ORIGINAL RESEARCH

Mechanisms governing avian phylosymbiosis: Genetic dissimilarity based on neutral and MHC regions exhibits little relationship with gut microbiome distributions of Galápagos mockingbirds

Ramona Fleischer[1] 🆔 | Alice Risely[1] 🆔 | Paquita E. A. Hoeck[2] | Lukas F. Keller[2,3] | Simone Sommer[1]

[1]Institute of Evolutionary Ecology and Conservation Genomics, University of Ulm, Ulm, Germany

[2]Zoological Museum, University of Zurich, Zurich, Switzerland

[3]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

Correspondence
Ramona Fleischer and Alice Risely, Institute of Evolutionary Ecology and Conservation Genomics, University of Ulm, Ulm, Germany.
Emails: riselya@gmail.com; ramona.fleischer@uni-ulm.de

Abstract

The gut microbiome of animals, which serves important functions but can also contain potential pathogens, is to varying degrees under host genetic control. This can generate signals of phylosymbiosis, whereby gut microbiome composition matches host phylogenetic structure. However, the genetic mechanisms that generate phylosymbiosis and the scale at which they act remain unclear. Two non-mutually exclusive hypotheses are that phylosymbiosis is driven by immunogenetic regions such as the major histocompatibility complex (MHC) controlling microbial composition, or by spatial structuring of neutral host genetic diversity via founder effects, genetic drift, or isolation by distance. Alternatively, associations between microbes and host phylogeny may be generated by their spatial autocorrelation across landscapes, rather than the direct effects of host genetics. In this study, we collected MHC, microsatellite, and gut microbiome data from separate individuals belonging to the Galápagos mockingbird species complex, which consists of four allopatrically distributed species. We applied multiple regression with distance matrices and Bayesian inference to test for correlations between average genetic and microbiome similarity across nine islands for which all three levels of data were available. Clustering of individuals by species was strongest when measured with microsatellite markers and weakest for gut microbiome distributions, with intermediate clustering of MHC allele frequencies. We found that while correlations between island-averaged gut microbiome composition and both microsatellite and MHC dissimilarity existed across species, these relationships were greatly weakened when accounting for geographic distance. Overall, our study finds little support for large-scale control of gut microbiome composition by neutral or adaptive genetic regions across closely related bird phylogenies, although this does not preclude the possibility that host genetics shapes gut microbiome at the individual level.

1 | INTRODUCTION

The gut microbiome of animals is a source of both functionality and pathogens, and is to various degrees under host genetic control (Davenport, 2016; Kubinak et al., 2015; Kurilshikov et al., 2017). However, the mechanisms by which host genetic variation shapes gut microbiota composition and the scale at which they act remain unclear (Kohl, 2020; Kurilshikov et al., 2017; Lim & Bordenstein, 2020). Within host species, pairwise similarity in taxonomic microbiome composition between individuals has been found to correlate with host genetic similarity (e.g., Griffiths et al., 2018, 2019; Kohl et al., 2017; Smith et al., 2015; Suzuki et al., 2019; Webster et al., 2018), yet generally the effect of genetics on the microbiome within species is thought to be modest (Kurilshikov et al., 2017) and not always detectable (e.g., Rothschild et al., 2018; Yuan et al., 2015). The influence of host genetics becomes stronger at cross-species phylogenetic scales, with large-scale studies demonstrating considerable host phylogenetic effects on the gut microbiome, although such affects are weaker in birds than mammals (Amato et al., 2018; Song et al., 2020; Trevelline et al., 2020; Youngblut et al., 2019).

A number of non-exclusive mechanisms could underpin the influence of host genetics on gut microbiome composition. Adaptive genetic diversity (i.e., genes or loci that affect host fitness and are therefore under selection; Holderegger et al., 2006) may be disproportionately important in controlling the gut microbiota. For example, genetically encoded components of the immune and endocrine system regulate the gut microbiome (Donaldson et al., 2018; Noguera et al., 2018), and variation in these regions could have disproportionally large effects on microbiota composition compared with neutral regions. One region of particular interest is the major histocompatibility complex (MHC), a highly diverse suite of genes that are crucial to the adaptive immune system across vertebrates (Piertney & Oliver, 2006; Sommer, 2005). Variation in the MHC has been repeatedly linked to host microbiome composition (Bolnick et al., 2014; Kubinak et al., 2015; Leclaire et al., 2019; Pearce et al., 2017; Silverman et al., 2017), and this regulation by the MHC may be adaptive by protecting against autoimmunity (Silverman et al., 2017) and reducing host susceptibility to infection (Kubinak et al., 2015). Spatial variation in parasite and pathogen communities across the host species range may therefore select for population-specific compositions in MHC genotypes that have the potential to generate divergent gut microbiome communities across the host population.

Alternatively, or in addition to MHC effects, differences in neutral host genetic diversity (genes or loci that have no effect on host fitness and are therefore not under selection; Holderegger et al., 2006) may contribute to the phylogenetic and spatial structuring of gut microbiomes observed both within and across species. While studies that link neutral gene markers to the microbiome are still uncommon in wild populations, there is evidence that variation in neutral regions of the host genome within species is associated with microbial composition of gut (Smith et al., 2015), skin (Griffiths et al., 2018), and coral surfaces, even in the presence of high gene flow (Griffiths et al., 2019). As host species diverge, the accumulation of such neutral effects via founder effects, genetic drift, or isolation by distance, would together generate a distinct phylogenetic signal on the host-associated microbiome.

While both neutral and adaptive genetic regions have been implicated in being important for shaping gut microbial communities, the extent to which these mechanisms act together has not been explicitly tested. Moreover, disentangling host genetic effects from environmental effects is statistically challenging since host-associated microbiomes, host genetics, and landscape characteristics (e.g., climate, soil properties, vegetation) are often spatially autocorrelated. Both host-associated microbiomes and host genetics are subject to isolation by distance due to limited dispersal of both hosts and microbes (Martiny et al., 2006; Moeller et al., 2017), potentially generating spurious correlations where no direct link exists (Grieneisen et al., 2019). Mantel tests and their derivatives (e.g., multiple regression on distance matrices; MRMs) are a common method in microbiome analysis to test for correlations between genetic and ecological variation across spatially structured populations. However, Mantel tests suffer from a number of statistical drawbacks and can generate false positives (Legendre et al., 2015). Recently, a Bayesian inference method called *BEDASSLE* that was developed to account for spatial autocorrelation when testing the contribution of ecological variables to spatially structured genetic data (Bradburd et al., 2013), was applied to microbiomes (Grieneisen et al., 2019), showing that this method can be used to understand drivers of gut microbiome variation across spatially distributed host populations.

In this study, we combine island-level data on gut microbial species diversity, MHC allele composition, and microsatellite diversity (as a measure of neutral genetic variation and phylogenetic distance) across four Galápagos mockingbird species to test the relative contribution of neutral and adaptive genetic markers in shaping the gut microbiome composition. The Galápagos mockingbirds comprise four generalist species distributed allopatrically across almost all islands of the Galápagos Archipelago, diverging into four genetically divergent clades approximately 500,000 (95% credible interval: 145,957–1,388,173) years ago after a single colonization event (Nietlisbach et al., 2013). Previous studies have found that genetic drift has acted to reduce genetic diversity of mockingbirds compared with other Galápagos bird species (Hoeck et al., 2010) and that the MHC is under selection across islands and species (Vlček & Štefka, 2020), making the mockingbird complex a particularly good

study system to investigate the role of host genetics in shaping gut microbiomes.

We combine gut microbiome data, MHC, and microsatellite data collected largely from different individuals; therefore, we use island averages to test whether island populations that are genetically similar also tend to be more similar in their gut microbiomes. If host genetics tends to have large effects on the gut microbiome, we would expect correlations between genetic similarity and microbiome similarity between islands. Effects of genetic variation on the gut microbiome at the individual level require paired samples collected from the same individual and therefore cannot be detected by this study. Nevertheless, community-wide genetic effects, whereby allele frequencies across islands drive differences in overall gut microbiome composition, are detectable within this framework.

Here, we apply both MRM tests and *BEDASSLE* to test for the effects of neutral and MHC dissimilarity while controlling for geographic distance. Since phylogenetic signals in gut microbiomes have been shown to be weak yet detectable in avian clades (Kropáčková et al., 2017; Song et al., 2020; Trevelline et al., 2020), including Galápagos finches (Michel et al., 2018), we expect that adaptive immune genetic regions may have larger effects on microbiome composition than neutral markers in this study system.

2 | MATERIAL AND METHODS

2.1 | Study sites, species distribution, and sample sizes

Four endemic mockingbird species (*Mimus parvulus*, *M. trifasciatus*, *M. macdonaldi*, and *M. melanotis*) that allopatrically inhabit islands of the Galápagos Archipelago were sampled between 2004 and 2008 as part of one overarching study. Blood samples were collected from 2004 to 2008 and used for MHC and microsatellite genotyping; fecal samples were collected largely from different individuals (overlap of 31 individuals between blood and fecal samples) between 2006 and 2008. Sample metadata is listed in Table S1. We compared nine island populations (Isabela, Marchena, Santiago, Santa Cruz, Rábida, Champion, Gardner, San Cristóbal, and Española, Figure 1), for which data on microsatellite genetic diversity and differentiation (Hoeck et al., 2010) and MHC divergence (Vlček et al., 2016) were already available, and generated the microbiome data in the present study. Blood and fecal samples were not necessarily obtained from the same individual birds but from the same islands and species; therefore, we focus on island averages of genetic parameters and gut microbiome composition in our analysis.
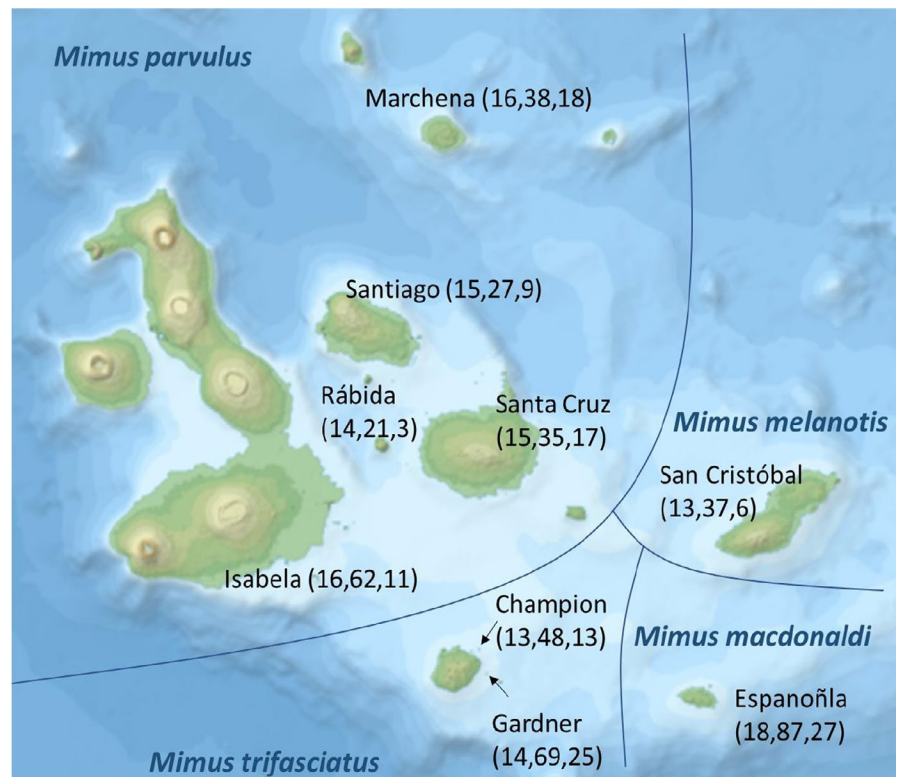

FIGURE 1 Map of the Galápagos Archipelago with species distributions, sample locations, and sample sizes. The numbers indicate the number of individuals sampled per island for MHC (first number), microsatellites (second number), and microbiome (third number)

2.2 | Assessment of neutral and adaptive genetic variation, and gut microbial species diversity per island

Variation in microsatellite genotypes among 135 individuals, based on 16 microsatellite loci, was estimated using the R package *adegenet* (Jombart, 2008) and visualized in a principal component analysis (PCA) including the first two axes that together explained 36% of variation. To calculate mean pairwise distances between islands, we generated an individual-based genetic distance matrix based on Nei's distance using the nei.dist() function in package *Poppr* (Kamvar et al., 2014) and calculated mean pairwise distance between islands using meandist() function in the package *vegan* (Oksanen et al., 2007).

We downloaded MHCII exon 2 B sequence data for 177 birds from Vlček et al. (2016). To estimate genetic distance between 119 MHC alleles, we generated a tree based on the genetic similarity of the whole MHC allele sequence using the packages *ape* (Paradis & Schliep, 2019) and *phangorn* (Schliep, 2011). This tree was rooted using an MHC allele sequence of a related species (*M. polyglottos*), which was afterward removed from the final tree. We then calculated pairwise dissimilarity of allele composition between all individuals per island, also taking allele genetic similarity into account, using the distance() function in *phyloseq* (McMurdie & Holmes, 2013), and calculated mean distance between islands using *vegan* package's meandist() function (Oksanen et al., 2007). While we also generated MHC supertypes to assess functional variation, these resulted in only seven supertypes, too few to calculate distances between islands (data not shown).

The microbiome data were derived from fecal samples of 129 birds. DNA was extracted from frozen fecal pellets and amplified using the paired primers 515F/ 806R for amplification of the V4 region of the bacterial 16S ribosomal RNA (16S rRNA) gene. Paired-end sequencing of the amplicons was performed with Illumina MiSeq technology over 2 × 250 cycles. The paired-end sequencing results were processed in QIIME 2 (Bolyen et al., 2019) using the DADA2 pipeline (Callahan et al., 2016). Before analysis, we excluded amplicon sequence variants (ASVs) that were not assigned to the kingdom bacteria (261), not assigned at phylum level (253), and assigned to chloroplasts (136) or mitochondria (3,017). All singletons (25,500) were excluded because they are likely to represent artefacts (Kopylova et al., 2016), but they represented only 3.4% of total reads. Alpha diversity was measured with three indices that emphasize different features of the individual microbial communities: the observed ASV richness which represents the count of ASVs, the Simpson index which takes into account the abundance of ASVs, and the Faith's phylogenetic diversity (PD) index which takes into account the phylogeny of ASVs. We calculated beta diversity based on unweighted UniFrac (which is based on presence/absence and controls for taxa phylogeny; Lozupone et al., 2011) and weighted UniFrac (which additionally considers abundance) and used multidimensional Scaling (MDS) based on the distance matrix for visualizing beta diversity. We then extracted mean distance between

island population centroids using the meandist() function in *vegan* (Oksanen et al., 2007). For beta diversity, samples were rarefied to a minimum sequence depth of 10,000 to normalize for sequencing depth (Weiss et al., 2017).

For statistical analyses, we generated two microbiome datasets, one representing the overall microbiome, excluding rare taxa with a prevalence of under 5% (leaving 1,228 ASVs), and one representing the island common core (each island microbiome dataset was filtered separately so that only taxa that had over 50% prevalence remained). We also applied 30% and 70% prevalence thresholds for the island core with no difference to the results. The island common core was applied to accentuate the differences in common gut microbes between islands, and because any effects of host genetics may act on local microbial communities. We excluded low prevalence taxa in the full microbiome dataset because they are difficult to model statistically and to limit computational time for the *BEDASSLE* analysis. However, sensitivity analyses with expanded number of ASVs did not change the results (Methods S1).

2.3 | Statistical modeling

We conducted analyses at both the individual level (to assess variation within each genetic layer) and island level (to test for associations between microbiome composition and host genetics). We estimated the effects of species and island on dissimilarity in gut microbiota ASV distributions (full microbiome and island core), genetic dissimilarity based on MHC allele frequencies, and genetic dissimilarity based on microsatellite markers, respectively, by using PERMANOVA tests on individual distance matrices described above, presenting standardized effect sizes (SES).

To assess the effects of MHC and microsatellite dissimilarity on microbiome composition at the average island level, we used two methods: First, we applied multiple regressions on distance matrices (MRM) that predict mean pairwise distance of microbiome composition between islands, using the package *ecodist* (Goslee & Urban, 2007). MRMs are a derivative of the Mantel test but allow multiple variables within the model. We included mean geographic distance between islands, as well a variable we term "comparison type," which codes whether the islands being compared are inhabited by the same species or different species, to control for these factors. Because the number of island populations being compared was small ($n = 9$ islands, generating 36 pairwise comparisons), there is a risk of overfitting the models, and model selection on MRMs using AIC generates spurious results (Franckowiak et al., 2017). Therefore, we present all possible models in Table S2a (predicting full microbiome) and Table S2b (predicting island core microbiome) and assess support for each variable across models. Statistics presented for PERMANOVAs and MRMs all predict microbiome dissimilarity based on unweighted UniFrac. Beta diversity plots based on weighted UniFrac are visualized in the supplementary material.

In addition to MRMs, which are prone to statistical problems and have low statistical power (Legendre et al., 2015), we applied

a Bayesian method called *BEDASSLE* (Bayesian estimation of differentiation in alleles by spatial structure and local ecology), which was developed to quantify the individual effects of geographic distance and ecological variables on population genetic structure (Bradburd et al., 2013). *BEDASSLE* has recently been applied to microbiomes (Grieneisen et al., 2019) and quantifies the geographic shift in microbiome distributions (in km) represented by a one-unit increase in the predictor variable (in this case, MHC and microsatellite genetic distances). If the effect size of microsatellites was 500 km, the interpretation would be that a one-unit shift in microsatellite genetic distance has the impact of approximately 500 km of extra pairwise geographic distance on gut microbiome distributions. The results should therefore be interpreted in the context of the Galápagos Islands, which span 270 km and have a mean distance between islands of about 80 km. For context, human populations separated by the Himalayas display genetic differences equivalent to 11–16,000 km lateral distance (Bradburd et al., 2013). We applied a beta-binomial model within the *BEDASSLE* R package to predict the occurrence of 1,228 ASVs and included distance matrices representing normalized mean pairwise dissimilarity in MHC and microsatellite markers between island populations as covariables, following the methodology of Grieneisen et al. (2019). We repeated this analysis for the island core microbiome (267 ASVs), which we ran for 20 million generations. Further information on *BEDASSLE* and its implementation is described in Supplementary Methods. All R code and

data can be downloaded at https://github.com/Riselya/Mockingbird-microbiome-paper.

3 | RESULTS

We first assessed variation in gut microbiome, MHC, and microsatellite at the individual level to examine the effects of species and island on the three levels of genetic data separately. Gut microbiome diversity demonstrated large interindividual variation (Figure S1). Alpha diversity was largely similar across species, though *M. melanotis* tended to have slightly lower diversity across all measures (Figure S2). Beta diversity clustered weakly by species and island when including all taxa (Figure 2a; see Figure S3 for weighted UniFrac). As expected (because the island core microbiome was generated per island), beta diversity clustered strongly by both species and island for the island core microbiome (Figure 2b). MHC allele composition exhibited somewhat stronger clustering by species and by island than the full microbiome (Figure 2c), while, in accordance with earlier results (Hoeck et al., 2010), microsatellite profiles showed the strongest clustering by species and more moderate clustering by island (Figure 2d).

We then examined associations between island centroids for the three layers of data. We estimated the effects of MHC and microsatellite dissimilarity on both the full microbiome (Figure 3a–c) and island
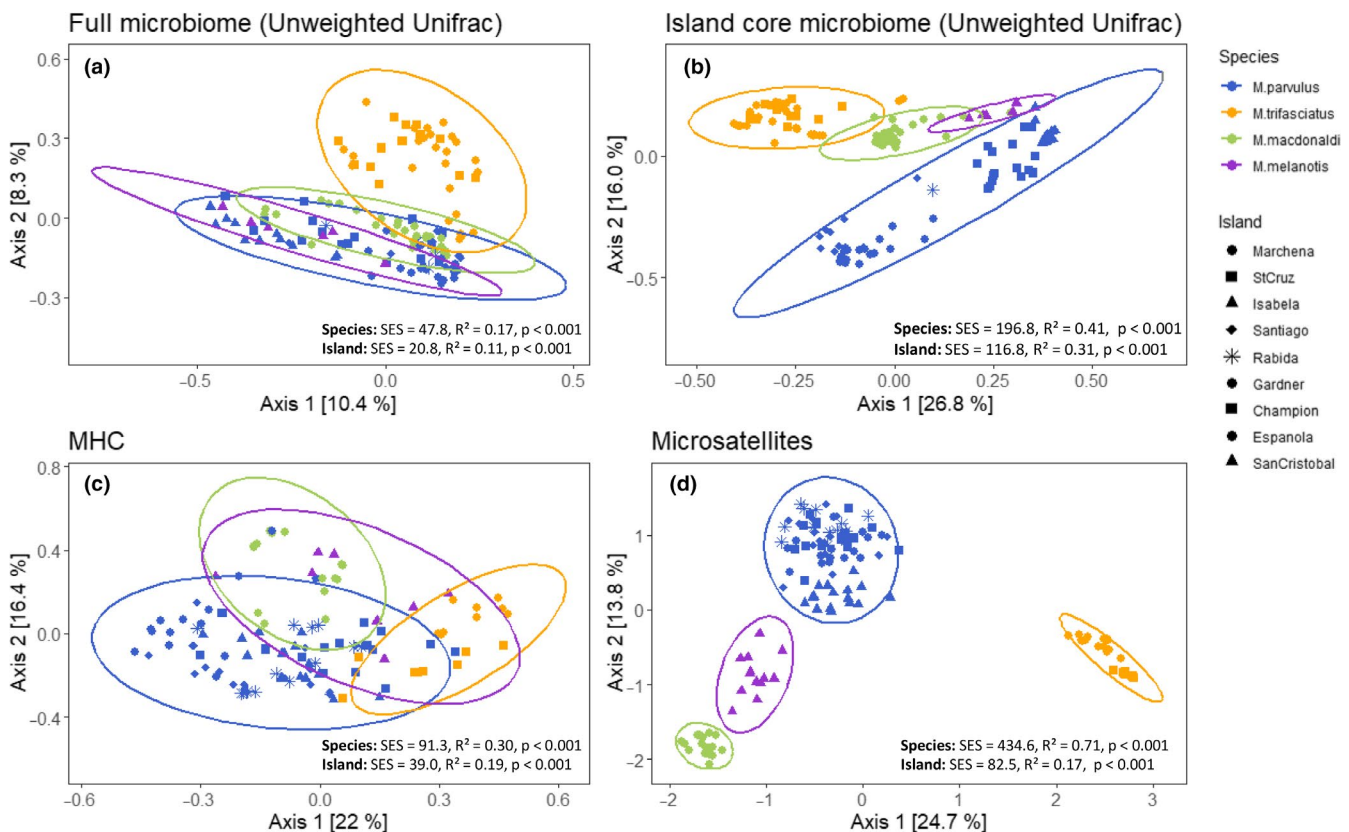

FIGURE 2 Ecological dissimilarity between species and islands for (a) the overall microbiome; (b) the island core microbiome; (c) MHC genotypes; and (d) microsatellite genotypes. Points represent sampled individuals and are colored by species and shaped by island. SES, standardized effect size
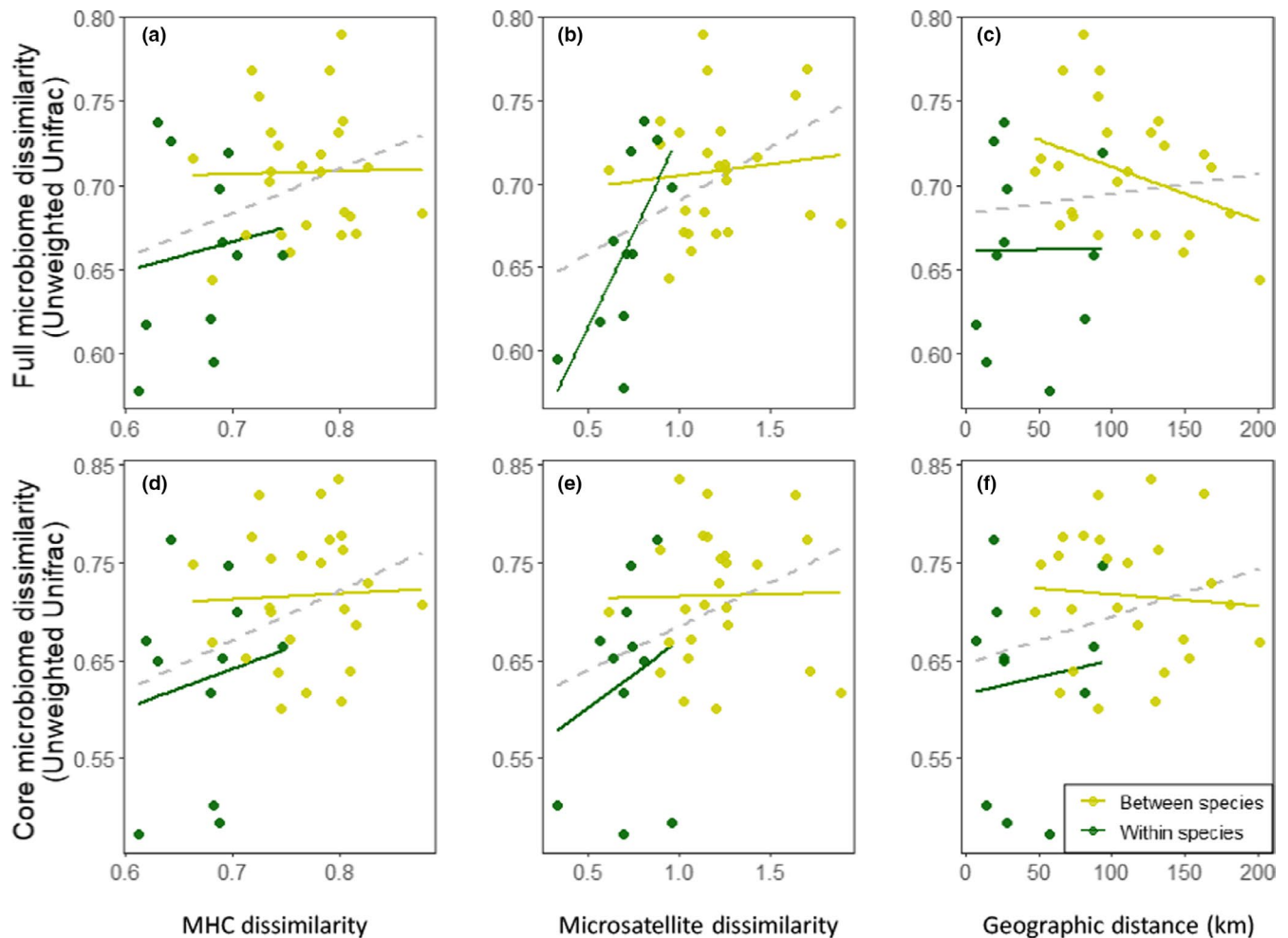
FIGURE 3 (a–c) Pairwise dissimilarity between overall microbiome beta diversity and (a) MHC dissimilarity; (b) microsatellite dissimilarity; and (c) geographic distance. Bottom panel (d–f) shows pairwise distances between the island core microbiome and (d) MHC dissimilarity; (e) microsatellite dissimilarity; and (f) geographic distance. Each point represents the mean pairwise distance between two islands and is colored by whether the comparison is within species (dark green) or between species (light green). The gray dashed line shows the overall trend (statistics found in Table S2).

core microbiome (Figure 3d,e) composition using both MRM tests, controlling for geographic distance and comparison type (within/ between species) and *BEDASSLE*. Genetic distances based on both MHC and microsatellite dissimilarity significantly correlated with gut microbiome dissimilarity in univariate MRM tests for both overall and island core microbiomes (Figure 3a,b,d,e; Table S2a,b, respectively). Across multivariate MRM models (models 5–15 Table S2a,b), however, there was little support for the effect of MHC distance on overall microbiome or island core composition (Figure 3a; Table S2a,b), and some support for the effect of microsatellite distance on overall microbiome composition (Table S2a) but little for island core microbiome composition (Table S2b). We noted a significant correlation between MHC and microsatellite data when applying Mantel tests, and both of these variables were also correlated with geographic distance (MHC—microsatellite: Mantel $R = 0.49$, $p = .004$; MHC—geographic distance: Mantel $R = 0.58$, $p = .004$; microsatellite—geographic distance: Mantel $R = 0.28$, $p = .002$; Figure S4).

To effectively control for landscape effects, we additionally applied *BEDASSLE* to estimate the relative contribution of MHC and microsatellite distances on gut microbiome distributions (1,228 ASVs with over 5% prevalence). We found that a one-unit increase in microsatellite genetic distance (i.e., one-unit increase in microsatellite Euclidean distance) shifted microbiome distributions by the equivalent of 45.4 km (Table 1a). When this effect size was normalized to fit the distribution of our microsatellite data (0.69–1.22; see Figure 2d), this was reduced to 24.1 km (Table 1a, right side panel). Given that the Galápagos islands span 270 km, and the mean distance between islands is 80 km, 24.1 km effectively denotes an effect size of zero. The effect of MHC distance (i.e., one-unit increase in MHC Euclidean distance) was predicted to have zero effect on gut microbiome distributions. We repeated this analysis on island core microbiomes (267 ASVs) and found similarly small normalized effects of microsatellite distance (4.2 km) and MHC distance (5.1 km).

TABLE 1 BEDASSLE effect sizes for the effect of microsatellite and MHC distances on (a) the full microbiome and (b) the island core microbiome

Predictor variables	Mean geographical distance equivalent (km) per 1 unit predictor variable (95% CI)	Effect size for each predictor (in km) normalized to reflect the difference between the 25th and 75th percentiles in our microsatellite and MHC data
(a) Full microbiome		
Microsatellite	45.4 (34.7–56.4)	24.1
MHC	6.1 (4.5–8.0)	0.6
(b) Island core microbiome		
Microsatellite	8.3 (0–20.1)	4.2
MHC	42.9 (19.0–76.0)	5.1

Note: The first panel shows effect sizes per one-unit increase in the predictor variable, while the second panel normalizes this to reflect the middle 50% of the data.

4 | DISCUSSION

This study aimed to quantify the effects of genetic dissimilarity based on host neutral and MHC allele frequencies on the gut microbiome ASV distributions of the Galápagos mockingbird species complex at the island population level. Our results support other studies that find relatively weak correlations between neutral genomic diversity and microbiome composition (e.g., Griffiths et al., 2019; Kohl et al., 2017; Smith et al., 2015; Suzuki et al., 2019; Webster et al., 2018). Analyses with *BEDASSLE* suggest that this correlation is a consequence of spatial autocorrelation between host genetic and microbiome clustering, rather than a direct consequence of host genetics acting upon gut microbial communities. Similar results were observed regarding the influence of MHC dissimilarity on microbiome distributions, contradicting our prediction that adaptive genetic regions may exert stronger influence on the gut microbiome than neutral markers in avian systems. Our results suggest that neither genetic differentiation based on neutral nor adaptive markers appear to exert detectable control over gut microbiome distributions in Galápagos mockingbirds at the population level. The weak clustering of gut microbiomes across species and island observed in this system is therefore likely to be the result of biogeography alone. Our results highlight that spatial autocorrelation between ecological and genetic variables due to limitations in both host and microbe dispersal (e.g., Moeller et al., 2017) may lead to an overestimate of the link between host genetics and microbiome composition.

Our results add to the increasing number of studies showing that phylogenetic signals in bird gut microbiomes are relatively weak (Michel et al., 2018; Song et al., 2020; Trevelline et al., 2020). They also suggest that in a species complex with a limited species range, and where pathogen diversification might be relatively low, regions of the genome that reflect adaptation to local pathogen-driven selection pressures, such as the MHC, are rather unimportant in explaining the diversification of the gut microbiome. These results may not apply to mammalian lineages or to other bird phylogenies covering larger biogeographic ranges. Mammals show much stronger signals of phylosymbiosis than birds, and demonstrate phylogenetic signals in their gut microbiomes even when accounting

for geographic distance (Amato et al., 2018; Kartzinel et al., 2019). Nevertheless, the development of methods that account for spatial autocorrelation between host genetics and microbiomes due to isolation by distance acting on both host and microbe communities may improve estimates of their direct relationship across host phylogenies, including mammals.

Our study applies microbiome and genetic data from different individuals, therefore is not able to detect genetic effects on the microbiome at the individual level. Paired samples, where gut microbiome, microsatellite, and MHC data all come from the same individual, would allow for more sophisticated statistical methods (e.g., redundancy analyses or joint species distribution modelling) with increased power to detect weak effects and associations between species alleles and microbial taxa, thereby helping to elucidate underlying mechanisms driving phylosymbiosis. For example, mockingbird species may share genetic elements that strongly affect the microbiome yet would not be detected here. Our results therefore do not preclude that host genetics influences gut microbe distributions in this system, yet suggest that if such control exists it is likely to be targeted towards specific microbial genera, governed by genetic regions not assessed in this study, or only detectable at larger phylogenetic scales. MHC polymorphism has been associated with the abundance of a limited number of microbial families (Bolnick et al., 2014; Kubinak et al., 2015), suggesting its influence may be obscured by environmental effects, or ecological differences in life history (such as differences in diet and food availability across or even within islands). Nevertheless, our study mirrors the results from other Galápagos bird species that show only weak clustering of species or islands microbiome(Michel et al., 2018), which indicates that any effects of host genetics on overall microbiome composition are likely to be weak at this phylogenetic scale. Our study confirms this lack of association at the island population level in a species complex with well-documented population genetics.

This study does not test for the effects of specific MHC alleles on gut microbiome distributions, and we note that specific MHC alleles are likely to have targeted effects on the gut microbiome. The detection of such effects remains challenging where genetic effects are weak and numbers of populations examined are necessarily limited

by species range and structure. In this study, we examined nine genetically isolated populations across four closely related species, with numbers of populations in other studies ranging from five to 15 (Griffiths et al., 2018, 2019; Kohl et al., 2017; Smith et al., 2015; Suzuki et al., 2019; Webster et al., 2018). Limited statistical power may therefore to some degree explain a lack of robust associations between genetics and microbiomes, especially for nonmammalian clades for which phylosymbiosis appears relatively weak. We also found that MHC and neutral dissimilarity were correlated in this study, which may also limit statistical power to distinguish between the two. Such correlations are common in small populations where genetic drift can act on both neutral and adaptive alleles (Radwan et al., 2010; Zeisset & Beebee, 2014). Therefore, while our results suggest no broad effects of genetics on gut microbiome distributions, weak or targeted effects of host genetics on the microbiome that scale up with biogeographic distance and host phylogeny may not have been detected within this study design.

5 | CONCLUSIONS

Our study brings together three layers of genetic data to test the relative contribution of neutral (microsatellite) and adaptive (MHC) genetic regions in shaping gut microbiome diversity in Galápagos mockingbirds. Whileour study is necessarily limited by the number of mockingbird populations, our study is the first to simultaneously test the independent contributions of neutral and MHC dissimilarity on gut microbiome distributions across a well-studied avian species complex. Our results support previous studies finding weak effects of phylogeny on bird gut microbiomes, yet go further to suggest that weak correlations between host genetics and microbiomes in allopatric populations may be a consequence of spatial autocorrelation between these factors. Future research that investigates the targeted effects of host genetics on specific microbial lineages within individuals may elucidate the relationship between avian genetics and the gut microbiome at a finer-scale resolution.

CONFLICT OF INTEREST
No authors have any conflicting interests in regard to this paper.

AUTHOR CONTRIBUTION
Ramona Fleischer: Conceptualization (supporting); Formal analysis (equal); Investigation (lead); Methodology (lead); Writing-original draft (equal); Writing-review & editing (equal). **Alice Risely:** Conceptualization (equal); Formal analysis (equal); Supervision (equal); Validation (equal); Writing-original draft (lead); Writing-review & editing (equal). **Paquita E. A. Hoeck:** Data curation (lead); Investigation (supporting); Methodology (supporting); Writing-review & editing (equal). **Lukas Fridolin Keller:** Data curation (lead); Investigation (supporting); Methodology (supporting); Writing-review & editing (equal). **Simone Sommer:** Conceptualization (equal); Resources (lead); Supervision (lead); Writing-review & editing (equal).

DATA AVAILABILITY STATEMENT
Data and code for this manuscript are publicly available to download at https://github.com/Riselya/Mockingbird-Microbiome-Phylo symbiosis-Project. Microbiome sequences are available under project PRJNA658437 on NCBI. MHC and microsatellite data are downloadable from the original studies.

ORCID
Ramona Fleischer https://orcid.org/0000-0003-1657-9347
Alice Risely https://orcid.org/0000-0002-0731-2934

REFERENCES
Amato, K. R., Sanders, J. G., Song, S. J., Nute, M., Metcalf, J. L., Thompson, L. R., & Humphrey, G. (2018). Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes. *ISME Journal*, 13, 576–587.

Bolnick, D. I., Snowberg, L. K., Caporaso, J. G., Lauber, C., Knight, R., & Stutz, W. E. (2014). Major Histocompatibility Complex class IIb polymorphism influences gut microbiota composition and diversity. *Molecular Ecology*, 23(19), 4831–4845.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C., Al-Ghalith, G. A., Asnicar, F. (2019). QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *Nature biotechnology*, 37(8), 852-857.

Bradburd, G. S., Ralph, P. L., & Coop, G. M. (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution*, 67(11), 3258–3273. https://doi.org/10.1111/evo.12193

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 Methods*, 13(7), 581. https://doi.org/10.1038/nmeth.3869

Davenport, E. R. (2016). Elucidating the role of the host genome in shaping microbiome composition. *Gut Microbes*, 7(2), 178–184. https://doi.org/10.1080/19490976.2016.1155022

Donaldson, G. P., Ladinsky, M. S., Yu, K. B., Sanders, J. G., Yoo, B. B., Chou, W.-C., Conner, M. E., Earl, A. M., Knight, R., Bjorkman, P. J., & Mazmanian, S. K. (2018). Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science*, 360(6390), 795–800. https://doi.org/10.1126/science.aaq0926

Franckowiak, R. P., Panasci, M., Jarvis, K. J., Acuna-Rodriguez, I. S., Landguth, E. L., Fortin, M.-J., & Wagner, H. H. (2017). Model selection with multiple regression on distance matrices leads to incorrect inferences. *PLoS One*, 12(4), e0175194. https://doi.org/10.1371/journal.pone.0175194

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*(7), 1–19.

Grieneisen, L. E., Charpentier, M. J., Alberts, S. C., Blekhman, R., Bradburd, G., Tung, J., & Archie, E. A. (2019). Genes, geology and germs: Gut microbiota across a primate hybrid zone are explained by site soil properties, not host species. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1901), 20190431. https://doi.org/10.1098/rspb.2019.0431

Griffiths, S. M., Antwis, R. E., Lenzi, L., Lucaci, A., Behringer, D. C., Butler, M. J. IV, & Preziosi, R. F. (2019). Host genetics and geography influence microbiome composition in the sponge *Ircinia campana*. *Journal of Animal Ecology*, *88*(11), 1684–1695.

Griffiths, S. M., Harrison, X. A., Weldon, C., Wood, M. D., Pretorius, A., Hopkins, K., Fox, G., Preziosi, R. F., & Antwis, R. E. (2018). Genetic variability and ontogeny predict microbiome structure in a disease-challenged montane amphibian. *ISME Journal*, *12*(10), 2506–2517. https://doi.org/10.1038/s41396-018-0167-0

Hoeck, P. E., Bollmer, J. L., Parker, P. G., & Keller, L. F. (2010). Differentiation with drift: A spatio-temporal genetic analysis of Galapagos mockingbird populations (*Mimus* spp.). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1543), 1127–1138.

Holderegger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology*, *21*(6), 797–807.

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Kartzinel, T. R., Hsing, J. C., Musili, P. M., Brown, B. R., & Pringle, R. M. (2019). Covariation of diet and gut microbiome in African megafauna. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(47), 23588–23593. https://doi.org/10.1073/pnas.1905666116

Kohl, K. D. (2020). Ecological and evolutionary mechanisms underlying patterns of phylosymbiosis in host-associated microbial communities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1798), 20190251. https://doi.org/10.1098/rstb.2019.0251

Kohl, K. D., Varner, J., Wilkening, J. L., & Dearing, M. D. (2017). Gut microbial communities of American pikas (*Ochotona princeps*): Evidence for phylosymbiosis and adaptations to novel diets. *Journal of Animal Ecology*, *87*(2), 323–330.

Kopylova, E., Navas-Molina, J. A., Mercier, C., Xu, Z. Z., Mahé, F., He, Y., Zhou, H.-W., Rognes, T., Caporaso, J. G., & Knight, R. (2016). Open-source sequence clustering methods improve the state of the art. *mSystems*, *1*(1), e00003-15. https://doi.org/10.1128/mSystems.00003-15

Kropáčková, L., Těšický, M., Albrecht, T., Kubovčiak, J., Čížková, D., Tomášek, O., & Procházka, P. (2017). Co-diversification of gastrointestinal microbiota and phylogeny in passerines is not explained by ecological divergence. *Molecular Ecology*, *26*(19), 5292–5304. https://doi.org/10.1111/mec.14144

Kubinak, J. L., Stephens, W. Z., Soto, R., Petersen, C., Chiaro, T., Gogokhia, L., Bell, R., Ajami, N. J., Petrosino, J. F., Morrison, L., Potts, W. K., Jensen, P. E., O'Connell, R. M., & Round, J. L. (2015). MHC variation sculpts individualized microbial communities that control susceptibility to enteric infection. *Nature Communications*, *6*, 8642. https://doi.org/10.1038/ncomms9642

Kurilshikov, A., Wijmenga, C., Fu, J., & Zhernakova, A. (2017). Host genetics and gut microbiome: Challenges and perspectives. *Trends in Immunology*, *38*(9), 633–647. https://doi.org/10.1016/j.it.2017.06.003

Leclaire, S., Strandh, M., Dell'Ariccia, G., Gabirot, M., Westerdahl, H., & Bonadonna, F. (2019). Plumage microbiota covaries with the major histocompatibility complex in blue petrels. *Molecular Ecology*, *28*(4), 833–846. https://doi.org/10.1111/mec.14993

Legendre, P., Fortin, M. J., & Borcard, D. (2015). Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, *6*(11), 1239–1247. https://doi.org/10.1111/2041-210X.12425

Lim, S. J., & Bordenstein, S. R. (2020). An introduction to phylosymbiosis. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1922), 20192900. https://doi.org/10.1098/rspb.2019.2900

Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *ISME Journal*, *5*(2), 169–172. https://doi.org/10.1038/ismej.2010.133

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Øvreås, L., Reysenbach, A.-L., Smith, V. H., & Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102–112. https://doi.org/10.1038/nrmicro1341

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. https://doi.org/10.1371/journal.pone.0061217

Michel, A. J., Ward, L. M., Goffredi, S. K., Dawson, K. S., Baldassarre, D. T., Brenner, A., Gotanda, K. M., McCormack, J. E., Mullin, S. W., O'Neill, A., Tender, G. S., Uy, J. A. C., Yu, K., Orphan, V. J., & Chaves, J. A. (2018). The gut of the finch: Uniqueness of the gut microbiome of the Galápagos vampire finch. *Microbiome*, *6*(1), 167. https://doi.org/10.1186/s40168-018-0555-8

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 of the United States of America*, *114*(52), 13768–13773. https://doi.org/10.1073/pnas.1700122114

Nietlisbach, P., Wandeler, P., Parker, P. G., Grant, P. R., Grant, B. R., Keller, L. F., & Hoeck, P. E. (2013). Hybrid ancestry of an island subspecies of Galápagos mockingbird explains discordant gene trees. *Molecular Phylogenetics and Evolution*, *69*(3), 581–592. https://doi.org/10.1016/j.ympev.2013.07.020

Noguera, J. C., Aira, M., Pérez-Losada, M., Domínguez, J., & Velando, A. (2018). Glucocorticoids modulate gastrointestinal microbiome in a wild bird. *Royal Society Open Science*, *5*(4), 171743. https://doi.org/10.1098/rsos.171743

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. (2007). *The vegan package. Community ecology package*, *10*(631-637), 719.

Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528.

Pearce, D. S., Hoover, B. A., Jennings, S., Nevitt, G. A., & Docherty, K. M. (2017). Morphological and genetic factors shape the microbiome of a seabird species (*Oceanodroma leucorhoa*) more than environmental and social factors. *Microbiome*, *5*(1), 146. https://doi.org/10.1186/s40168-017-0365-4

Piertney, S., & Oliver, M. (2006). The evolutionary ecology of the major histocompatibility complex. *Heredity*, *96*(1), 7–21. https://doi.org/10.1038/sj.hdy.6800724

Radwan, J., Biedrzycka, A., & Babik, W. (2010). Does reduced MHC diversity decrease viability of vertebrate populations? *Biological Conservation*, *143*(3), 537–544. https://doi.org/10.1016/j.biocon.2009.07.026

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., Costea, P. I., Godneva, A., Kalka, I. N., Bar, N., Shilo, S., Lador, D., Vila, A. V., Zmora, N., Pevsner-Fischer, M., Israeli, D., Kosower,

N., Malka, G., Wolf, B. C., ... Segal, E. (2018). Environment domi- nates over host genetics in shaping human gut microbiota. *Nature*, *555*(7695), 210–215. https://doi.org/10.1038/nature25973

Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, *27*(4), 592–593. https://doi.org/10.1093/bioinformatics/btq706

Silverman, M., Kua, L., Tanca, A., Pala, M., Palomba, A., Tanes, C., Bittinger, K., Uzzau, S., Benoist, C., & Mathis, D. (2017). Protective major histocompatibility complex allele prevents type 1 diabetes by shaping the intestinal microbiota early in ontogeny. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(36), 9671–9676. https://doi.org/10.1073/pnas.1712280114

Smith, C. C., Snowberg, L. K., Caporaso, J. G., Knight, R., & Bolnick, D. I. (2015). Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *ISME Journal*, *9*(11), 2515–2526. https://doi.org/10.1038/ismej.2015.64

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, *2*(1), 16.

Song, S. J., Sanders, J. G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M. W., Mazel, F., Lutz, H. L., Winker, K., Graves, G. R., Humphrey, G., Gilbert, J. A., Hackett, S. J., White, K. P., Skeen, H. R., Kurtis, S. M., Withrow, J., Braile, T., Miller, M., ... Knight, R. (2020). Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *MBio*, *11*(1), e02901-19. https://doi.org/10.1128/mBio.02901 -19

Suzuki, T. A., Phifer-Rixey, M., Mack, K. L., Sheehan, M. J., Lin, D., Bi, K., & Nachman, M. W. (2019). Host genetic determinants of the gut mi- crobiota of wild mice. *Molecular Ecology*, *28*(13), 3197–3207. https:// doi.org/10.1111/mec.15139

Trevelline, B. K., Sosa, J., Hartup, B. K., & Kohl, K. D. (2020). A bird's- eye view of phylosymbiosis: Weak signatures of phylosymbiosis among all 15 species of cranes. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1923), 20192988. https://doi.org/10.1098/ rspb.2019.2988

Vlček, J., Hoeck, P. E., Keller, L. F., Wayhart, J. P., Dolinová, I., & Štefka, J. (2016). Balancing selection and genetic drift create unusual patterns of MHCII β variation in Galápagos mockingbirds. *Molecular Ecology*, *25*(19), 4757–4772. https://doi.org/10.1111/mec.13807

Vlček, J., & Štefka, J. (2020). Association between louse abundance and MHC II supertypes in Galápagos mockingbirds. *Parasitology Research*, *119*(5), 1–9. https://doi.org/10.1007/s00436-020-06617-3

Webster, T. M. U., Consuegra, S., Hitchings, M., & de Leaniz, C. G. (2018). Interpopulation variation in the Atlantic salmon microbiome re- flects environmental and genetic diversity. *Applied and Environment Microbiology*, *84*(16), e00691-18. https://doi.org/10.1128/ AEM.00691-18

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J. R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E. R., & Knight, R. (2017). Normalization and microbial dif- ferential abundance strategies depend upon data characteristics. *Microbiome*, *5*(1), 27. https://doi.org/10.1186/s40168-017-0237-y

Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., Ley, R. E., & Farnleitner, A. H. (2019). Host diet and evo- lutionary history explain different aspects of gut microbiome diver- sity among vertebrate clades. *Nature Communications*, *10*(1), 2200. https://doi.org/10.1038/s41467-019-10191-3

Yuan, M. L., Dean, S. H., Longo, A. V., Rothermel, B. B., Tuberville, T. D., & Zamudio, K. R. (2015). Kinship, inbreeding and fine-scale spatial struc- ture influence gut microbiota in a hindgut-fermenting tortoise. *Molecular Ecology*, *24*(10), 2521–2536. https://doi.org/10.1111/mec.13169

Zeisset, I., & Beebee, T. J. (2014). Drift rather than selection dominates MHC class II allelic diversity patterns at the biogeographical range scale in natterjack toads *Bufo calamita*. *PLoS One*, *9*(6), e100176. https://doi.org/10.1371/journal.pone.0100176

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.