Musolin, D. L., Hari, N., Hosokawa, T., Fujisaki, K., & Fukatsu, T. (2016). Collapse of insect gut symbiosis under simulated climate change. *Mbio*, *7*(5), e01578-16.
46. Kingsolver, J. G., & Huey, R. B. (2008). Size, temperature, and fitness: three rules. *Evolutionary Ecology Research, 10(2),* 251-268.
47. Kok, R. G., Christoffels, V. M., Vosman, B., & Hellingwerf, K. J. (1993). Growth-phase-dependent expression of the lipolytic system of Acinetobacter calcoaceticus BD413: cloning of a gene encoding one of the esterases. *Microbiology*, *139*(10), 2329-2342.
48. Lacap, D. C., Lau, M. C., & Pointing, S. B. (2011). Biogeography of prokaryotes. *Biogeography of microscopic organisms: Is everything small everywhere*, 35-42.
49. Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological monographs*, *69*(1), 1-24.
50. Lozupone, C., & Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Applied and environmental microbiology, 71(12), 8228-8235.
51. Macagno, A. L., Moczek, A. P., & Pizzo, A. (2016). Rapid divergence of nesting depth and digging appendages among tunneling dung beetle populations and species. The American Naturalist, 187(5), E143-E151.
52. MacArthur, R. H. (1958). Population ecology of some warblers of northeastern coniferous forests. *Ecology*, *39*(4), 599-619.
53. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, *17*(1), 10-12.
54. Martiny, J. B. H., Bohannan, B. J., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., … & Morin, P. J. (2006). Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102.
55. Matsen, F. A., Hoffman, N. G., Gallagher, A., & Stamatakis, A. (2012). A format for phylogenetic placements. *PLoS One*, *7*(2), e31009.
56. Matsen, F. A., Kodner, R. B., & Armbrust, E. V. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC bioinformatics*, *11*(1), 538.
57. McMurdie, P. J., & Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, *8*(4), e61217.
58. Mirarab, S., Nguyen, N., & Warnow, T. (2012). SEPP: SATé-enabled phylogenetic placement. In *Biocomputing 2012* (pp. 247-258).
59. Moeller, A. H., Peeters, M., Ndjango, J. B., Li, Y., Hahn, B. H., & Ochman, H. (2013). Sympatric chimpanzees and gorillas harbor convergent gut microbial communities. *Genome research*, *23*(10), 1715-1720.
60. Moeller, A. H., Suzuki, T. A., Lin, D., Lacey, E. A., Wasser, S. K., & Nachman, M. W. (2017). Dispersal limitation promotes the diversification of the mammalian gut microbiota. *Proceedings of the National Academy of Sciences*, *114*(52), 13768-13773.
61. Morales-Jiménez, J., de León, A. V. P., García-Domínguez, A., Martínez-Romero, E., Zúñiga, G., & Hernández-Rodríguez, C. (2013). Nitrogen-fixing and uricolytic bacteria associated with the gut of *Dendroctonus rhizophagus* and *Dendroctonus valens* (Curculionidae: Scolytinae). *Microbial ecology*, *66*(1), 200-210.
62. Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin P.R., O'Hara, R., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H.J.C. (2013). Vegan: community ecology package. R Package version 2.0. http:// CRAN.Rproject.org/package=vegan.
63. Palmer-Young, E. C., Raffel, T. R., & McFrederick, Q. S. (2018). Temperature-mediated inhibition of a bumblebee parasite by an intestinal symbiont. *Proceedings of the Royal Society B*, *285*(1890), 20182041.
64. Parker, E. S., Newton, I. L., & Moczek, A. P. (2020). (My Microbiome) Would Walk 10,000 miles: Maintenance and Turnover of Microbial Communities in Introduced Dung Beetles. *Microbial Ecology*, 1-12.
65. Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, *87*(10), 2614-2625.
66. Price, D. L. & May, M. (2009). Behavioral ecology of *Phanaeus* dung beetles (Coleoptera: Scarabaeidae): review and new observations. *Acta Zoológica Mexicana (ns)*, *25*(1), 211-238.

67. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.
68. Rosa, E., Minard, G., Lindholm, J., & Saastamoinen, M. (2019). Moderate plant water stress improves larval development, and impacts immunity and gut microbiota of a specialist herbivore. *PLoS One*, *14*(2), e0204292.
69. Russell, J. A., & Moran, N. A. (2006). Costs and benefits of symbiont infection in aphids: variation among symbionts and across temperatures. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1586), 603-610.
70. Schoener, T. W. (1974). Resource partitioning in ecological communities. *Science*, *185*(4145), 27-39.
71. Schwab, D. B., Riggs, H. E., Newton, I. L., & Moczek, A. P. (2016). Developmental and ecological benefits of the maternally transmitted microbiota in a dung beetle. *The American Naturalist*, 188(6), 679-692.
72. Sepulveda, J., & Moeller, A. H. (2020). The effects of temperature on animal gut microbiomes. *Frontiers in Microbiology*, *11*.
73. Shafquat, A., Joice, R., Simmons, S. L., & Huttenhower, C. (2014). Functional and phylogenetic assembly of microbial communities in the human microbiome. *Trends in microbiology*, *22*(5), 261-266.
74. Shukla, S. P., Sanders, J. G., Byrne, M. J., & Pierce, N. E. (2016). Gut microbiota of dung beetles correspond to dietary specializations of adults and larvae. *Molecular Ecology*, 25(24), 6092-6106.
75. Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, *62*(318), 626-633.
76. Simmons, L. W., & Ridsdill-Smith, T. J. (2011). Reproductive competition and its impact on the evolution and ecology of dung beetles. *Ecology and evolution of dung beetles. Oxford: Blackwell Publishing*, 1-20.
77. Snell-Rood, E. C., Burger, M., Hutton, Q., & Moczek, A. P. (2016). Effects of parental care on the accumulation and release of cryptic genetic variation: review of mechanisms and a case study of dung beetles. *Evolutionary ecology*, *30*(2), 251-265.
78. Soil Science Division Staff. (2017). Examination and Description of Soil Profiles. In C. Ditzler, K. Scheffe, and H.C. Monger (Eds.), *Soil Survey Manual*, *USDA Handbook 18* (pp. 83-233). Washington, D.C.: Government Printing Office.
79. Sun, X., Yang, Y., Zhang, N., Shen, Y., & Ni, J. (2015). Draft genome sequence of *Dysgonomonas macrotermitis* strain JCM 19375T, isolated from the gut of a termite. *Genome announcements*, *3*(4).
80. Tiede, J., Scherber, C., Mutschler, J., McMahon, K. D., & Gratton, C. (2017). Gut microbiomes of mobile predators vary with landscape context and species identity. *Ecology and evolution*, *7*(20), 8545-8557.
81. Walters, A. W., Hughes, R. C., Call, T. B., Walker, C. J., Wilcox, H., Petersen, S. C., Rudman, S.M., Newell, P.D., Douglas, A.E., Schmidt, P.S., & Chaston, J. M. (2020). The microbiota influences the Drosophila melanogaster life history strategy. *Molecular Ecology*, *29*(3), 639-653.
82. Wang, Y., Kapun, M., Waidele, L., Kuenzel, S., Bergland, A. O., & Staubach, F. (2020). Common structuring principles of the Drosophila melanogaster microbiome on a continental scale and between host and substrate. *Environmental Microbiology Reports*, *12*(2), 220-228.
83. Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California. *Ecological monographs*, *30*(3), 279-338.
84. Wilck, N., Matus, M. G., Kearney, S. M., Olesen, S. W., Forslund, K., Bartolomaeus, H., Hasse, S., Mäher, A., Balogh, A., Markó, L., & Vvedenskaya, O. (2017). Salt-responsive gut commensal modulates T H 17 axis and disease. *Nature*, *551*(7682), 585.
85. Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., … & Glöckner, F. O. (2014). The SILVA and "all-species living tree project (LTP)" taxonomic frameworks. *Nucleic acids research*, 42(D1), D643-D648.
86. Zhang, B., Leonard, S. P., Li, Y., & Moran, N. A. (2019). Obligate bacterial endosymbionts limit thermal tolerance of insect host species. *Proceedings of the National Academy of Sciences*, *116*(49), 24712-24718.

APPENDICES

Appendix A: Figures

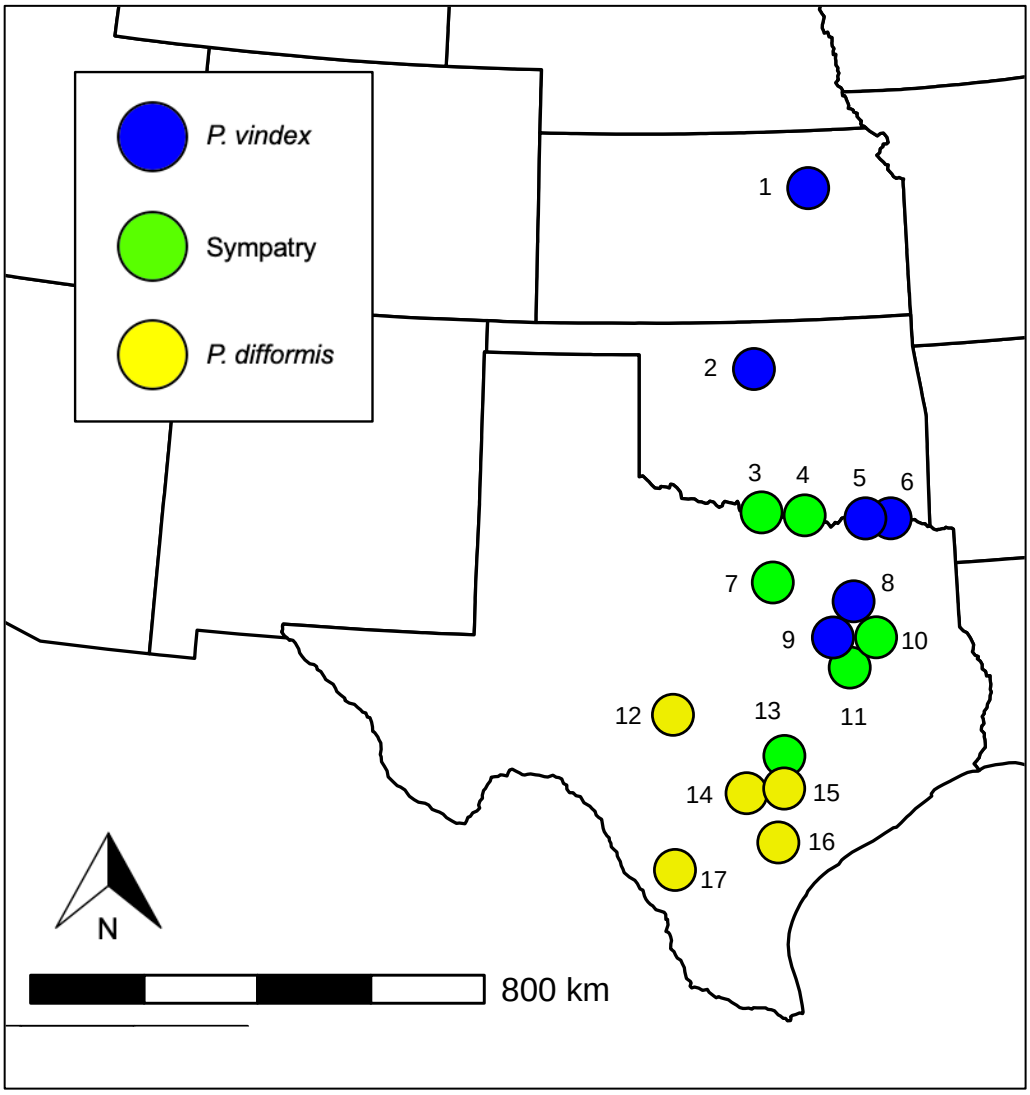


Figure 1: Map showing collection localities of *Phanaeus vindex* and *P. difformis* samples in the states of Kansas, Oklahoma, and Texas (USA). For legibility, points close to each other have been moved slightly. See Table S1 for details about sampling sites.

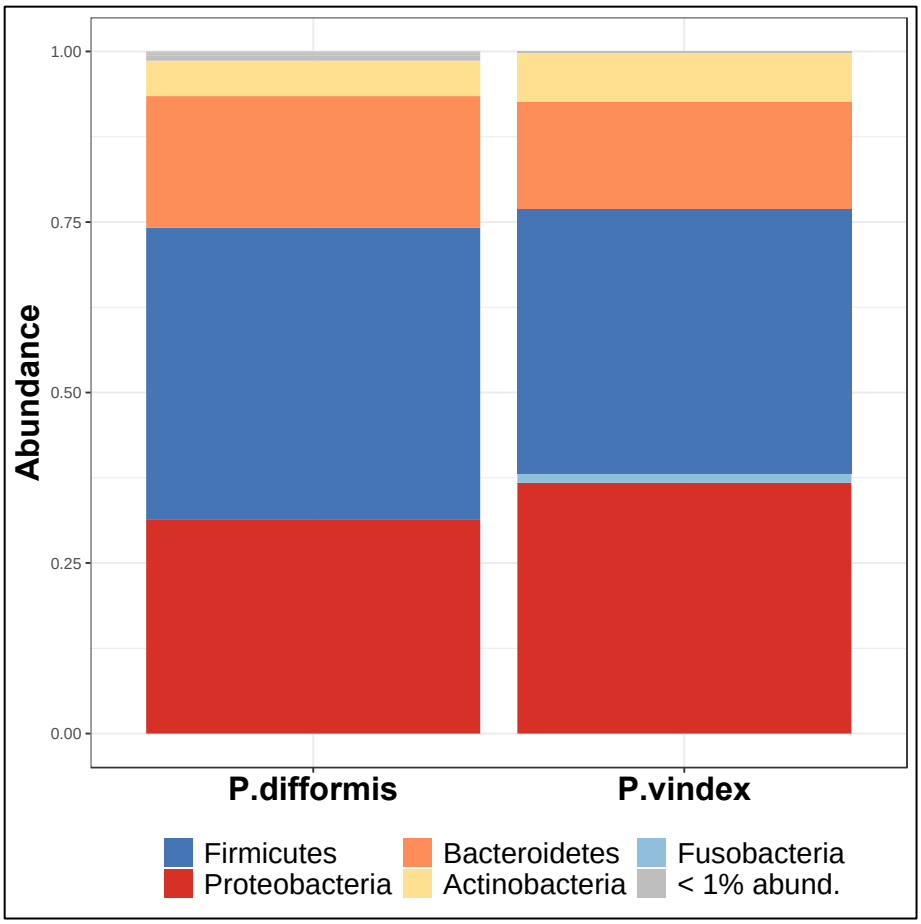


Figure 2: Relative abundance of taxonomic phyla present in the guts of *Phanaeus difformis* and *P. vindex* samples. Taxonomic groups representing less than 1% abundance have been combined.

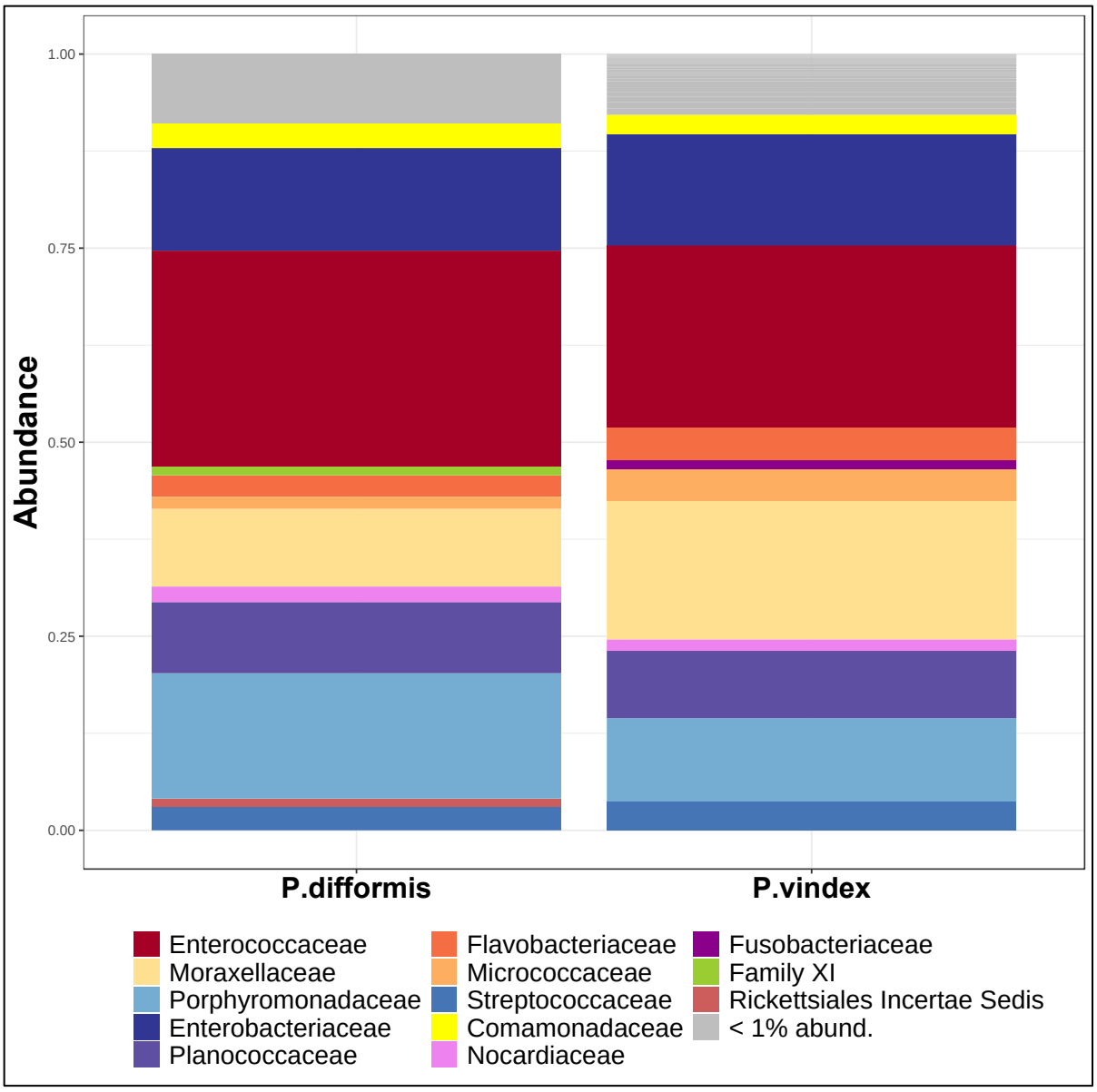
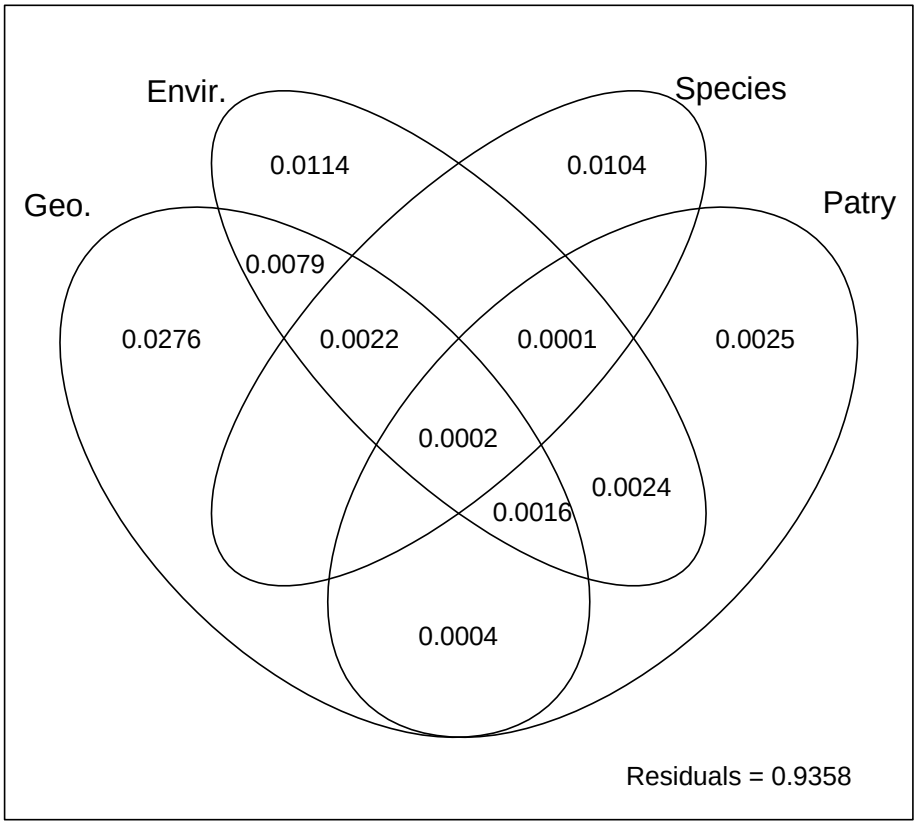
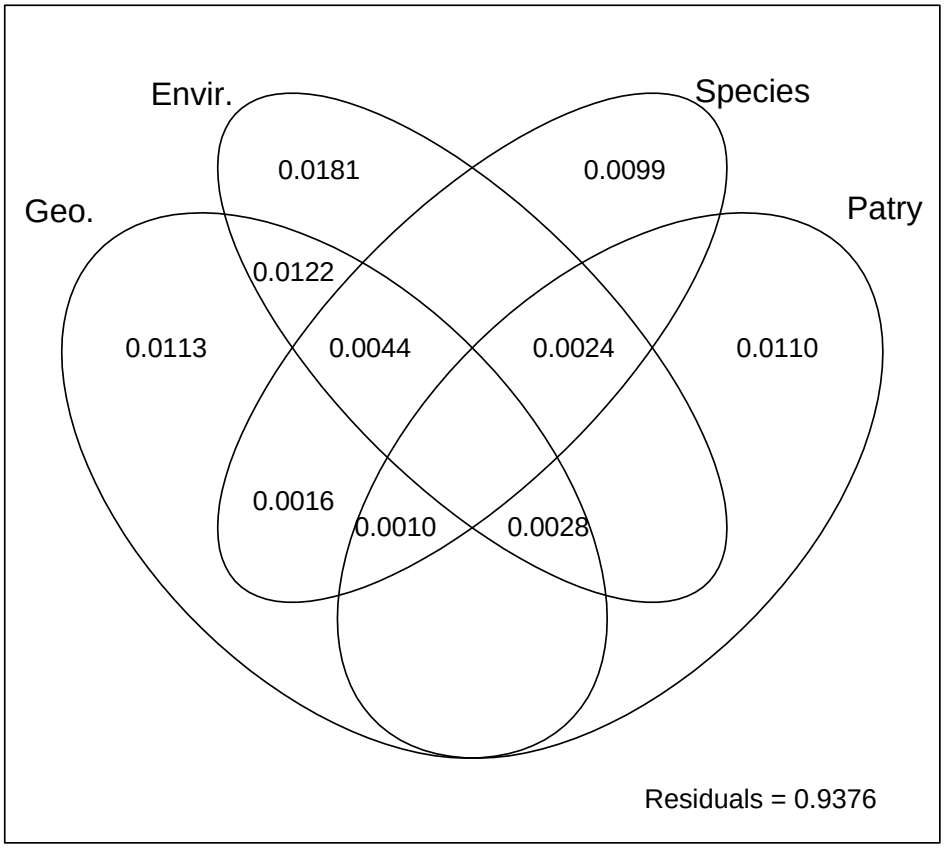


Figure 3: Relative abundance of taxonomic families present in the guts of *Phanaeus difformis* and *P. vindex* samples. Taxonomic groups representing less than 1% abundance have been combined.



(A) Quantitative Jaccard

Figure 4: Variation partitioning of *Phanaeus vindex* and *P. difformis* gut microbiome data into various components. Response variables are (A) quantitative Jaccard dissimilarity and (B) weighted UniFrac distances**.** In both plots, Species refers to the amount of variation explained by *Phanaeus* species (*P. vindex* or *P.* difformis) and Patry refers to range overlap (sympatry or allopatry of *Phanaeus* species). In A, geographic variables (Geo.) include X,Y coordinates (longitude and latitude) and six negative distance-based Moran's eigenvector mapping variables (MEMs 2,5,8,10,11, and 12); environmental variables (Envir.) include average temperature, average precipitation, the average Shannon index of soil samples taken from the sampling area, and the amount of cattle present. In B), Geo includes X,Y coordinates and one negative MEM (MEM 6); and Envir. variables include average temperature, average precipitation, and amount of cattle present. Numbers within the Venn diagram represent the adjusted R^2 of each fraction. Blank fractions have very small negative adjusted R^2 values and are not shown.



(B) Weighted UniFrac

Figure 4 (continued)

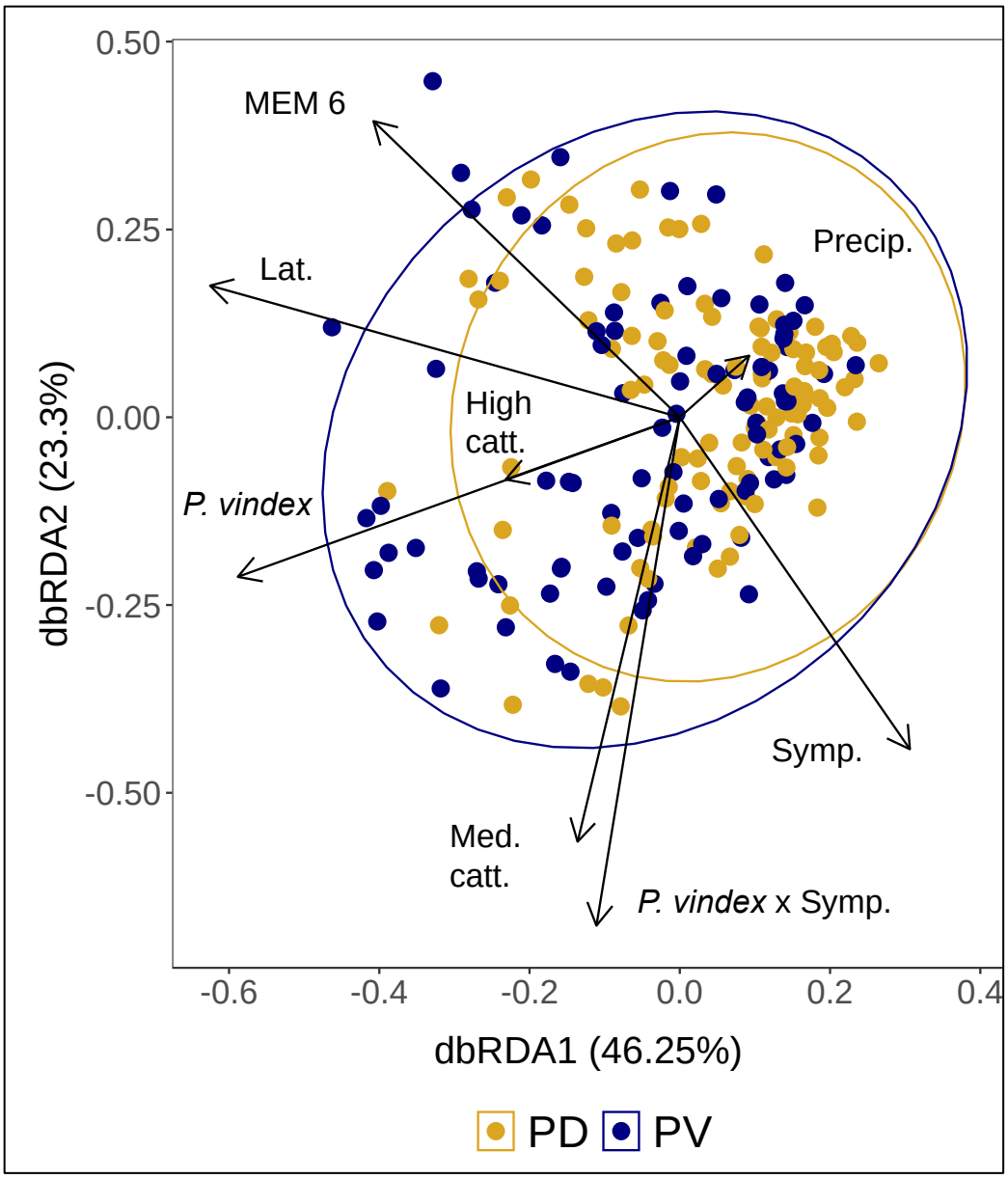


Figure 5: Plot of distance-based redundancy analysis (db-RDA) based on weighted UniFrac dissimilarity matrix of *Phanaeus difformis* and *P. vindex* samples. Points on the plot indicate individual *Phanaeus* beetles. Ellipses represent 95% confidence intervals around centroids of each *Phanaeus* species. All constrained axes together explain 10.00% of the total variation in multivariate space. Arrows indicate the strength of correlation of variables with dbRDA axes 1 and 2. Numerical variables constraining the ordination include one negative distance-based Moran's eigenvector mapping eigenvalues (MEM 6), latitude (Lat.), and precipitation of the sampling location (Precip.). Constraining factor variables include cattle presence in the sampling area (low, medium, or high), *Phanaeus* species, range overlap (i.e. sympatry or allopatry) and the interaction between *Phanaeus* species and range overlap (*P. vindex* x Symp.). Factor variables are automatically dummy coded; thus, the factor level coded as the intercept does not have a corresponding axis. A permutational anova (99,999 permutations) showed that geographic predictors (i.e. MEM 6 and latitude), precipitation, and the abundance of cattle at the sampling location were significant predictors in the model.

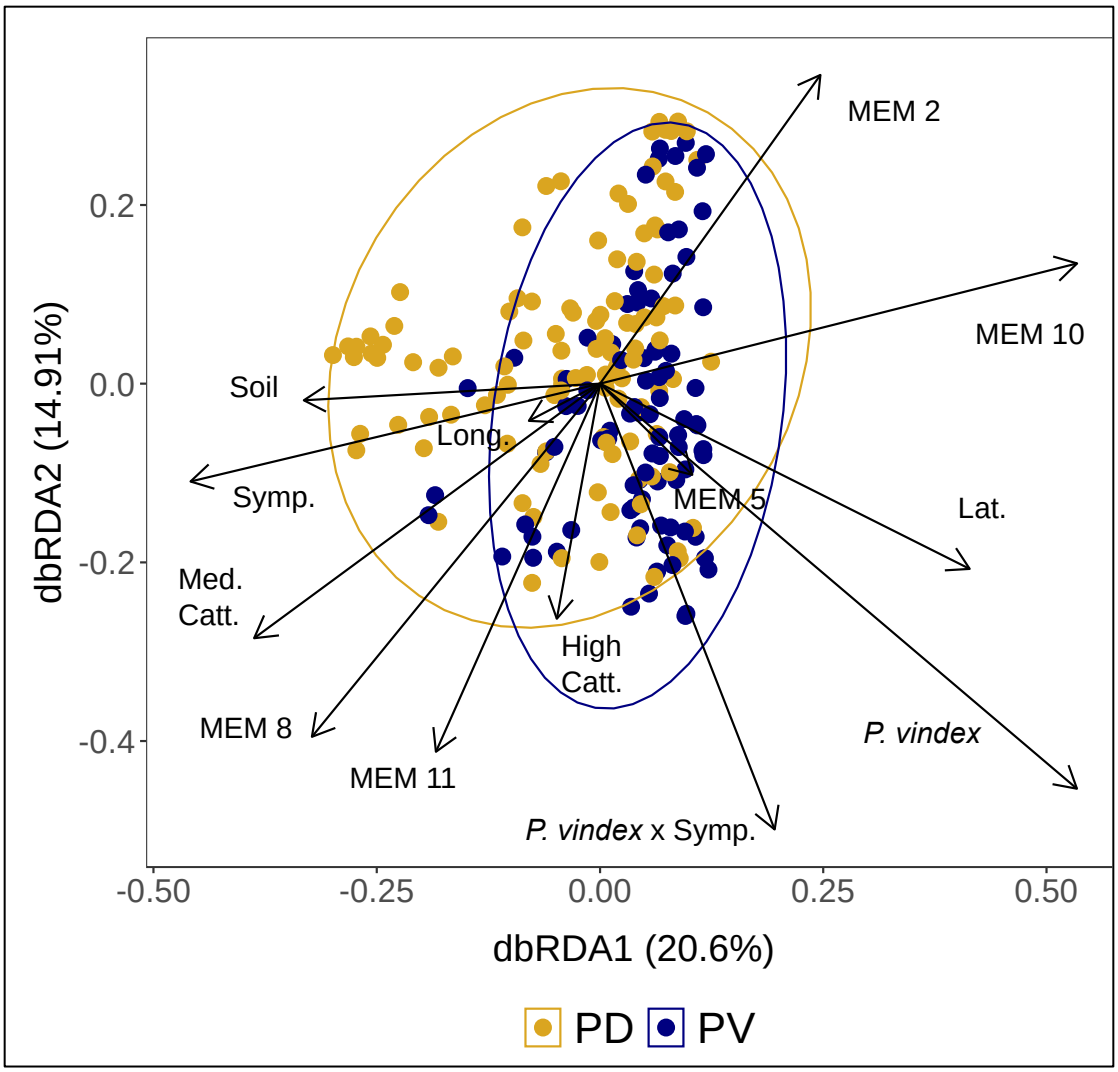
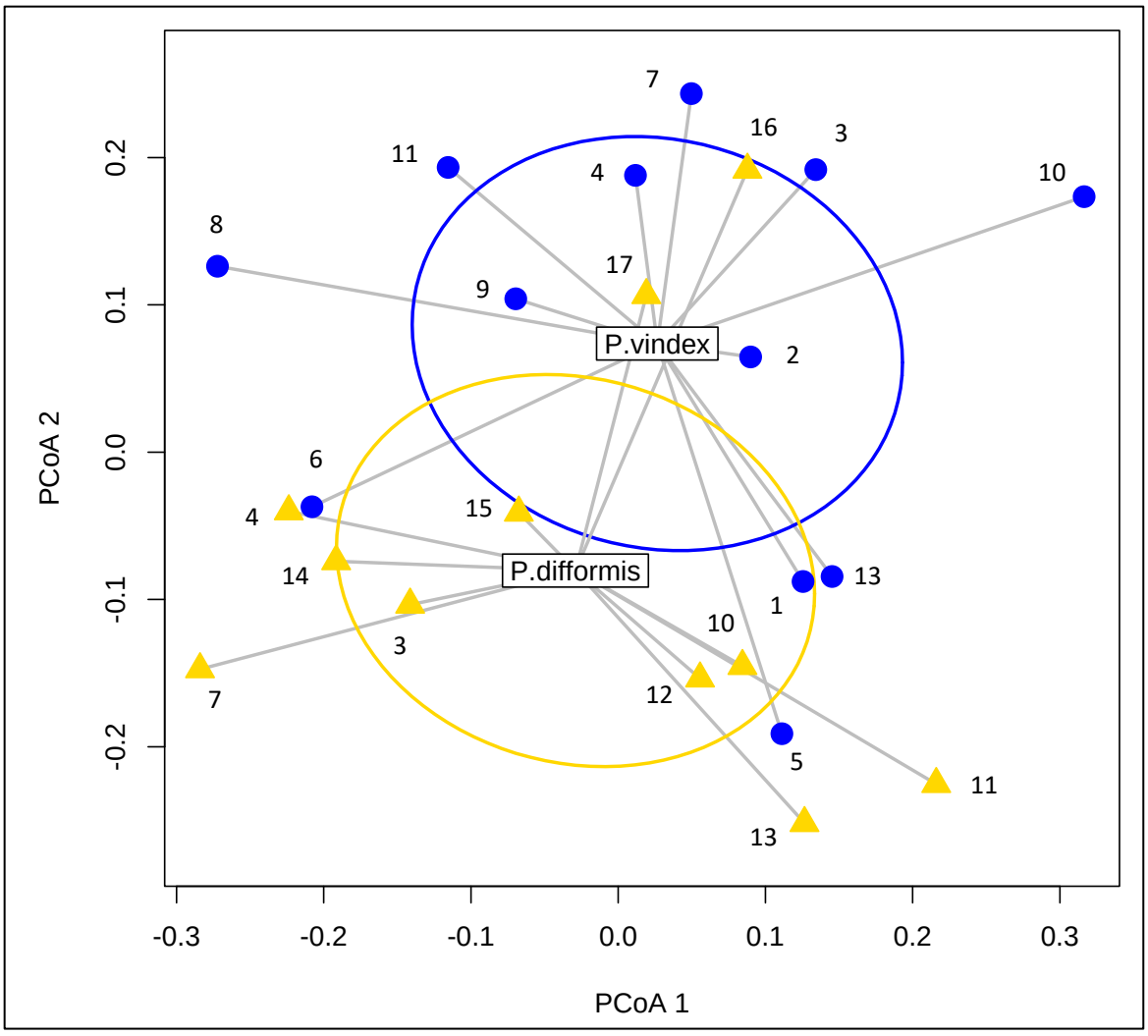
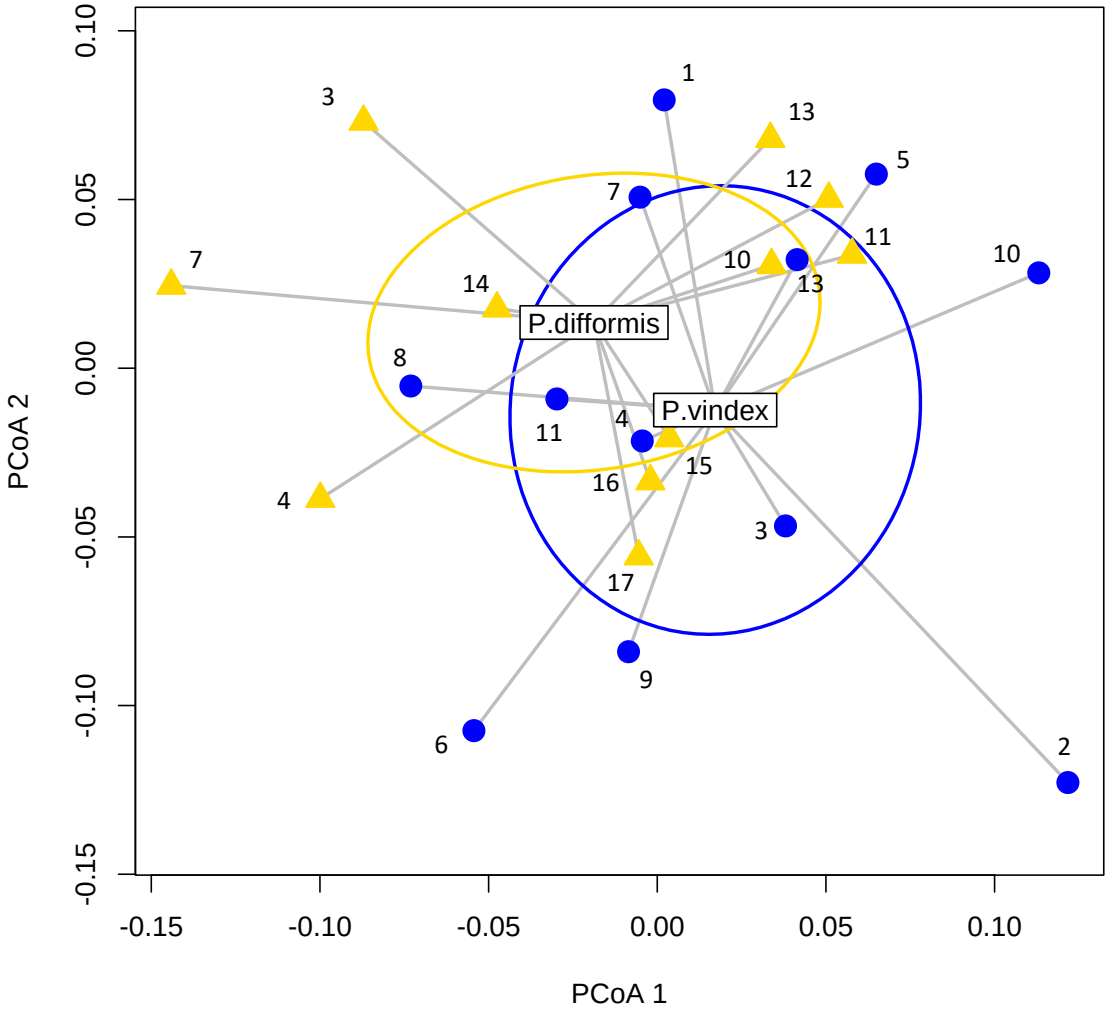


Figure 6: Plot of distance-based redundancy analysis (db-RDA) based on the quantitative Jaccard dissimilarity matrix of *Phanaeus difformis* and *P. vindex* samples. Points on the plot indicate individual *Phanaeus* beetles. Ellipses represent 95% confidence intervals around centroids of each *Phanaeus* species. All constrained axes together explain 12.27% of the total variation in multivariate space. Arrows indicate the strength of correlation of variables with dbRDA axes 1 and 2. Numerical variables constraining the ordination include five negative distance-based Moran's eigenvector mapping eigenvalues (MEMs 2, 5, 8, 10, and 11), latitude (Lat.), longitude (Long), and the average Shannon index of soil samples taken from each sampling location (Soil). Constraining factor variables include cattle presence in the sampling area (low, medium, or high), *Phanaeus* species, range overlap (i.e. sympatry or allopatry) and the interaction between *Phanaeus* species and range overlap (*P. vindex* x Symp). Factor variables are automatically dummy coded; thus, the factor level coded as the intercept does not have a corresponding axis. A permutational anova (99,999 permutations) indicated that geographic variables (i.e. dbMEMs, latitude, and longitude), cattle abundance, and the interaction effect between *P. vindex* and *P. difformis* were significant in the model.



(A) Quantitative Jaccard

Figure 7: Beta diversity of *Phanaeus vindex* and *P. difformis* represented as multivariate dispersions. (A) was performed on Jaccard dissimilarity dissimilarities and (B) was performed on weighted UniFrac dissimilarities. Populations of *P. vindex* are represented by blue dots; gold triangles indicate populations of *P. difformis*. Numbers represent different populations by location (see Table 1 for metadata associated with each population). The words "*P. vindex*" and "*P. difformis*" are positioned in the species' respective centroids in multivariate space. Ellipses represent one standard deviation away from the centroid of *P. vindex* and *P. difformis*.



(B) Weighted UniFrac

Figure 7 (continued)

Appendix B: Supplementary Data and Figures

Table S1. Metadata associated with each sampling location. PV and PD refer to *Phanaeus vindex* and *P. difformis*, respectively. Allo. and symp. stand for allopatry and sympatry, respectively. Latitude and longitude reported here are to the nearest tenth of a degree to respect the privacy of landowners. Average beetle mass refers to that of the beetle samples retained after rarefying. Average temperature and temperature variation are based on average monthly data over ten years, whereas average precipitation is the average monthly precipitation over ten years. We used the United States Department of Agriculture's (USDA) Web Soil Survey to assign the soil at each successful trap to one of the twelve major soil textural classes based on proportions of silt, sand, and clay as defined by the USDA (Soil Science Division Staff, 2017). Soil types reported here are the two most common soil types at each location (some locations had up to 3 soil types).

Location	Range overlap	Beetles		Avg. Beetle Mass (g)		Lat.	Long.	Cattle	Avg. Temp. (°C)	Temp. variation (°C)	Avg. precip. (mm)	Soil Type 1	Soil Type 2	Avg. soil Shan-non
		PV	PD	PV	PD									
1. KP	Allo. PV	3	—	0.88	—	39.1	-96.6	high	13.14	9.89	70.32	silty clay loam	silt loam	5.2339
2. DOK	Allo. PV	6	—	0.84	—	36.2	-97.7	high	15.58	9.74	65.16	sandy loam	silty clay loam	6.2502
3. CTW	Symp.	7	8	0.87	1.09	34.0	-97.4	medium	16.22	9.20	81.18	loamy sand	sandy loam	6.6027
4. FBR	Symp.	9	12	0.95	1.00	33.9	-96.9	high	16.92	9.28	88.49	loamy sand	loamy sand	6.5900
5. PMW	Allo. PV	7	—	0.66	—	33.8	-95.7	low	18.47	9.21	95.05	silt loam	sandy loam	6.1704
6. PME	Allo. PV	5	—	0.91	—	33.8	-95.6	low	18.47	9.21	95.05	sandy loam	sandy loam	6.0668
7. FW	Symp.	11	11	0.88	1.14	32.8	-97.5	low	19.37	8.17	80.85	sandy loam	loamy sand	6.4095
8. WTX	Allo. PV	6	—	1.17	—	32.5	-96.0	high	18.65	8.12	93.91	sandy loam	sandy loam	5.7322
9. RCW	Allo. PV	10	—	0.81	—	31.9	-96.1	low	18.87	9.44	82.02	sandy loam	clay loam	6.2151
10. GEW	Symp.	11	12	1.04	1.26	31.9	-95.9	low	19.58	7.35	89.40	loamy sand	loamy sand	6.4127
11. BUTX	Symp.	8	12	0.90	0.97	31.5	-96.1	high	19.08	7.47	88.92	sand	sand	6.6602
12. STR	Allo. PD	—	12	—	0.82	30.8	-99.4	med.	19.06	7.52	55.57	sandy loam	sandy loam	6.1468
13. BATX	Symp.	6	12	0.78	1.23	30.1	-97.4	high	19.908	8.73	80.57	loamy sand	loamy sand	6.2731
14. GCR	Allo. PD	—	6	—	0.96	29.5	-97.8	high	18.959	10.06	70.98	sand	sand	6.2240
15. NW	Allo. PD	—	9	—	1.23	29.6	-97.7	medium	18.959	10.06	70.98	sand	sand	6.0804
16. YTX	Allo. PD	—	9	—	0.99	29.0	-97.5	high	21.326	7.73	68.47	loamy sand	loamy sand	5.6862
17. CWE	Allo. PD	—	7	—	1.46	28.3	-99.4	medium	22.718	7.61	42.03	sandy loam	sandy loam	5.8813

Table S2. Results of forward model selection based on adjusted R^2 and P-values, performed prior to variation partitioning. Jaccard dissimilarities and weighted UniFrac dissimilarities were used as response variables. **(**A)Model selection on negative distance-based Moran's eigenvector map (dbMEM) variables. Because a linear trend was detected, detrended dissimilarity matrices were used as the response variables. B) Model selection on environmental variables.

(A) Model selection on negative distance-based Moran's eigenvector map variables

		After double-stopping forward model selection		After Sidák correction
Dissimilarity matrix	**MEM:**	**adjusted R^2**	**p-value**	**adjusted p-value**
Jaccard	MEM1	0.036909	< 0.05 *	0.19132
	MEM2	0.016477	< 0.01 **	< 0.05 *
	MEM5	0.024535	< 0.01 **	< 0.05 *
	MEM6	0.03904	< 0.05 *	0.29557
	MEM7	0.034445	< 0.01 **	0.09164
	MEM8	0.028196	< 0.01 **	< 0.05 *
	MEM9	0.031439	< 0.01 **	0.06198
	MEM10	0.006983	< 0.0001 ***	< 0.001 ***
	MEM11	0.012124	< 0.001 ***	< 0.01 **
	MEM12	0.020868	< 0.01 **	< 0.05 *
Weighted UniFrac	MEM6	0.010804	< 0.01 **	< 0.05 *
	MEM7	0.019963	< 0.05 *	0.06797
	MEM8	0.033577	< 0.05 *	0.11256
	MEM11	0.027603	< 0.05 *	0.19922

(B). Model selection on environmental variables

Dissimilarity matrix	Environmental variable	After double-stopping forward model selection	
		Adjusted R^2	P-value
Jaccard	Cattle presence	0.010721	0.0001 ***
	Average soil Shannon entropy	0.015704	< 0.001 ***
	Average monthly temperature	0.020614	< 0.001 ***
	Average monthly precipitation	0.025039	< 0.01 **
	All variables	0.027232	N/A
weighted UniFrac	Cattle presence	0.020423	< 0.05 *
	Average monthly temperature	0.012315	< 0.01 **
	Average monthly precipitation	0.033710	< 0.01 **
	All variables	0.038557	N/A

Table S3. Results of forward model selection based on adjusted R^2 and P-value, performed prior to implementing distance-based redundancy analyses (db-RDAs) and additional variation partitioning. Quantitative Jaccard and weighted UniFrac distances were used as response variables**.** All variables except precipitation in the Jaccard model were included in final db-RDAs and second round of variation partitioning analyses.

		After double-stopping forward model selection	
Dissimilarity matrix	**Predictor**	**Adjusted R^2**	**P-value**
Jaccard	X coordinates (longitude)	0.062416	< 0.01 **
	Y coordinates (latitude)	0.037077	< 0.001 ***
	Negative dbMEM 2	0.053218	< 0.05 *
	Negative dbMEM 5	0.050680	< 0.01 **
	Negative dbMEM 8	0.057153	< 0.01 **
	Negative dbMEM 10	0.046759	< 0.01 **
	Negative dbMEM 11	0.032073	< 0.001 ***
	Ranked cattle abundance	0.010721	< 0.001 ***
	Average precipitation	0.042212	< 0.001 ***
	Average Shannon entropy of soil	0.058946	< 0.05 *
	Phanaeus species	0.019942	< 0.001 ***
	Range overlap (“patry”)	0.026400	< 0.001 ***
	All variables	0.064215	N/A
Weighted UniFrac	Y coordinates (latitude)	0.016370	< 0.001 ***
	Negative dbMEM 6	0.060728	< 0.05 *
	Ranked cattle abundance	0.037549	< 0.05 *
	Average precipitation	0.053021	< 0.05 *
	Phanaeus species	0.044704	< 0.05 *
	Range overlap (“patry”)	0.028341	< 0.01 **
	All variables	0.064637	N/A

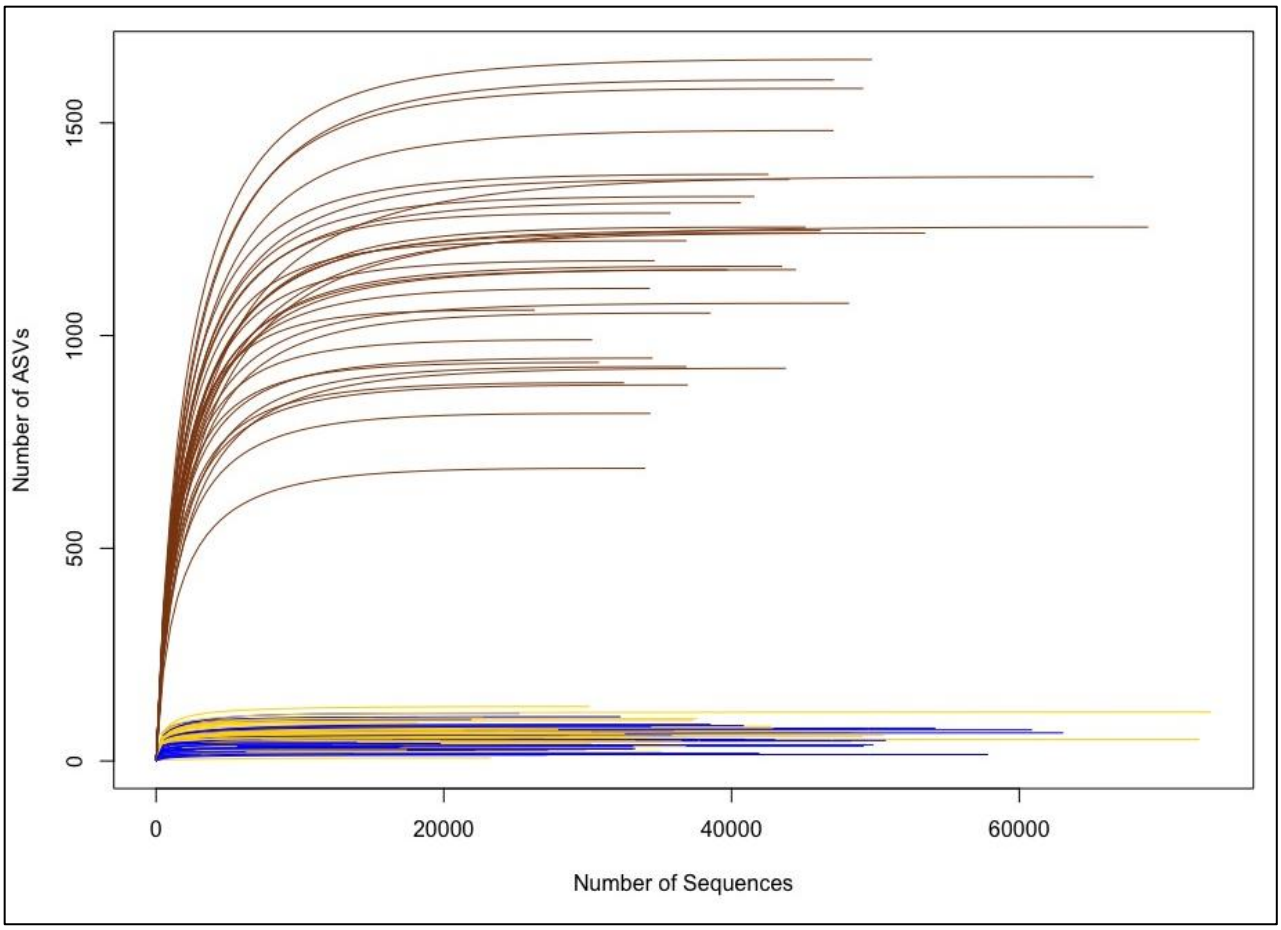


Figure S4. Rarefaction curves indicate that rarefaction depths chosen were adequate to capture ASV diversity in our samples. Brown lines represent soil samples, whereas blue and gold lines are *P. vindex* and *P. difformis* samples, respectively. Soil samples were rarefied at 26,336 reads and gut samples were rarefied at 3,500 reads because soils were far more diverse than gut samples and overall had better sampling coverage.

Table S5. Indicator species analyses for different groupings of *Phanaeus difformis* and *P. vindex* gut microbiome samples in allopatry and in sympatry. Asterisks next to r_g values indicate significance levels; * p < 0.05, ** p < 0.01, and *** p < 0.001. PD and PV stand for *P. difformis* and *P. vindex*, respectively. No species characterizing the combination of sympatric *P. vindex* and sympatric *P. difformis* gut microbiome samples were identified.

Group	Phylum, family, genus, species	r_g	Percent Relative Abundance			
			Allo. PD	Symp. PD	Allo. PV	Symp. PV
Allopatric *P. difformis*	Proteobacteria, Orbaceae, *Gilliamella,* Unidentified species	0.279 ***	0.697	0.161	0	0.030
	Firmicutes, Planococcaceae, *Solibacillus*, unidentified species	0.278 ***	0.688	0.120	0.054	0.030
	Firmicutes, Planococcaceae, unidentified genus, unidentified species	0.267 **	2.349	0.367	0.341	0.679
	Proteobacteria, Rhodobacteraceae, unidentified genus, unidentified species	0.258 **	0.014	0.001	0	0
	Tenericutes, Spiroplasmataceae, *Spiroplasma*, unidentified species	0.249 **	0.006	0.001	0.001	0.001
	Proteobacteria, Comamonadaceae, *Comamonas*, unidentified species	0.243 **	2.041	0.704	0.011	0.451
	Bacteroidetes, Porphyromonadaceae, *Proteiniphilum*, unidentified species	0.223 *	0.032	0	0	0
	Proteobacteria, Neisseriaceae, unidentified genus, unidentified species	0.219 **	0.098	0.005	0	0
	Firmicutes, Enterococcaceae, *Enterococcus*, unidentified species	0.212 *	0.122	0.028	0	0
	Actinobacteria, Micrococcales Incertae Sedis, *Timonella,* uncultured bacterium	0.208 *	0.543	0.262	0.059	0.245
	Proteobacteria, Enterobacteriaceae, unidentified genus, unidentified species	0.192 *	1.218	0	0	0
	Bacteroidetes, Sphingobacteriaceae, *Sphingobacterium,* uncultured bacterium	0.189 *	0.003	0.001	0	0
	Actinobacteria, Segniliparaceae, *Segniliparus*, Unidentified species	0.187 *	0.152	0.033	0.069	0.044
	Bacteroidetes, Porphyromonadaceae, *Dysgonomonas*, uncultured bacterium	0.18 *	1.194	0	0	0
	Firmicutes, Enterococcaceae, *Enterococcus*, unidentified species	0.18 *	0.013	0	0	0
	Tenericutes, Mycoplasmataceae, *Echinogammarus veneris*	0.178 *	0.646	0.023	0.041	0.032
	Proteobacteria, Enterobacteriaceae, *Morganella,* unidentified species	0.17 *	0.101	0	0	0
	Proteobacteria, Rhodobacteraceae, unidentified genus, unidentified species	0.167 *	0.039	0.002	0	0
	Saccharibacteria, unidentified family, unidentified genus, unidentified species	0.163 **	0.091	0	0	0
	Spirochaetae, Brevinemataceae, *Brevinema*, unidentified species	0.163 *	0.063	0	0	0
	Proteobacteria, Moraxellaceae, *Acinetobacter*, unidentified species	0.148 *	0.142	0	0.002	0
Sympatric *P. difformis*	Firmicutes, Enterococcaceae, *Vagococcus*, unidentified species	0.363 ***	4.360	13.304	0.135	2.401
	Proteobacteria, Neisseriaceae, *Vitreoscilla*, Unidentified species	0.295 ***	0.060	0.963	0.156	0.289
	Bacteroidetes, Porphyromonadaceae, *Dysgonomonas*, uncultured bacterium	0.244 **	0.308	1.590	0.092	0.499
	Firmicutes, Enterococcaceae, *Vagococcus*, unidentified species	0.202 *	1.767	3.459	0.743	1.702
	Firmicutes, Family XI, *Gallicola*, uncultured bacterium	0.199 *	0.017	1.145	0.107	0.376
	Bacteroidetes, Porphyromonadaceae, *Dysgonomonas*, uncultured bacterium	0.199 *	0.001	0.055	0	0.013

Table S5 (Continued)

Group	Phylum, family, genus, species	r_g	Percent Relative Abundance			
			Allo. PD	Symp. PD	Allo. PV	Symp. PV
Sympatric *P. difformis*	Proteobacteria, Rhodobacteraceae, uncultured, bacterium endosymbiont of *Onthophagus taurus*	0.196 *	0.002	0.013	0	0
	Firmicutes, Ruminococcaceae, uncultured, uncultured Firmicutes bacterium	0.196 *	0	0.013	0	0
	Proteobacteria, Orbaceae, *Gilliamella*, Unidentified species	0.192 *	0	0.001	0	0
	Proteobacteria, Coxiellaceae, *Rickettsiella*, Candidatus Rickettsiella viridis	0.19 *	0	0.003	0	0
	Firmicutes, Enterococcaceae, *Vagococcus,* unidentified species	0.179 *	0	0.061	0	0
Allopatric *P. vindex*	Proteobacteria, Enterobacteriaceae, unidentified genus, unidentified species	0.318 ***	1.242	0.923	4.854	0.902
	Proteobacteria, Moraxellaceae, *Acinetobacter*, unidentified species	0.278 ***	0.029	0.117	7.161	0.648
	Proteobacteria, Moraxellaceae, *Acinetobacter*, unidentified species	0.248 **	0.002	0.000	0.156	0.000
	Proteobacteria, Pseudomo nadaceae, *Pseudomonas*, unidentified species	0.239 **	0.069	0.090	0.211	0.086
	Proteobacteria, Comamonadaceae, *Delftia,* unidentified species	0.211 *	0.003	0.004	0.034	0.010
	Actinobacteria, Micrococcaceae, unidentified genus, unidenti fied species	0.207 **	0.933	0.529	3.197	1.418
	Proteobacteria, Burkholderiaceae, *Ralstonia*, unidentified species	0.202 *	0	0	0.003	0
	Proteobacteria, Enterobacteriaceae, *Hafnia-Obesumbacterium*, unidentified species	0.202 *	0.023	0	2.471	0.015
	Actinobacteria, Nocardiaceae, *Rhodococcus*, unidentified species	0.197 *	0	0.001	0.013	0.003
	Proteobacteria, Methylobacteriaceae, *Methylobacterium*, *Methylobacterium aquaticum*	0.195 *	0.022	0.023	0.073	0.039
	Actinobacteria, Nocardioidaceae, *Nocardioides*, unidentified species	0.194 **	0	0	0.092	0
	Actinobacteria, Micrococcaceae, unidentified genus, unidentified species	0.189 *	0	0.001	0.101	0
	Bacteroidetes, Flavobacteriaceae, *Empedobacter,* unidentified species	0.187 *	0	0	0.011	0
	Bacteroidetes, Flavobacteriaceae, *Flavobacterium*, unidentified species	0.186 *	0.006	0	0.083	0
	Firmicutes, Lachnospiraceae, *Lachnoclostridium 5,*Unidentified species	0.184 *	0	0	0.016	0
	Firmicutes, Enterococcaceae, *Vagococcus*, unidentified species	0.183 *	0	0	0.036	0
	Proteobacteria, Enterobacteriaceae, *Morganella*, unidentified species	0.169 *	0.005	0.022	0.222	0
	Unassigned phylum, unassigned genus, unassigned species	0.163*	0	0	0.004	0
	Verrucomicrobia, DA101 soil group, unidentified genus, unidentified species	0.154 *	0	0	0.003	0
	Proteobacteria, Comamonadaceae, unidentified genus, unidentified species	0.153 *	0	0	0.049	0
Sympatric *P. vindex*	Actinobacteria, Micrococcaceae, *Glutamicibacter,* Unidentified species	0.311 ***	0.035	0.466	0.414	1.221
	Actinobacteria, Microbacteriaceae, *Leucobacter,* unidentified species	0.273 **	0	0.002	0	0.015

Table S5 (Continued)

Group	Phylum, family, genus, species	r_g	Percent Relative Abundance			
			Allo. PD	Symp. PD	Allo. PV	Symp. PV
Sympatric *P. vindex*	Firmicutes, Enterococcaceae, *Enterococcus*, unidentified species	0.259 **	0	0.053	2.927	9.900
	Proteobacteria, Moraxellaceae, *Acinetobacter,* unidentified species	0.243 **	0.003	0.006	0.012	0.027
	Actinobacteria, Promicromonosporaceae, *Cellulosimicrobium*, unidentified species	0.237 **	0	0	0	0.006
	Actinobacteria, Micrococcaceae, *Glutamicibacter,* unidentified species	0.23 **	0.363	0.368	0.212	0.751
	Bacteroidetes, Sphingobacteriaceae, *Sphingobacterium*, unidentified species	0.224 **	0	0	0	0.042
	Bacteroidetes, Flavobacteriaceae, *Empedobacter*, uncultured bacterium	0.202 *	0	0.084	0	0.451
	Firmicutes, Veillonellaceae, uncultured, uncultured bacterium	0.199 *	0	0.003	0	0.026
	Fusobacteria, Fusobacteriaceae, *Fusobacterium*, unidentified species	0.194 *	0	0.025	0	0.264
	Actinobacteria, Micrococcaceae, unidentified genus, unidentified species	0.187 *	0	0.013	0	0.178
	Actinobacteria, Dermacoccaceae, unidentified genus, unidentified species	0.184 *	0	0.012	0.011	0.054
	Proteobacteria, Comamonadaceae, *Comamonas*, unidentified genus	0.174 *	0.021	0.101	0.024	0.365
	Bacteroidetes, Sphingobacteriaceae, *Sphingobacterium*, unidentified genus	0.17 *	0.108	0.110	0.041	0.637
	Bacteroidetes, Porphyromonadaceae, *Dysgonomonas*, Unidentified species	0.169 *	0	0	0	0.019
Allopatric *P. difformis* + Sympatric *P. difformis*	Actinobacteria, Propionibacteriaceae, *Propioniciclava*, uncultured bacterium	0.22 *	0.189	0.183	0.070	0.049
	Bacteroidetes, Flavobacteriaceae, *Cloacibacterium*, Unidentified species	0.184 *	0.011	0.007	0	0
	Firmicutes, Enterococcaceae, *Enterococcus*, unidentified species	0.173 *	0.108	0.079	0	0
Allopatric *P. difformis* + Allopatric *P. vindex*	Firmicutes, Planococcaceae, *Kurthia,* unidentified species	0.187 *	9.286	3.631	9.118	4.997
Allopatric *P. vindex* + Sympatric *P. vindex*	Proteobacteria, Enterobacteriaceae, unidentified genus, unidentified species	0.228 **	0.491	1.087	2.329	3.343
	Actinobacteria, Micrococcaceae, unidentified genus, unidentified species	0.189 *	0.094	0.056	0.359	0.231
Allopatric *P. difformis* + Sympatric *P. vindex*	Proteobacteria, Rhodobacteraceae, unidentified genus, unidentified species	0.234 **	0.093	0.027	0.005	0.175
	Proteobacteria, Moraxellaceae, Acinetobacter, unidentified species	0.2 *	1.108	0.192	0.028	0.980

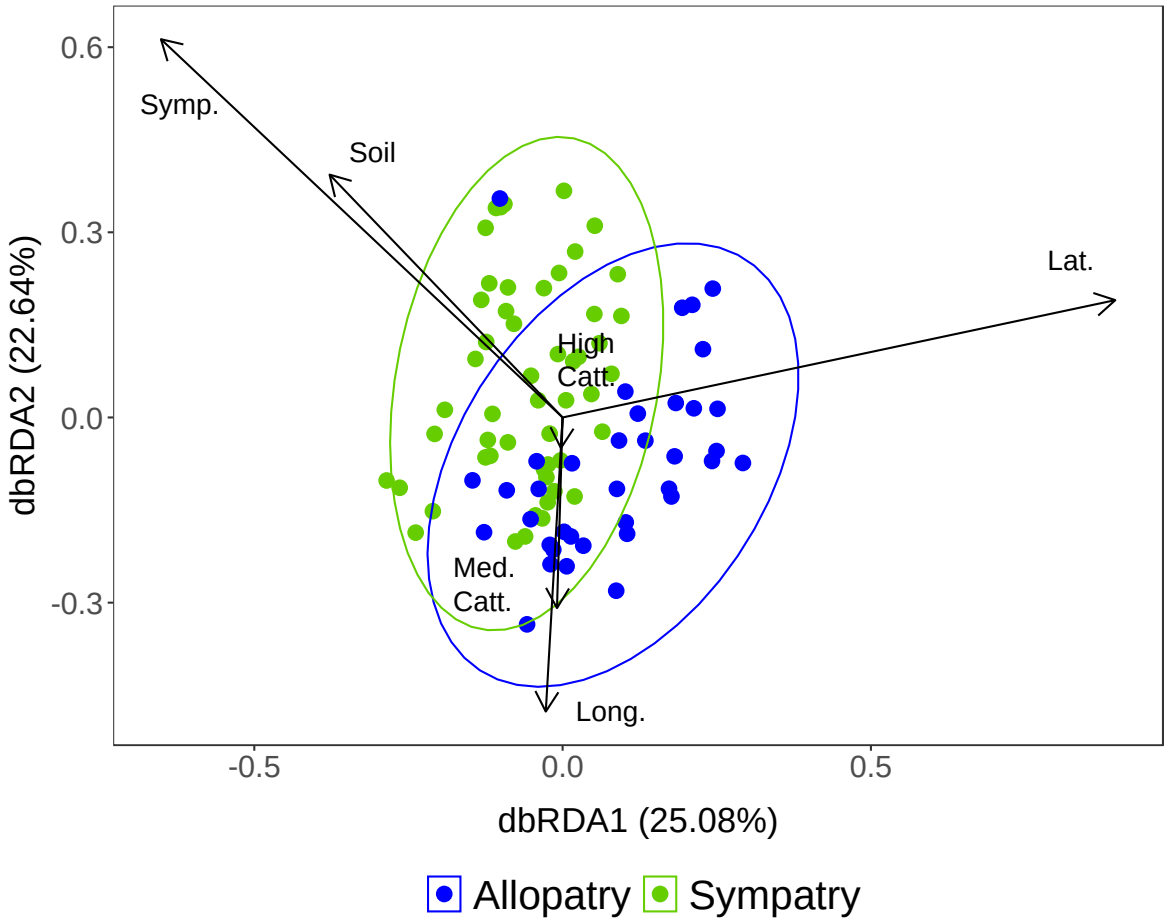


Figure S6: Plot of distance-based redundancy analysis (db-RDA) based on the quantitative Jaccard dissimilarity among *Phanaeus vindex* samples in sympatry and in allopatry. Points on the plot indicate individual *P. vindex* beetles. Ellipses represent 95% confidence intervals around centroids of *P. vindex* in allopatry and in sympatry with *P. difformis*. All constrained axes together explain 10.89% of the total variation in multivariate space. Arrows indicate the strength of correlation of variables with dbRDA axes 1 and 2. Numerical variables constraining the ordination latitude (Lat), longitude (Long.), and average Shannon index of soil samples taken from the sampling areas (Soil). Constraining factor variables include cattle presence in the sampling area (low, medium, or high), and range overlap (i.e. sympatry or allopatry). Factor variables are automatically dummy coded; thus, the factor level coded as the intercept does not have a corresponding axis. All variables shown were significant following anovas (99,999 permutations).

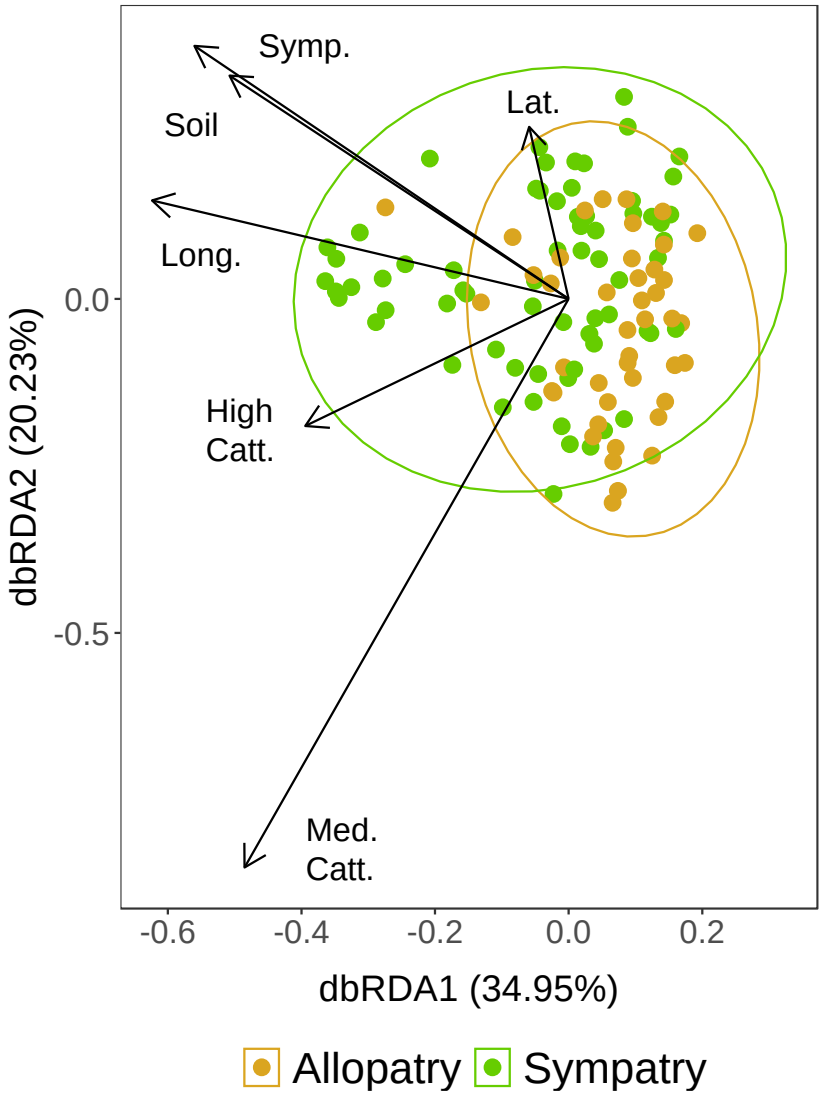


Figure S7: Plot of distance-based redundancy analysis (db-RDA) based on the quantitative Jaccard dissimilarity among *Phanaeus difformis* samples in sympatry and in allopatry. Points on the plot indicate individual *P. difformis* beetles. Ellipses represent 95% confidence intervals around centroids of *P. difformis* in allopatry and in sympatry with *P. vindex*. All constrained axes together explain 9.47% of the total variation in multivariate space. Arrows indicate the strength of correlation of variables with dbRDA axes 1 and 2. Numerical variables constraining the ordination latitude (Lat), longitude (Long.), and average Shannon index of soil samples taken from the sampling areas (Soil). Constraining factor variables include cattle presence in the sampling area (low, medium, or high), and range overlap (i.e. sympatry or allopatry). Factor variables are automatically dummy coded; thus, the factor level coded as the intercept does not have a corresponding axis. All variables shown were significant following anovas (99,999 permutations).

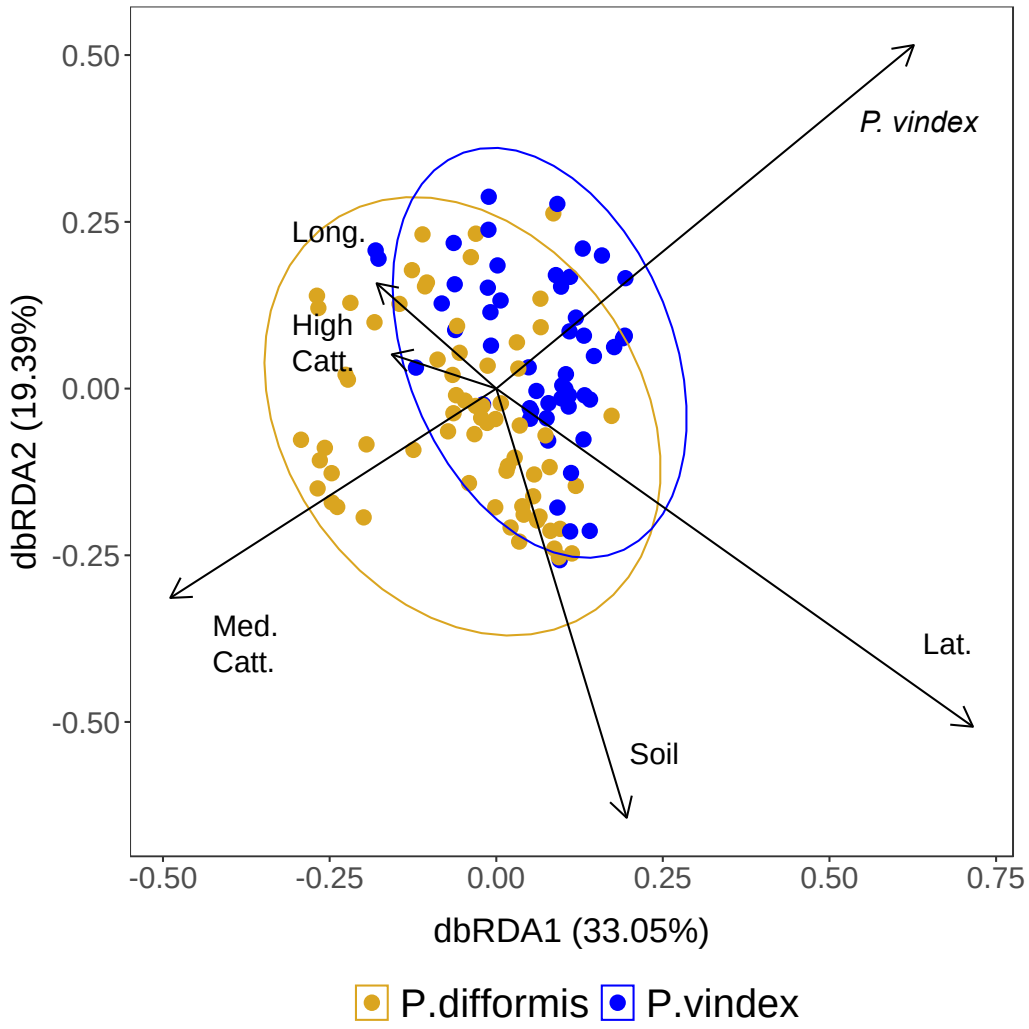


Figure S8: Plot of distance-based redundancy analysis (db-RDA) based on the quantitative Jaccard dissimilarity among samples of *Phanaeus vindex* and *P. difformis* in sympatry. Points on the plot indicate individual *P. difformis* and *P. vindex* beetles. Ellipses represent 95% confidence intervals around centroids of *P. difformis* in allopatry and in sympatry with *P. vindex*. All constrained axes together explain 11.24% of the total variation in multivariate space. Arrows indicate the strength of correlation of variables with dbRDA axes 1 and 2. Numerical variables constraining the ordination latitude (Lat), longitude (Long.), and average Shannon index of soil samples taken from the sampling areas (Soil). Constraining factor variables include cattle presence in the sampling area (low, medium, or high), and range overlap (i.e. sympatry or allopatry). Factor variables are automatically dummy coded; thus, the factor level coded as the intercept does not have a corresponding axis. All variables shown were significant following anovas (99,999 permutations).

VITA

Claire Winfrey grew up in Oklahoma City, Oklahoma. She attended the University of Oklahoma in Norman (OU) from the years of 2011-2016, obtaining a Bachelor of Science in Biology, *summa cum laude.* At OU, she was an active member of the labs of Dr. Ola Fincke and Dr. Cameron Siler. Under the mentorship of Dr. Fincke, she designed and performed experiments to investigate the mating cues of damselflies which resulted in her first peer-reviewed publication. In the lab of Dr. Siler, Claire helped describe a new species of skink and worked on a project quantifying the prevalence of amphibian pathogens across Oklahoma. Upon completion of her MS from the University of Tennessee, Knoxville, she plans to pursue a Ph.D. in Ecology & Evolutionary Biology at the University of Colorado Boulder.