Body size poorly predicts host-associated microbial diversity in wild birds

Elizabeth Herder [1], Holly Lutz[2], and Sarah Hird[1]

[1]University of Connecticut
[2]Field Museum of Natural History

January 25, 2022

Abstract

The extent to which the avian microbiome is shaped by host phylogeny relative to other factors is largely unknown. In this study, we examine microbial biodiversity across multiple body sites of 211 bird species sampled in Malawi. Microbial community dissimilarity differed significantly across body sites, which included blood, buccal cavity, gizzard, intestine, cloaca, liver, and spleen. With these data, we tested the hypothesis that the avian microbiota follow a Species-Area Relationship by using a comparative phylogenetic method to examine the correlation between microbiota richness and host weight. Using Pagel's lambda, we confirmed that bird mass is significantly correlated with host phylogeny but found that few microbial diversity metrics showed such a correlation. Phylogenetic Generalized Least Squares identified a significant but weak negative correlation between host weight and microbial richness of the blood and a similarly significant but weak positive correlation between the cloacal microbiota and host weight among birds within the order Passeriformes. Taken together, these results suggest that the avian microbiome does not follow a traditional species-area relationship when phylogenetic relatedness is considered, rather, microbial diversity is influenced by factors beyond host phylogeny and size.

INTRODUCTION

Microbiomes are both a trait of the host (Benson et al., 2010) and ecological communities comprised of microorganisms capable of complex and dynamic interactions (Kodera et al. 2022). Ecological theory provides specific hypotheses for testing and inferring rules of life that apply to all organisms (Koskella, Hall, & Metcalf, 2017), and ecological processes are essential to understanding the composition, stability, and evolution of the microbiome (Miller, Svanbäck, & Bohannan, 2018). The Theory of Island Biogeography (MacArthur & Wilson, 2001) is an ecological theory that relates the size of an island and its biodiversity through parameterizing factors such as immigration and extinction (MacArthur & Wilson, 2001). With larger islands, extinction events are predicted to be less likely to occur due to the greater availability of space and resources. With more isolated islands, the arrival of new immigrants is predicted to be less likely and thus fewer new species come into the space. Conceptualizing hosts as "islands" is reasonable: both geographic islands and vertebrate hosts harbor complex biological communities that are connected by ecological processes (e.g., dispersal and immigration) and that are limited by resources and space. Communities can be significantly impacted by random environmental events and follow successional processes in the face of disturbance (Karl et al., 2018). The difference is scale. Another ecological theory, Metacommunity Theory (Leibold et al., 2004), incorporates scale in how communities interact with local and regional processes (Miller et al., 2018). According to Metacommunity Theory, communities exist in patches that are connected by dispersal and are hierarchically nested within larger patches. Importantly, both (1) properties of the physical space and (2) traits of the organisms within the community impact successful colonization of a new habitat and probability of survival (Miller & Bohannan, 2019).

The consideration of hosts and their specific body sites as patches of biodiversity that are affected by processes

shared by all life is a powerful way to test and identify "universal" rules of life (Ma & Li, 2018; Li et al., 2020; L. Li & Ma, 2016). Body size has been shown to positively correlate with bacterial richness, implying adherence to Species-Area Relationships (Sherrill-Mix et al., 2018), although the non-independence of host species was not accounted for in this case. The Theory of Island Biogeography has been demonstrated in the human lung microbiota, where sites farther from the "mainland source" are less diverse (Dickson et al., 2015) and tests of microbial composition against a neutral assembly model have identified diseased lung microbiomes as under selection (Venkataraman et al., 2015).

Birds (class: Aves) are important members of Earth's biosphere and to fully understand their biology requires knowledge of their microbiota. Furthermore, bird body sizes span five orders of magnitude by weight, making them an excellent clade for exploring species-area relationships in host-associated microbiota. Larger birds exhibit a greater area for microorganisms to occupy than smaller birds, which may provide increased ecological niches and lead to fewer extinction events as bacteria are less likely to compete for resources. Microbial colonization may also be higher in larger birds due to intrinsic qualities and life history traits (e.g. greater food requirements, larger territories), leading to increased exposure to diverse microorganisms. Alternatively, higher immune cell output of larger birds (Ruhs, Martin, & Downs, 2020) may inhibit the establishment of new microbial colonizers.

The microbiome is not only an ecological community, it is also a trait of its host (Benson et al., 2010). To understand the evolution of any trait in a comparative context, we must also consider the underlying phylogeny (Felsenstein, 1985), as many organismal traits are not independent of evolutionary relationships. To appropriately test the relationship between island size and (micro)biodiversity, and to ensure any correlations are not simply a factor of the relatedness of the host species, we use phylogenetic comparative analyses. Five body sites were the focus of our analyses: four are distinct sites along the gastrointestinal tract connected to each other through digestion (buccal, gizzard, intestines, cloaca) and the fifth is blood. These sample types encompass diverse environmental conditions and may follow Island Biogeography principles more or less strongly. The buccal, gizzard, intestine, and cloacal samples are frequently exposed to external microbes through the intake of food containing distinct microbiota that may be able to newly colonize those areas. The blood (and liver and spleen) sample types are in contact with new microbes rarely and thus will have fewer potential "immigrants" into their microbial communities.

There were several goals of this paper. First, we describe the taxonomic composition and diversity of the microbiomes of hundreds of wild birds at various body sites. Second, we compare the microbiota of the body sites, and identify conserved and unique members. Third, we estimate the phylogenetic signal of microbiome diversity using Pagel's lambda, and fourth, we address the relationship between host body size and microbiome diversity using a phylogenetically controlled method. Together, these aims expand what is known about the microorganisms, the birds and the processes structuring the avian microbiome.

METHODS

Sampling and Sequencing

Sampling was conducted as previously described (Lutz et al., 2015). Briefly, birds were collected using a combination of mist-netting and ballistic capture in 2009 and 2011. Blood samples were obtained via brachial venipuncture; oral (buccal) and cloaca samples were collected by swabbing with sterile cotton swabs; liver, spleen, and intestinal tracts were dissected from euthanized animals. All samples were stored in cryogenic vials and flash-frozen in liquid nitrogen until nucleic acid extraction. Samples collected for this project were processed following the Earth Microbiome Project standard processing protocols (Thompson et al., 2017). Briefly, DNA was extracted using 96-well PowerSoil PowerMag DNA extraction plates (Qiagen), which were homogenized using a TissueLyser beadbeater (Qiagen). From eluted DNA, triplicate PCRs using the 515f/806r EMP primers amplified the V4 region of the 16S rRNA gene, pooled amplicons were sequenced on Illumina MiSeq and HiSeq instruments, and data were uploaded to the Qiita web-based microbiome analysis platform for initial processing (made available previously via QIITA ID 11166 by (Song et al., 2020). All sequences were demultiplexed and quality filtered using the Qiime2 pipeline (Bolyen et al., 2019) and forward

and reverse reads were joined and processed using Deblur to remove sequencing errors and refine sequences to amplicon sequence variants (ASVs) (Amir et al., 2017). Deblurred ASV tables and sample metadata were further processed using Qiime2 (Bolyen et al., 2019). To assess the presence of artefactual batch effects associated with microbiota diversity, statistical significance of the plate on which the samples were placed on were calculated on each sample type using the adonis2 function (PERMANOVA) from the vegan package using unweighted UniFrac and weighted UniFrac (McArdle & Anderson, 2001); (Oksanen et al., 2018). Only the unweighted UniFrac distances of the intestine samples showed significant (p <0.05) differences due to plate (platename R^2= 1.714%, p = 0.05).

Data analysis

Rarefaction : Statistical analyses were conducted using the programming language R version 4.1.1 (Team, 2021). Rarefaction curves from the vegan package (Oksanen et al., 2018) were produced to find the number to which to rarefy for sufficient sequencing depth. For microbial analyses and visualization of results, the phyloseq and ggplot2 packages were used (McMurdie & Holmes, 2013); (Wickham, 2009). For initial sample analyses, all samples were combined. They were pruned and rarefied to 4,000 sequences per sample leading to 1,606 samples. For sample-type specific analyses, all samples were divided into one of seven sample types, pruned, and rarefied at a sample-type specific value (Blood: 4,030 sequences, Buccal: 10,961 sequences, Cloaca: 10,329 sequences, Intestines: 9,094 sequences, Gizzard: 22,306 sequences, Liver: 8,658 sequences, Spleen: 11,701 sequences). Rarefaction curves are available in the supplementary material (Supplemental Fig. S1).

Taxonomic composition: Taxonomy was assigned using the Qiime2 naive Bayes feature classifier trained against the Greengenes 13_8 reference (DeSantis et al., 2006). Phylum level taxonomic information was exported to Microsoft Excel version Microsoft Office Professional Plus 2016 to create bar plots and determine phylum percentages; phyla represented at less than 10% of the sequences in each distinct sample were grouped into the "other phyla" category.

Alpha diversity: We calculated alpha diversity using the Observed number of ASVs (richness), the Shannon diversity index (Shannon & Weaver, 1949), and the Simpson Diversity Index (Simpson, 1949); these were calculated using the phyloseq package (McMurdie & Holmes, 2013). The three alpha diversity measurements that will be used are observed ASVs, the Shannon index (Shannon & Weaver, 1949), and the Simpson index (Simpson, 1949). Observed ASVs informs how many different types of operational taxonomic units (OTUs) are within the sample without accounting for abundance. Both the Shannon index and the Simpson index account for both the quantity of distinct OTUs and the abundances of those OTUs. However, the Shannon Index quantifies the amount of randomness within the dataset by considering the amount of distinct OTUs and how evenly the abundance of sequences are distributed across them and the Simpson index investigates the proportion of OTUs in the samples.

Beta diversity: Non-metric Multidimensional Scaling (NMDS) visualized unweighted UniFrac and weighted UniFrac distances (Lozupone & Knight, 2005). Statistical significance was determined using the adonis2 function (PERMANOVA) within the vegan package (McArdle & Anderson, 2001); (Oksanen et al., 2018).

Differential abundance: Plots showing the logarithmic differences in the quantities of taxa between sample types were created using the package DESeq2 in R (Love, Huber, & Anders, 2014). Each sample type pair was rarefied to the lower rarefaction point used for the sample types separately (i.e. comparisons between Blood and Buccal samples were rarefied to 4,030 sequences per sample). An alpha cutoff of 0.01 was used. Venn diagrams were created using BioVenn (Hulsen, 2021).

Phylogenetic comparative analyses: BirdTree.org was used to create phylogenetic trees of the bird species represented in each sample type (Fig. 1) (W. Jetz, Thomas, Joy, Hartmann, & Mooers, 2012), (Walter Jetz et al., 2014). Hackett All Species was used within this tool (Hackett et al., 2008). To quantify phylogenetic signal in the microbiome, Pagel's lambda was estimated for the three alpha diversity metrics (observed ASVs (after rarefying), Shannon and Simpson), using phytools in R and the phylogeny of the species estimated from birdtree.org (Revell, 2012).

To compare two host traits (bird size and microbiome diversity) while controlling for phylogeny, Phylogenetic Generalized Least Squares (PGLS) was performed independently on each of the rarefied sample types. Average weight for each species was calculated using averaged field measurements of all birds sampled within each species in the entire dataset. For species that had no recorded weight, the average species weight in Birds of the World (Billerman, Keeney, Rodewald, & Schulenberg, 2020) was used. Following that, the samples were divided by body site and the samples were rarefied as described above. Alpha diversity metrics were calculated for each sample and then averaged for every species, resulting in one Observed, one Shannon, and one Simpson value for each species within each body site.

First, each alpha diversity metric was plotted against the average species weight without controlling for host phylogeny. Pearson's correlation coefficient, measuring the linear correlation between two variables, was used to determine the correlation between the two variables with a 95% confidence interval (Freedman, Pisani, & Purves, 2007). Next the same variables were tested against each other while controlling for phylogeny using a Phylogenetic Generalized Least Squares analysis in R using the package Geiger (Pennell et al., 2014).

Because the phylogenetic signal of bird size varies when analyzing size across orders and within orders (Harmon et al., 2010), each sample type was further subset to include only the largest bird order, Passeriformes. The new subsets were tested using the same methods as above.

RESULTS

Sequencing and Taxonomy

Following sample processing, sequencing, and quality control, a total of 1,740 16S rRNA amplicon sequence libraries were generated, with an average read depth of 65,595 reads per library (standard devion [SD]. +/- 47,441 reads). A total of 1,606 libraries were retained for analyses following rarefaction (Supplemental Table S1). All subsequent results were derived from rarefied data. Taxonomic composition across all body sites was dominated by Proteobacteria and Firmicutes, although relative abundance of these and other phyla differed between body sites (Fig. 2, Supplemental Table S2).

Diversity and Abundance

Three independent measures of alpha diversity (observed number of ASVs, the Shannon Diversity Index, and the Simpson Index) were largely congruent and revealed no significant differences in alpha diversity between body sites (Fig. 3). Samples from different body sites (i.e. sample types) were then combined to investigate beta diversity of the data as a whole (Fig. 4). For this combined dataset, samples were pruned and rarefied to 4,030 sequences resulting in a comparison among 606 blood samples, 353 buccal samples, 369 cloacal samples, 213 intestines samples, and 45 gizzard samples for a total of 1586 samples. Sample type significantly explained the variation of the microbiomes in both metrics tested (both p =0.001) but the variation explained varied: unweighted UniFrac R^2 = 6.7%, weighted UniFrac R^2 = 17%.

To assess differential abundance of microbial taxa between body sites, we conducted pairwise comparison of individual sites (excluding liver and spleen due to low sample counts), rarefying to the lower value of rarefaction between each of the two sample types being compared in each case. Overall, 56 genera were differentially abundant across all comparisons (Fig. 5, 6).

Pagel's Lambda

To assess whether our traits contained phylogenetic signals, Pagel's lambda was calculated for bird weight and for each alpha diversity metric in each dataset. Bird weight was always significantly different from 0 (p <0.0001) and lambda was between 0.9 and 1.0 (Supplemental Table S3). The exception was the Passeriformes-only gizzard dataset (lambda = 0.63, p =0.27), but we note that this dataset has extremely low sample size (N=9 individuals from 8 species).

Five microbiota diversity lambdas were significantly different from zero in the full dataset: Blood-observed, Blood-Shannon, Buccal-Shannon, Intestines-Shannon, Intestines-Simpson (Table 1). Three were significantly

different from zero in the Passeriformes-only dataset: Buccal-observed, Buccal-Shannon, Cloaca-observed (Table 1).

Phylogenetic Generalized Least Squares

Samples analyzed by sample type and each alpha diversity measurement (observed, Shannon, and Simpson) from all samples from the same species were averaged together to get a single set of measurements per host species. In the full dataset, the non-phylogenetically controlled regression found a single test to be marginally significant (cloaca, Simpson, p =0.048); all other tests had p >0.05 (Table 1, Supplemental Fig. S2). Using PGLS on the full dataset, all three alpha diversity metrics for the blood microbiota were significantly associated with bird mass (p <0.001); all other comparisons were not significant (p >0.17). When looking at the Passeriformes-only dataset, the non-phylogenetically controlled regression found a single test to be marginally significant (cloaca, Shannon, p =0.022); all other tests had$p$ >0.05 (Table 1, Supplemental Fig. S3). Using PGLS on the Passeriformes-only dataset, all three alpha diversity metrics for the cloacal microbiota were significantly associated with bird mass (p <0.05). Two of the gizzard microbiota tests were also significant (p <0.05) but we note that this dataset has extremely low sample size (N=9 individuals from 8 species). All other comparisons were not significant.

DISCUSSION

Biological understanding of vertebrates is incomplete without a thorough knowledge of the microbiome. One of the specific goals of evolutionary biology is to discover and describe biodiversity (see (Hird, 2017)); herein, 1740 samples from 779 wild birds expand what we know about both birds and their associated microbes. These samples, from a total of seven body sites, provide a fundamental description of the microbiomes of diverse and previously undescribed body sites in over 200 species. The cloacal and intestinal microbiome results support many previous studies that show a dominance of Proteobacteria and Firmicutes (Figure 2, Supplemental Table S2, reviewed in (Grond, Sandercock, Jumpponen, & Zeglin, 2018)).

The avian blood microbiome has only been characterized in chickens, where it was shown to be comprised of 60.58% Proteobacteria, 13.99% Bacteriodetes, 11.45% Firmicutes, 10.21% Actinobacteria, and 1.96% Cyanobacteria (Mandal et al., 2016). Our blood samples exhibited over 20% more Proteobacteria, with correspondingly smaller percentages of the remaining phyla (Fig. 2, Supplemental Table S2). The differences between the two findings may be unsurprising given that wild birds tend to have higher percentages of Proteobacteria (Grond et al., 2018). *Clostridium* was more abundant in the blood than in the cloaca, intestines, and buccal cavity (Fig. 6a, b, c) and is known to cause infections in birds (Crespo, Fisher, Shivaprasad, Fernández-Miyakawa, & Uzal, 2007). *Janthinobacterium* was more abundant in the blood than in the cloaca and buccal cavities (Fig. 6a, b). This genus has previously been identified in the cloaca of shorebirds (Santos et al., 2012). Known to contain pathogenic species, *Mycobacterium* was more abundant in the blood than in the cloaca, intestines, gizzard, and buccal cavity (Fig. 6a, b, c, d) , potentially showing that some birds were infected in this study or that non-pathogenic species are part of the healthy microbiota (Dhama et al., 2011). *Phenylobacterium* was also more common in the blood than in the cloaca, intestines, gizzard, and buccal cavity and has been previously shown to be in the uropygial glands of house sparrows with malaria (Fig. 6a, b, c, d) (Videvall et al., 2021).*Legionella* was more abundant in the blood than in the cloaca and the buccal cavity (Fig. 6a, b). This genus has been found in the cloaca of birds and is known to cause opportunistic infections in humans (Santos et al., 2012; Fields, Benson, & Besser, 2002). *Parvibaculum* (more abundant in the blood than in cloaca, intestines, gizzard, and buccal cavity and more abundant in the buccal cavity than in the cloaca (Fig. 6a, b, c, d, e)), *Planctomyces*(more abundant in the blood than in the cloaca, intestines, and buccal cavity (Fig. 6a, b, c)) and *Sediminibacterium* (more abundant in the blood than in the cloaca, intestines, gizzard, and buccal cavity (Fig. 6a, b, c, d)) were abundant in this study but are not commonly described members of the avian microbiota.

The buccal microbiome contained over 50% Proteobacteria and smaller percentages of Firmicutes, Actinobacteria, Tenericutes, and Bacteriodetes (Fig. 2, Supplemental Table S2). These results are similar to what was found in the oral microbiome of the Great Tit (Kropáčková et al., 2017) but are quite different from

those found in the Cooper's hawk, which contains higher relative abundance of Firmicutes than the quantities of Actinobacteria, Bacteriodetes, and Tenericutes (Taylor et al., 2019). Further sampling is required to determine how uniform the oral microbiome is in wild birds. *Hylemonella* was more abundant in the buccal cavity than in the blood, gizzard, intestines, and cloaca (Fig. 6a, e, f, g). This genus has been identified in the skin microbiome of vultures, raising the possibility of preening transferring skin and buccal microbes (Zepeda Mendoza et al., 2018). *Gallibacterium* was more abundant in the buccal cavity than in the blood, gizzard, intestines, and cloaca (Fig. 6a, e, f, g) and has been identified in the respiratory tract of healthy and unhealthy birds (Bisgaard, 1977), (Mushin, Weisman, & Singer, 1980). *Leucobacter* was more abundant in the buccal cavity than in the blood or cloaca in this dataset (Figure 6a, e) and has been identified in the feces of swiftlets (Sien, Lihan, Yee, Chuan, & Koon, 2013). *Rothia* was more abundant in the buccal cavity than in the gizzard or cloaca (Figure 6e, g) and is also a common member of the human saliva microbiome (Tsuzukibashi et al., 2017).

The gizzard samples were composed mostly of Firmicutes, Proteobacteria, and Cyanobacteria (Fig. 2, Supplemental Table S2). This is not consistent with the only published study on gizzard microbiomes that were mainly composed of Bacteroidetes, Cyanobacteria, Planctomycetes, Verrucomicrobia, and Alpha and Gammaproteobacteria (García-Amado et al., 2018). The liver and spleen microbiomes were both dominated by Firmicutes and Proteobacteria (Fig. 2, Supplemental Table S2); this is similar to the microbiome of wild mouse spleens which are also composed primarily of Firmicutes and Proteobacteria (Ge, Guo, Ge, Yin, & Yin, 2018). The relative abundances of these two phyla were very different in the liver and spleen, as compared to the blood. Spleens filter blood, so similarities between these two sample types may be expected. However, the vast differences between them show that these environments may be hospitable to different communities of bacteria, although this requires further confirmation, as our spleen and liver sample sizes were quite low.

The majority of our samples were of the blood, buccal, gizzard, intestines, and cloaca. These sample types were significantly different from each other in all three beta diversity metrics measured ($p < 0.001$) and the variation explained by each of these metrics was relatively high (6.7%- 17%). This shows that the types of taxa in each sample type are different from each other and that taxa those that are more phylogenetically distinct are more abundant. These measurements are lower than some other bird body site studies (Grond, Guilani, & Hird, 2020), but perhaps to be expected as this study incorporates dozens to hundreds of bird species that may have distinct microbiomes.

Many ornithologists are interested in studying the microbiome without harming the bird and therefore many have asked whether non-destructive sampling (e.g., oral swabs, cloacal swabs, feces) is adequate to describe the gut microbiome (Videvall, Strandh, Engelbrecht, Cloete, & Cornwallis, 2018). At the ASV level, our cloacal and intestine samples had substantial, and roughly equivalent, unique components; however, at the level of the sequencing reads, almost all the diversity was shared (Fig. 5b). Therefore, the unique ASVs contribute far less to the total microbiota than the shared ASVs. Compositionally, ten genera were more abundant in the cloaca than in the intestines: *Acinetobacter* , *Aerococcus* ,*Cloacibacterium* , *Cupriavidus* , *Limnohabitans* ,*Micrococcus* , *Propionibacterium* , *Rheinheimera* ,*Staphylococcus* , and *Stenotrophomas* (Figure 5a). Notably,*Cloacibacterium* has been found in avian blood (Mandal et al., 2016),*Micrococcus* in healthy conjunctiva and nasal passages (Silvanose, Bailey, Naldo, & Howlett, 2001), and *Staphylococcus* can cause infections in birds (Hermans, Devriese, De Herdt, Godard, & Haesebrouck, 2000). Six genera were more abundant in the intestines than in the cloaca: *Balneimonas* , *Enterococcus* ,*Lactobacillus* , *Lactococcus* , *Psychrobacter* , and*Rickettsiella* (Figure 5a). *Enterococcus* has been found in the cloaca of birds (Jørgensen et al., 2017) and*Psychrobacter* has been found in the throats and guts of birds (Kämpfer et al., 2015), (Kämpfer et al., 2020).*Lactobacillus* and *Lactococcus* have been previously identified in the cloaca of birds (Allegretti et al., 2014; Gunasekaran, Trabelcy, Izhaki, & Halpern, 2021), but we found them at higher abundances in the intestines, which is similar to previous comparative studies (Hird et al., 2015; Capunitan et al., 2020). *Enterococcus* was more abundant in the intestines than the cloaca, gizzard, blood, and buccal cavity and more common in the blood than the gizzard (Figs. 5a, 6c, d, f, i). This genus has been found in the cloaca of chickens (Jørgensen et al., 2017).*Lactobacillus* similarly was more common in the intestines than

the cloacal, gizzard, blood, and buccal cavity (Figs. 5a, 6c, f, i). This matched a finding in parrots that found *Lactobacillus* in their cloaca (Allegretti et al., 2014). In a comparison of body site microbiota in ostriches, several families that include our differentially abundant intestinal microbes were also significantly higher at internal gastrointestinal sites, as compared to the cloaca (*Lactobacillaceae* ,*Streptococcaceae* , *Enterococcaceae* ; Videvall et al., 2018).

The comparison of body sites also identified many additional taxa common to bird microbiomes and/or which are known pathogens:*Campylobacter*(Kapperud & Rosef, 1983; Hird et al., 2018),*Cloacibacterium*(Mandal et al., 2016).*Comamonas* (Kropáčková et al., 2017), *Enhydrobacter*(Kreisinger, Čížková, Kropáčková, & Albrecht, 2015), *Methylotenera*(Boukerb et al., 2021),*Pseudomonas* (Oprea, Crivineanu, Tudor, łOGOE, & Popa, 2010), *Psychrobacter*(Kämpfer et al., 2015), (Kämpfer et al., 2020).*Sphingobacterium*(Gunasekaran et al., 2021),*Streptococcus*(Devriese et al., 1994),*Curtobacterium*(Giorgio, De Bonis, Balestrieri, Rossi, & Guida, 2018), *Kocuria*(Braun, Wang, Zimmermann, Boutin, & Wink, 2018), *Brevundimonas*(Giorgio et al., 2018), Kingella (Foster et al., 2005), *Micrococcus*(Silvanose et al., 2001),*Staphylococcus*(Hermans et al., 2000),*Lactococcus*(Gunasekaran et al., 2021).*Methylobacterium* was found to be more abundant in the blood and buccal cavity more than the gizzard and more abundant in the buccal cavity than in the intestines and cloaca (Figure 6d, e, f, g). This is noteworthy as *methylobacterium* is known to be a contaminant in kits (Salter et al., 2014).

Although alpha diversity did not vary significantly across anatomical sites, we did identify specific microbial taxa that were differentially abundant between sites. (Fig. 3). Together with the beta diversity results, this shows that while the sample types contain similar levels of diversity, the composition of those communities is different across the body sites.

The microbiome is a trait of the host that may not be independent of the underlying phylogeny. Strictly speaking, phylogeny describes the evolutionary history of organisms. For a variety of reasons, phylogeny captures more than just evolutionary history and more closely related organisms frequently have more similar traits ("phylogenetic signal"). Therefore, comparisons of microbial "traits" across species need to control for phylogeny of the hosts.

Bird weight and one third of the alpha diversity measurements contained significant phylogenetic signal, as assessed by Pagel's Lambda (Table 1: two blood, one buccal, two intestine). When subsetting to include only Passeriformes, only three of the 15 tests contained significant phylogenetic signal: two buccal and one cloacal (although the three gizzard tests were likely affected by extremely low sample size). This shows an inconsistent or low level of association between phylogeny and microbiota richness and diversity, meaning that factors beyond phylogeny impact the microbial communities. This is similar to previous comparative work that found a White Noise model (of no phylogenetic signal) may fit the avian microbiome better than a neutral model or a model that includes selection (Capunitan et al., 2020).

How does size of a host influence the richness and diversity of the microbiota? Birds can be conceptualized as "islands" containing communities of microorganisms and their composition could potentially be driven by the Theory of Island Biogeography. We found eight significant correlations between bird size and microbiota, when using phylogenetic comparative methods (Table 1). In the full dataset, the blood samples show a significant negative correlation between their microbial diversity and the average host weight for all three of the diversity metrics ($p < 0.001$), indicating that larger birds actually have significantly lower diversity in their blood than smaller birds.

Because bird orders diversified quickly, there can be different associations of host traits within and across orders (Harmon et al., 2010). Therefore, we restricted analyses to a single order (Passeriformes) and reran the PGLS analyses. In the subset data, all three of the microbiome metrics for the cloacal samples exhibited significant negative correlations ($p < 0.05$) (Table 1, Supplemental Fig. S3). The Theory of Island Biogeography predicts that larger islands will house more diversity; our results appear to show the opposite: either a significant negative correlation or no correlation at all.

In broader terms, our results show that as host "islands" increase in size, the number of microbial taxa

immigrating decreases and/or the number of extinction events grows larger, indicating that regulatory mechanisms associated with host body size may be influencing the Species-Area Relationship. For example, avian microbiomes may be more prone to colonization by bacteria that inhibit conspecific proliferation, the immune systems of the host may be more sensitive to new "immigrant" members of the microbiome, different anatomical sites may be limited by nutrients, or other possibilities may be working to prevent a significant and large difference in alpha diversity due to host average weight. However, and importantly, even when the$p$ -values showed there was a negative or positive significant correlation between average host weight and alpha diversity, the value was small in every case. This suggests that there are only minor significant changes in the alpha diversity corresponding to average host weight. Taken together, our findings do not support an unexamined application of traditional Species-Area Relationships to the avian microbiome.

ACKNOWLEDGMENTS

We thank the Malawi Department of National Parks and Wildlife and staff of Vwaza Marsh Wildlife Reserve and Nyika National Park for logistical support. Original sampling and sequencing were supported by the Field Museum of Natural History's Emerging Pathogens Project, funded by the Davee Foundation and Dr. Ralph and Marian Falk Medical Research Trust. Sampling and research were conducted under permits from the Department of Forestry (License No. 3/12/ 2007/1, granted 10 July 2009, valid 10 July 2009–10 July 2010) and the Department of National Parks and Wildlife (Ref. No. NPW/2/1/12, granted 6 October 2009, valid October-November 2009). Staff of the Museums of Malawi, including Potiphar Kaliba, Christopher Salema, and Nyson Gawani, participated in the fieldwork and preparation of specimens. We also thank Field Museum of Natural History (Chicago, IL) staff and associates (current and former) for support of data collection, including Shannon Hackett, John Bates, Thomas Gnoske, Heather Skeen, Sarah Kurtis, and Zifeng Jiang. We thank Jason Weckstein for helpful conversations during the collection and analyses of microbiome data.

REFERENCES

Allegretti, L., Revolledo, L., Astolfi-Ferreira, C. S., Chacón, J. L., Martins, L. M., Seixas, G. H. F., & Ferreira, A. J. P. (2014). Isolation and molecular identification of lactic acid bacteria and Bifidobacterium spp. from faeces of the blue-fronted Amazon parrot in Brazil. *Beneficial Microbes,5*(4), 497–503. doi: 10.3920/BM2013.0082

Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Zech Xu, Z., ... Knight, R. (2017). Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems*, *2*(2). doi: 10.1128/mSystems.00191-16

Benson, A. K., Kelly, S. A., Legge, R., Ma, F., Low, S. J., Kim, J., ... Pomp, 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 of the United States of America*, *107*(44), 18933–18938. doi: 10.1073/pnas.1007028107

Billerman, S. M., Keeney, B. K., Rodewald, P. G., & Schulenberg, T. S. (2020). Birds of the World.*Cornell Laboratory of Ornithology, Ithaca, NY, USA*.

Bisgaard, M. (1977). Incidence of Pasteurella haemolytica in the respiratory tract of apparently healthy chickens and chickens with infectious bronchitis. characterisation of 213 strains. *Avian Pathology: Journal of the W.V.P.A*, *6*(4), 285–292. doi: 10.1080/03079457708418238

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., ... Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. doi: 10.1038/s41587-019-0209-9

Boukerb, A. M., Noël, C., Quenot, E., Cadiou, B., Chevé, J., Quintric, L., ... Gourmelon, M. (2021). Comparative Analysis of Fecal Microbiomes From Wild Waterbirds to Poultry, Cattle, Pigs, and Wastewater Treatment Plants for a Microbial Source Tracking Approach. *Frontiers in Microbiology,12*, 1851. doi: 10.3389/fmicb.2021.697553

Braun, M. S., Wang, E., Zimmermann, S., Boutin, S., & Wink, M. (2018). Kocuria uropygioeca sp. nov. and Kocuria uropygialis sp. nov., isolated from the preen glands of Great Spotted Woodpeckers (Dendrocopos major). *Systematic and Applied Microbiology*, *41*(1), 38–43. doi: 10.1016/j.syapm.2017.09.005

Capunitan, D. C., Johnson, O., Terrill, R. S., & Hird, S. M. (2020). Evolutionary signal in the gut microbiomes of 74 bird species from Equatorial Guinea. Molecular Ecology, 29(4), 829–847.

Crespo, R., Fisher, D. J., Shivaprasad, H. L., Fernández-Miyakawa, M. E., & Uzal, F. A. (2007). Toxinotypes of Clostridium perfringens isolated from sick and healthy avian species. *Journal of Veterinary Diagnostic Investigation: Official Publication of the American Association of Veterinary Laboratory Diagnosticians, Inc*, *19*(3), 329–333. doi: 10.1177/104063870701900321

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., ... Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, *72*(7), 5069–5072. doi: 10.1128/AEM.03006-05

Devriese, L. A., Haesebrouck, F., de Herdt, P., Dom, P., Ducatelle, R., Desmidt, M., ... Higgins, R. (1994). Streptococcus suis infections in birds. *Avian Pathology: Journal of the W.V.P.A*, *23*(4), 721–724. doi: 10.1080/03079459408419040

Dhama, K., Mahendran, M., Tiwari, R., Dayal Singh, S., Kumar, D., Singh, S., & Sawant, P. M. (2011). Tuberculosis in Birds: Insights into the Mycobacterium avium Infections. *Veterinary Medicine International*, *2011*, 712369. doi: 10.4061/2011/712369

Dickson, R. P., Erb-Downward, J. R., Freeman, C. M., McCloskey, L., Beck, J. M., Huffnagle, G. B., & Curtis, J. L. (2015). Spatial Variation in the Healthy Human Lung Microbiome and the Adapted Island Model of Lung Biogeography.*Annals of the American Thoracic Society*, *12*(6), 821–830. doi: 10.1513/AnnalsATS.201501-029OC

Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*,*125*(1), 1–15. doi: 10.1086/284325

Fields, B. S., Benson, R. F., & Besser, R. E. (2002). Legionella and Legionnaires' disease: 25 years of investigation. *Clinical Microbiology Reviews*, *15*(3), 506–526. doi: 10.1128/CMR.15.3.506-526.2002

Foster, G., Malnick, H., Lawson, P. A., Kirkwood, J., MacGregor, S. K., & Collins, M. D. (2005). Suttonella ornithocola sp. nov., from birds of the tit families, and emended description of the genus Suttonella. *International Journal of Systematic and Evolutionary Microbiology*, *55*(Pt 6), 2269–2272. doi: 10.1099/ijs.0.63681-0

Freedman, D., Pisani, R., & Purves, R. (2007). *Statistics: Fourth International Student Edition*. W.W. Norton & Company. Retrieved from https://play.google.com/store/books/details?id=vWeQDAAAQBAJ

García-Amado, M. A., Shin, H., Sanz, V., Lentino, M., Martínez, L. M., Contreras, M., ... Domínguez-Bello, M. G. (2018). Comparison of gizzard and intestinal microbiota of wild neotropical birds. *PloS One*, *13*(3), e0194857. doi: 10.1371/journal.pone.0194857

Ge, Y., Guo, G., Ge, B., Yin, H., & Yin, H. (2018). The spleen microbiota of small wild mammals reveals distinct patterns with tick-borne bacteria. *PLoS Neglected Tropical Diseases*, *12*(7), e0006499. doi: 10.1371/journal.pntd.0006499

Giorgio, A., De Bonis, S., Balestrieri, R., Rossi, G., & Guida, M. (2018). The Isolation and Identification of Bacteria on Feathers of Migratory Bird Species.*Microorganisms*, *6*(4). doi: 10.3390/microorganisms6040124

Grond, K., Guilani, H., & Hird, S. M. (2020). Spatial heterogeneity of the shorebird gastrointestinal microbiome. *Royal Society Open Science*,*7*(1), 191609. doi: 10.1098/rsos.191609

Grond, K., Sandercock, B. K., Jumpponen, A., & Zeglin, L. H. (2018). The avian gut microbiota: community, physiology and function in wild birds. *Journal of Avian Biology*, *49*(11), e01788. doi: 10.1111/jav.01788

Gunasekaran, M., Trabelcy, B., Izhaki, I., & Halpern, M. (2021). Direct Evidence That Sunbirds' Gut Microbiota Degrades Floral Nectar's Toxic Alkaloids. *Frontiers in Microbiology*, *12*, 639808. doi: 10.3389/fmicb.2021.639808

Hackett, S. J., Kimball, R. T., Reddy, S., Bowie, R. C. K., Braun, E. L., Braun, M. J., ... Yuri, T. (2008). A phylogenomic study of birds reveals their evolutionary history. *Science*, *320*(5884), 1763–1768. doi: 10.1126/science.1157704

Harmon, L. J., Losos, J. B., Jonathan Davies, T., Gillespie, R. G., Gittleman, J. L., Bryan Jennings, W., ... Mooers, A. Ø. (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolution; International Journal of Organic Evolution*, *64*(8), 2385–2396. doi: 10.1111/j.1558-5646.2010.01025.x

[dataset] Herder E, Lutz H, Hird SM; 2022; Comparative Microbiota from Malawian Birds. Deposition on Dryad, Qiita, and EBI pending. DOI: pending.

Hermans, K., Devriese, L. A., De Herdt, P., Godard, C., & Haesebrouck, F. (2000). Staphylococcus aureus infections in psittacine birds. *Avian Pathology: Journal of the W.V.P.A*, *29*(5), 411–415. doi: 10.1080/030794500750047153

Hird, S. M., Sánchez, C., Carstens, B. C., & Brumfield, R. T. (2015). Comparative Gut Microbiota of 59 Neotropical Bird Species. Frontiers in Microbiology, 6. doi: 10.3389/fmicb.2015.01403

Hird, S. M. (2017). Evolutionary Biology Needs Wild Microbiomes. *Frontiers in Microbiology*, *8*, 725. doi: 10.3389/fmicb.2017.00725

Hird, S. M., Ganz, H., Eisen, J. A., & Boyce, W. M. (2018). The Cloacal Microbiome of Five Wild Duck Species Varies by Species and Influenza A Virus Infection Status. mSphere, 3(5). doi: 10.1128/mSphere.00382-18

Hulsen, T. (2021). BioVenn – an R and Python package for the comparison and visualization of biological lists using area-proportional Venn diagrams. *Data Science*, Vol. 4, pp. 51–61. doi: 10.3233/ds-210032

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, *491*(7424), 444–448. doi: 10.1038/nature11631

Jetz, W., Thomas, G. H., Joy, J. B., Redding, D. W., Hartmann, K., & Mooers, A. O. (2014). Global distribution and conservation of evolutionary distinctness in birds. *Current Biology: CB*, *24*(9), 919–930. doi: 10.1016/j.cub.2014.03.011

Jørgensen, S. L., Poulsen, L. L., Thorndal, L., Ronaghinia, A. A., Bisgaard, M., & Christensen, H. (2017). Characterization of Enterococcus faecalis isolated from the cloaca of "fancy breeds" and confined chickens. *Journal of Applied Microbiology*, *122*(5), 1149–1158. doi: 10.1111/jam.13416

Kämpfer, P., Glaeser, S. P., Irgang, R., Fernández-Negrete, G., Poblete-Morales, M., Fuentes-Messina, D., ... Avendaño-Herrera, R. (2020). Psychrobacter pygoscelis sp. nov. isolated from the penguin Pygoscelis papua. *International Journal of Systematic and Evolutionary Microbiology*, *70*(1), 211–219. doi: 10.1099/ijsem.0.003739

Kämpfer, P., Jerzak, L., Wilharm, G., Golke, J., Busse, H.-J., & Glaeser, S. P. (2015). Psychrobacter ciconiae sp. nov., isolated from white storks (Ciconia ciconia). *International Journal of Systematic and Evolutionary Microbiology*, *65*(Pt 3), 772–777. doi: 10.1099/ijs.0.000013

Kapperud, G., & Rosef, O. (1983). Avian wildlife reservoir of Campylobacter fetus subsp. jejuni, Yersinia spp., and Salmonella spp. in Norway. *Applied and Environmental Microbiology*, *45*(2), 375–380. doi: 10.1128/aem.45.2.375-380.1983

Karl, J. P., Hatch, A. M., Arcidiacono, S. M., Pearce, S. C., Pantoja-Feliciano, I. G., Doherty, L. A., & Soares, J. W. (2018). Effects of Psychological, Environmental and Physical Stressors on the Gut Microbiota. *Frontiers in Microbiology*, *9*, 2013. doi: 10.3389/fmicb.2018.02013

Kodera, S., Das, P., Gilbert, J. A., Lutz, H. L. IN PRESS. Conceptual strategies for characterizing interactions in microbial communities.*iScience* .

Koskella, B., Hall, L. J., & Metcalf, C. J. E. (2017). The microbiome beyond the horizon of ecological and evolutionary theory. *Nature Ecology & Evolution*,*1*(11), 1606–1615. doi: 10.1038/s41559-017-0340-2

Kreisinger, J., Čížková, D., Kropáčková, L., & Albrecht, T. (2015). Cloacal Microbiome Structure in a Long-Distance Migratory Bird Assessed Using Deep 16sRNA Pyrosequencing. *PLOS ONE*, Vol. 10, p. e0137401. doi: 10.1371/journal.pone.0137401

Kropáčková, L., Pechmanová, H., Vinkler, M., Svobodová, J., Velová, H., Těšičký, M., ... Kreisinger, J. (2017). Variation between the oral and faecal microbiota in a free-living passerine bird, the great tit (Parus major). *PloS One*, *12*(6), e0179945. doi: 10.1371/journal.pone.0179945

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., ... Gonzalez, A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, Vol. 7, pp. 601–613. doi: 10.1111/j.1461-0248.2004.00608.x

Li, L., & Ma, Z. (sam). (2016). Testing the Neutral Theory of Biodiversity with Human Microbiome Datasets. *Scientific Reports*, Vol. 6. doi: 10.1038/srep31448

Li, S.-P., Wang, P., Chen, Y., Wilson, M. C., Yang, X., Ma, C., ... Jiang, L. (2020). Island biogeography of soil bacteria and fungi: similar patterns, but different mechanisms. *The ISME Journal*, *14*(7), 1886–1896. doi: 10.1038/s41396-020-0657-8

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. doi: 10.1186/s13059-014-0550-8

Lozupone, C., & Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*,*71*(12), 8228–8235. doi: 10.1128/AEM.71.12.8228-8235.2005

Lutz, H. L., Hochachka, W. M., Engel, J. I., Bell, J. A., Tkach, V. V., Bates, J. M., ... Weckstein, J. D. (2015). Parasite prevalence corresponds to host life history in a diverse assemblage of afrotropical birds and haemosporidian parasites. *PloS One*, *10*(4), e0121254. doi: 10.1371/journal.pone.0121254

MacArthur, R. H., & Wilson, E. O. (2001). *The Theory of Island Biogeography*. Princeton University Press. Retrieved from https://play.google.com/store/books/details?id=a10cdkywhVgC

Mandal, R. K., Jiang, T., Al-Rubaye, A. A., Rhoads, D. D., Wideman, R. F., Zhao, J., ... Kwon, Y. M. (2016). An investigation into blood microbiota and its potential association with Bacterial Chondronecrosis with Osteomyelitis (BCO) in Broilers. *Scientific Reports*, *6*, 25882. doi: 10.1038/srep25882

Ma, Z., & Li, L. (2018). Semen Microbiome Biogeography: An Analysis Based on a Chinese Population Study. *Frontiers in Microbiology*, *9*, 3333. doi: 10.3389/fmicb.2018.03333

McArdle, B. H., & Anderson, M. J. (2001). Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology*, *82*(1), 290–297.

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. doi: 10.1371/journal.pone.0061217

Miller, E. T., & Bohannan, B. J. M. (2019). Life Between Patches: Incorporating Microbiome Biology Alters the Predictions of Metacommunity Models. *Frontiers in Ecology and Evolution*, *7*, 276. doi: 10.3389/fevo.2019.00276

Miller, E. T., Svanbäck, R., & Bohannan, B. J. M. (2018). Microbiomes as Metacommunities: Understanding Host-Associated Microbes through Metacommunity Ecology. *Trends in Ecology & Evolution*, *33*(12), 926–935. doi: 10.1016/j.tree.2018.09.002

Mushin, R., Weisman, Y., & Singer, N. (1980). Pasteurella haemolytica Found in the Respiratory Tract of Fowl. *Avian Diseases*, *24*(1), 162–168. doi: 10.2307/1589775

Oksanen, J., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., ... Others. (2018). Vegan: community ecology package. *R Package Version*,*2*(6).

Oprea, A., Crivineanu, M., Tudor, L., łOGOE, I., & Popa, R. (2010). PSEUDOMONAS AERUGINOSA INFECTIONS IN A PET PARROTS. *LUCRĂRI ŞTIINłIFICE MEDICINĂ VETERINARĂ*, *XLIII*(1), 86–89.

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., ... Harmon, L. J. (2014). geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees.*Bioinformatics* , *30*(15), 2216–2218. doi: 10.1093/bioinformatics/btu181

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*, 217–223. doi: 10.1111/j.2041-210X.2011.00169.x

Ruhs, E. C., Martin, L. B., & Downs, C. J. (2020). The impacts of body mass on immune cell concentrations in birds. *Proceedings. Biological Sciences / The Royal Society*, *287*(1934), 20200655. doi: 10.1098/rspb.2020.0655

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., ... Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, *12*, 87. doi: 10.1186/s12915-014-0087-z

Santos, S. S., Pardal, S., Proença, D. N., Lopes, R. J., Ramos, J. A., Mendes, L., & Morais, P. V. (2012). Diversity of cloacal microbial community in migratory shorebirds that use the Tagus estuary as stopover habitat and their potential to harbor and disperse pathogenic microorganisms. *FEMS Microbiology Ecology*, *82*(1), 63–74. Retrieved from https://academic.oup.com/femsec/article-abstract/82/1/63/567606

Sekelja, M., Rud, I., Knutsen, S. H., Denstadli, V., Westereng, B., Næs, T., & Rudi, K. (2012). Abrupt temporal fluctuations in the chicken cloacal microbiota are explained by its gastrointestinal origin. *Applied and Environmental Microbiology*, *78*(8), 2941–2948. doi: 10.1128/AEM.05391-11

Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. University of Illinois. *Urbana* , *117*.

Sherrill-Mix, S., McCormick, K., Lauder, A., Bailey, A., Zimmerman, L., Li, Y., ... Bushman, F. D. (2018). Allometry and Ecology of the Bilaterian Gut Microbiome.*mBio*, *9*(2). doi: 10.1128/mBio.00319-18

Sien, L. S., Lihan, S., Yee, L. T., Chuan, C. H., & Koon, L. C. (2013). Isolation and Characterization of Antibiotic Resistant Bacteria from Swiftlet Feces in Swiftlet Farm Houses in Sarawak, Malaysia. *Microbiology Indonesia*, *7*(4), 1–1. doi: 10.5454/mi.7.4.1

Silvanose, C. D., Bailey, T. A., Naldo, J. L., & Howlett, J. C. (2001). Bacterial flora of the conjunctiva and nasal cavity in normal and diseased captive bustards.*Avian Diseases*, *45*(2), 447–451. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11417826

Simpson, E. H. (1949). Measurement of Diversity. *Nature*, *163*(4148), 688–688. doi: 10.1038/163688a0

Song, S. J., Sanders, J. G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M. W., ... Knight, R. (2020). Comparative Analyses of Vertebrate Gut Microbiomes Reveal Convergence between Birds and Bats. *mBio*, *11*(1). doi: 10.1128/mBio.02901-19

Taylor, M. J., Mannan, R. W., U'Ren, J. M., Garber, N. P., Gallery, R. E., & Arnold, A. E. (2019). Age-related variation in the oral microbiome of urban Cooper's hawks (Accipiter cooperii). *BMC Microbiology*, *19*(1), 47. doi: 10.1186/s12866-019-1413-y

Team, R. D. C. (2021). *R Programming. R Development Core Team.*

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., ... Earth Microbiome Project Consortium. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463. doi: 10.1038/nature24621

Tsuzukibashi, O., Uchibori, S., Kobayashi, T., Umezawa, K., Mashimo, C., Nambu, T., ... Ochiai, T. (2017). Isolation and identification methods of Rothia species in oral cavities. *Journal of Microbiological Methods*, *134*, 21–26. doi: 10.1016/j.mimet.2017.01.005

Venkataraman, A., Bassis, C. M., Beck, J. M., Young, V. B., Curtis, J. L., Huffnagle, G. B., & Schmidt, T. M. (2015). Application of a neutral community model to assess structuring of the human lung microbiome. *mBio*,*6*(1). doi: 10.1128/mBio.02284-14

Videvall, E., Marzal, A., Magallanes, S., Fleischer, R. C., Espinoza, K., & García-Longoria, L. (2021). The uropygial gland microbiome of house sparrows with malaria infection. *Journal of Avian Biology*, *52*(2). doi: 10.1111/jav.02686

Videvall, E., Strandh, M., Engelbrecht, A., Cloete, S., & Cornwallis, C. K. (2018). Measuring the gut microbiome in birds: Comparison of faecal and cloacal sampling.*Molecular Ecology Resources*, *18*(3), 424–434. doi: 10.1111/1755-0998.12744

Wickham, H. (2009).*ggplot2: Elegant Graphics for Data Analysis*. Springer Science & Business Media. Retrieved from https://play.google.com/store/books/details?id=bes-AAAAQBAJ

Zepeda Mendoza, M. L., Roggenbuck, M., Manzano Vargas, K., Hansen, L. H., Brunak, S., Gilbert, M. T. P., & Sicheritz-Pontén, T. (2018). Protective role of the vulture facial skin and gut microbiomes aid adaptation to scavenging. *Acta Veterinaria Scandinavica*, *60*(1), 61. doi: 10.1186/s13028-018-0415-3

DATA ACCESSIBILITY

For the cloacal and intestinal data: 16S rRNA sequences are publicly available via the QIITA platform under study identifiers (ID) 11166 (httpa://qiita.ucsd.edu/study/description/11166) and #### TBD (TBD pending publication), and the European Bioinformatics Institute (EBI) under accession numbers PRJEB35449 and #### TBD (TBD pending publication). For all other body site data: Qiita and EBI (or SRA) submission are pending and will be provided prior to publication. Metadata files associated with sequence data are available in this study in Supplemental Table 1. All sequence and metadata are also available per request.

BENEFITS SHARING STATEMENT

Benefits generated from this research accrue from the sharing of our data and results on public databases as described above. Physical specimens have been accessioned in biological repositories at the Field Museum of Natural History and the Museums of Malawi through which they are accessible for scientific research.

AUTHOR CONTRIBUTIONS

Designed research (SH, EH); collected data (HL); analyzed data (EH, HL); interpreted results (EH, SH, HL); wrote manuscript (EH, SH); Approved and edited manuscript (EH, SH, HL).

TABLES AND FIGURES

Table 1. Regression, Phylogenetic Generalized Least Squares (PGLS) and Pagel's Lambda (and significance) for the various body sites using all data (left) and a subset of data, the Passeriformes only.

	Full Dataset						Passeriformes Only					
	non-PGLS		PGLS		Lambda		non-PGLS		PGLS		Lambda	
	R	*p*	Value	*p*	λ	*p*	R	*p*	Value	*p*	λ	*p*
	Blood (N=606 individuals, 187 species)						**Blood** (N=505 individuals, 130 species)					
Observed	-0.0025	0.97	**-0.0451**	**0.000**	**0.2939**	**0.032**	-0.039	0.66	-0.0062	0.861	0.0001	1.000
Shannon	-0.0058	0.94	**-0.0009**	**0.000**	**0.2538**	**0.019**	0.03	0.74	-0.0002	0.749	0.0001	1.000
Simpson	-0.046	0.53	**-0.0003**	**0.000**	0.0411	0.606182	0.053	0.55	-0.0001	0.688	0.0000	1.000
	Buccal (N=346 individuals, 132 species)						**Buccal** (N=272 individuals, 86 species)					
Observed	-0.048	0.59	-0.0231	0.622	0.2283	0.227	0.13	0.23	0.6110	0.362	**0.9192**	**0.001**
Shannon	-0.082	0.35	-0.0004	0.177	**0.5299**	**0.000**	0.19	0.083	0.0050	0.442	**0.3331**	**0.021**
Simpson	0.019	0.83	0.0000	0.764	0.3625	0.123	0.091	0.41	0.0002	0.834	0.0001	1.000
	Gizzard (43 individuals, 32 species)						**Gizzard** (N=9 individuals, 8 species)					
Observed	-0.24	0.19	-0.0416	0.437	0.0001	1.000	-0.26	0.53	0.2103	0.178	0.0001	1.000
Shannon	-0.33	0.061	-0.0004	0.237	0.2410	0.583	-0.026	0.95	**0.0028**	**0.025**	0.0001	1.000
Simpson	-0.3	0.093	0.0000	0.681	0.0001	1.000	0.23	0.5	**0.0011**	**0.008**	0.0001	1.000
	Intestine (N=200 individuals, 115 species)						**Intestine** (N=151 individuals, 79 species)					
Observed	-0.1	0.28	-0.0297	0.692	0.3158	0.161	-0.032	0.78	-0.0140	0.956	0.0001	1.000
Shannon	-0.066	0.48	0.0029	0.705	**0.3506**	**0.001**	-0.026	0.82	-0.0010	0.718	0.0001	1.000
Simpson	-0.014	0.88	0.0020	0.286	**0.2668**	**0.004**	-0.0013	0.99	-0.0001	0.889	0.0497	1.000
	Cloaca (N=353 samples, 133 species)						**Cloaca** (N=275 samples, 85 species)					
Observed	0.016	0.86	0.0014	0.932	0.3393	0.077	-0.048	0.66	**-0.9538**	**0.044**	**0.5819**	**0.022**
Shannon	-0.15	0.082	-0.0003	0.225	0.2786	0.197	**-0.25**	**0.022**	**-0.0425**	**0.000**	0.2055	0.185
Simpson	**-0.17**	**0.048**	-0.0001	0.263	0.1654	1.000	-0.2	0.062	**-0.0078**	**0.000**	0.1153	0.534

Hosted file

`image2.emf` available at https://authorea.com/users/457250/articles/554075-body-size-poorly-predicts-host-associated-microbial-diversity-in-wild-birds

Figure 1: Phylogeny of all of the bird species represented in this study (left) and a table with gray cells to indicate which sample types were sampled from each bird species (right).

Hosted file

`image3.emf` available at https://authorea.com/users/457250/articles/554075-body-size-poorly-predicts-host-associated-microbial-diversity-in-wild-birds

Figure 2: Phylum level bar plots of all samples grouped by body sites ordered by sample type and, within that, by average species size.

Hosted file

`image4.emf` available at https://authorea.com/users/457250/articles/554075-body-size-poorly-predicts-host-associated-microbial-diversity-in-wild-birds

Figure 3: Box plots of Alpha Diversity metrics (A) Observed ASV, (B) Shannon, and (C) Simpson for all samples used in this study and separated by Body Site. All samples were rarefied to 4,000 for comparison among the samples.

Hosted file

`image5.emf` available at https://authorea.com/users/457250/articles/554075-body-size-poorly-predicts-host-associated-microbial-diversity-in-wild-birds

Figure 4: Non-metric Multidimensional Scaling Ordination constructed from unweighted UniFrac and weighted UniFrac matrices of wild bird microbiomes collected from Malawi. Colors represent the five body sites sampled; results from the adonis tests for significance of body sites are shown.

Hosted file

`image6.emf` available at https://authorea.com/users/457250/articles/554075-body-size-poorly-predicts-host-associated-microbial-diversity-in-wild-birds

Figure 5: Comparison of cloacal and intestine samples after rarefaction to 9,000 sequences per sample. A. DESeq2 Plot showing differential abundance of ASVs, alpha=0.01, B. Venn diagram of shared and unique ASVs (left) and sequencing reads (right).

Hosted file

`image7.emf` available at https://authorea.com/users/457250/articles/554075-body-size-poorly-predicts-host-associated-microbial-diversity-in-wild-birds

Figure 6: Differential abundance of ASVs between sample types using an alpha of 0.01: A. Blood and buccal, rarefied to 4,000 sequences. B. Blood and cloaca, rarefied to 4,000 sequences. C. Blood and intestine, rarefied to 4,000. D. Blood and gizzard, rarefied to 4,000 sequences. E. Cloaca and buccal, rarefied to 10,000 sequences. F. Intestine and buccal, rarefied to 9,000 sequences. G. Gizzard and buccal, rarefied to 10,000 sequences. H. Cloaca and gizzard, rarefied to 9,000 sequences. I. Intestine and gizzard, rarefied to 9,000.

SUPPLEMENTAL MATERIAL

Figure S1. Rarefaction curves for each sample type with red dotted line indicating the number of sequences to which the samples were rarefied: A. blood, B. buccal, C. cloaca, D. intestine, E. gizzard, F. liver, G. spleen.

Figure S2. Correlations between microbiota alpha diversity and average bird weight using Observed (upper left), Shannon (upper right), and Simpson diversity (bottom left) metrics; correlation and significance values with and without controlling for phylogeny are also reported (bottom right) for the (A) blood, (B) buccal, (C) cloaca, (D) intestine, (E) gizzard body sites for all samples.

Figure S3. Correlations between microbiota alpha diversity and average bird weight using Observed (upper left), Shannon (upper right), and Simpson diversity (bottom left) metrics; correlation and significance values (in green) with and without controlling for phylogeny are also reported (bottom right) for the (A) blood, (B) buccal, (C) cloaca, (D) intestine, (E) gizzard body sites for only samples belonging to the order Passeriformes.

Table S1 . Metadata for all samples, including FMNH#, species, body site and raw reads.

Table S2. Relative abundance of bacterial phyla for the seven body sites. Phyla that were represented at below 10% of the sequences within the sample were placed in the "Other" category. Samples are separated by sample type.

Table S3. Pagel's lambda (and significance) for bird weight, using the species in each of the body site specific datasets.