DR. RYAN S TERRILL (Orcid ID : 0000-0003-3834-9761)
DR. SARAH M HIRD (Orcid ID : 0000-0002-1998-9387)

Article type : Original Article

Evolutionary signal in the gut microbiomes of 74 bird species from Equatorial Guinea

Darien C. Capunitan[1*], Oscar Johnson[2], Ryan S. Terrill[2,3], Sarah M. Hird[1,4*†]

Affiliations

[1] Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA

[2] Museum of Natural Science, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA, USA

[3] Moore Lab of Zoology, Occidental College, Los Angeles, CA, USA

[4] Institute for Systems Genomics, University of Connecticut, Storrs, CT, USA

*These authors contributed equally to the work.

†Correspondence: Dr. Sarah Hird, sarah.hird@uconn.edu; ORCID 0000-0002-1998-9387

Running title: Gut Microbiomes of 74 Equatoguinean Birds

Keywords: Microbiomes, microbiota, birds, gut, comparative analysis, evolution

ABSTRACT

How the microbiome interacts with hosts across evolutionary time is poorly understood. Datasets including many host species are required to conduct comparative analyses. Here, we have analyzed 142 intestinal microbiome samples from 92 birds belonging to 74 species from Equatorial Guinea, using the 16S rRNA gene. Using four definitions for microbial taxonomic units (97%OTU, 99%OTU, 99%OTU with singletons removed, ASV), we conducted alpha and beta diversity analyses. We find that raw abundances and diversity varied between the datasets but relative patterns were largely consistent across datasets. Host taxonomy, diet and locality were significantly associated with microbiomes, at generally similar levels using three distance metrics. Phylogenetic comparative methods assessed the evolutionary relationship between the microbiome as a trait of a host species and the underlying bird phylogeny. Using multiple ways of defining "microbiome traits", we find that a neutral Brownian motion model does not explain variation in microbiomes. Instead, we find a White Noise model (indicating little phylogenetic signal), is most likely. There was some support for the Ornstein-Uhlenbeck model (that invokes selection), but the level of support was similar to that of a White Noise simulation, further supporting the White Noise model as the best explanation for the evolution of the microbiome as a trait of avian hosts. Our study demonstrates that both environment and evolution play a role in the gut microbiome and the relationship does not follow a neutral model; these biological results are qualitatively robust to analytical choices.

INTRODUCTION

A microbiome is the collection of microorganisms that live in a particular environment (Whipps, Lewis, & Cooke, 1988). The vertebrate gut microbiome can be involved in host health, physiology, development, and adaptation (e.g., Al-Asmakh et al., 2014; Alberdi, Aizpurua, Bohmann, Zepeda-Mendoza, & Gilbert, 2016; Cho & Blaser, 2012; Diaz Heijtz et al., 2011; Kohl, Varner, Wilkening, & Dearing, 2018). While we are beginning to appreciate the importance of the microbiome in a diversity of host taxa (Engel & Moran, 2013; Givens, Ransom, Bano, & Hollibaugh, 2015; Hird, Sánchez, Carstens, & Brumfield, 2015; Spor, Koren, & Ley, 2011; Wang, Brelsfoard, Wu, & Aksoy, 2013), we lack microbiome information for the vast majority of host species and have a poor understanding of the evolutionary and ecological processes relevant to microbiome composition and function, especially in wild organisms and non-model taxa.

Most of our knowledge of gut microbiomes comes from studies on mammals (Colston & Jackson, 2016). However, mammals have adaptations that affect their microbiomes, specifically live birth through the vagina (except for Monotremes) and milk (Cho & Blaser, 2012; Dominguez-Bello et al., 2010). Birds are a globally

distributed class of organisms that comprise ~15% of vertebrate species, yet in 2015 wild birds comprised only 0.6% of the vertebrate microbiome literature (Colston & Jackson, 2016). Birds have their own set of adaptations that could affect and/or be affected by the microbiome. For example, all birds lay eggs that hatch in nests; nests range from minimalistic (e.g., slight depressions on the ground, bare rocks) to elaborate (e.g., the bowers of bowerbirds, communal nests of weavers). The nest, eggs, nestlings and parents form a complex and unique microbial network: nest building influences parental feather microbiome (Kilgas, Saag, Mägi, Tilgar, & Mänd, 2012) and nest materials influence eggshell microbiomes and nestling development (Jacob et al., 2015). Cross fostering experiments show that the nest can significantly affect nestling microbiome (Teyssier, Lens, Matthysen, & White, 2018). Parental care by birds also regulates microbes: egg incubation influences bacterial pathogen load on eggs (Brandl et al., 2014); nestling fecal sacs are coated in a mucus that allows parents to easily remove nestling feces and prevents contamination of the nest with harmful bacteria (Ibáñez-Álamo, Ruiz-Rodríguez, & Soler, 2014) and parental saliva (during feeding) can vertically transmit essential microbes to offspring (Kyle & Kyle, 2004). Many notable bird adaptations are also related to the microbiome. The low diversity of vultures' gut microbiomes may allow them to competitively exclude pathogens residing on their carrion-based diet (Roggenbuck et al., 2014). The folivorous diet of the hoatzin is facilitated by an enlarged crop that houses bacterial communities that digest leaves in a manner that is convergent to a cow's rumen (Godoy-Vitorino et al., 2012, 2008). Birds exhibit immense diversity in their diets, behaviors, and ecologies. Unsurprisingly, they also exhibit diversity in their microbiomes. Bird microbiomes appear to be dominated by a small number of bacterial phyla, namely Proteobacteria, Firmicutes, and Bacteroidetes (Grond, Sandercock, Jumpponen, & Zeglin, 2018; Hird et al., 2015; Waite & Taylor, 2014) and are influenced by environment (Wienemann et al., 2011), geographic location (Hird, Carstens, Cardiff, Dittmann, & Brumfield, 2014), diet (Rubio et al., 1998), age (Grond, Lanctot, Jumpponen, & Sandercock, 2017; Teyssier et al., 2018; Van Dongen et al., 2013), and host genetics (Kropáčková et al., 2017). Despite this, little is known about the role of the microbiome in avian evolution.

Understanding the role of the microbiome across host evolution is a difficult task. Phylosymbiosis has been proposed as an eco-evolutionary test for congruence between host species and microbiome divergence (Brooks, Kohl, Brucker, van Opstal, & Bordenstein, 2016). However, phylosymbiosis is agnostic to mechanism. Additional methods exist for determining the relationship between host phylogeny and the microbiome (e.g., using simulations in Mazel et al., 2018; Mantel tests in Pollock et al., 2018; multiple regression tests in Youngblut et al., 2019), however the evolutionary processes that generate the microbiome have been relatively unexplored. Here, we propose utilizing the microbiome as a trait of a host (Benson, 2016; Benson et al., 2010),

and extrapolating that trait to host species. The microbiome could be defined as a host trait in many ways, as this is a complex community that can be summarized using many diversity and compositional measures (e.g., relative or absolute abundance of bacterial taxa, alpha diversity metrics, relative or absolute abundance of functional gene pathways). For this study, we use nine microbiome traits: the average relative abundance of the six most abundant bacterial phyla and one measure of alpha diversity and two measures of beta diversity. Importantly, many traits are not independent of phylogeny and this relationship must be accounted for to understand trait evolution (Felsenstein, 1985). Since Felsenstein published the phylogenetic comparative method (PCM), many models have been proposed for the evolution of traits on phylogenies. Comparison of the fit of these different models of evolution for different traits (Pennell & Harmon, 2013) can allow for statistical comparison of phylogenetically non-independent traits, as well as give insight into the patterns of evolution of traits. Trait values at the tips of a phylogeny are used to infer the state of the most recent common ancestor. Models of evolution are then applied to that initial trait value across the length of the tree to predict trait values at the tips under different models. Comparison between the expected values and the actual values provides support for a model, thereby indicating which evolutionary process(es) may be responsible for the observed patterns.

The major evolutionary processes we were interested in distinguishing were neutrality and selection. We tested four different models to study host-microbiome evolution: Brownian motion, Ornstein-Uhlenbeck, Early Burst, and White Noise. Brownian motion models are frequently used as neutral models of trait evolution; conceptually, they are a "random walk" that allows a trait to be influenced by many small, random forces. A neutral model will explain a trait's contemporary values on a phylogenetic tree when it has diverged at a rate consistent with random mutations and drift only; for example, body size within many clades of birds fits a neutral model best (Harmon et al., 2010). The Ornstein-Uhlenbeck model (Uhlenbeck & Ornstein, 1930) is also a "random walk" model but with some degree of pull towards a local optimum or multiple optima (i.e., selection). This model has found support in the study of quantitative traits and gene expression because it incorporates both drift and selection (Rohlfs, Harrigan, & Nielsen, 2014). Early Burst models test for an early burst of rapid evolution that slows over the length of a tree as niches become filled, akin to adaptive radiation (Harmon et al., 2010). This model may be supported by microbiome data, as microbial taxa within gut microbiomes display phylogenetic patterns consistent with adaptive radiations (Ley, Peterson, & Gordon, 2006). The last model we test is the White Noise model wherein the contemporary values of a trait appear to be a random draw of trait values with no relationship to phylogeny. This could be the case when a trait evolves very quickly (i.e., along the branches of the phylogeny) or when a trait is influenced strongly by the

environment. Model fit can be assessed using the Akaike information criterion, or AIC (Akaike, 1974), a statistical means of model selection. We also calculate Pagel's lambda (**λ**) (Pagel, 1999) as a measure of how well the trait values adhere to the neutral expectation of evolution along a given phylogeny. A **λ** of 0 means the phylogeny has to be reduced to a star phylogeny to fit the distribution of traits (because the data contain no phylogenetic signal). A **λ** of 1 means the phylogeny fits the trait data perfectly (because the data have been neutrally evolving along the phylogeny).

Describing the microbiome of wild species is a basic biodiversity goal in itself and can provide essential information for evolutionary biology (Hird, 2017). Equatorial Guinea is comprised of two distinct Afrotropical biomes; the mainland is part of the biodiverse Lower Guinean Forest and is contiguous with adjacent Congolese Basin forests, while Bioko Island is a continental island of the Cameroon Mountain Chain (Borrow & Demey 2014). The avifauna of the country was documented by early European explorers of western Africa (Strickland 1844; Alexander 1903), but subsequently received little attention until recently (Pérez del Val 1994; Cooper et al. 2016, 2017). The lowland avifauna of Bioko Island is broadly similar to that of mainland Equatorial Guinea though less diverse, while the montane avifauna is more closely allied to that of the mainland Cameroon Mountains (Borrow & Demey 2014). The isolated nature of the Bioko highlands has resulted in a high level of endemism that is less than that of more distant islands such as São Tomé and Príncipe (Borrow & Demey 2014). The relatively recent isolation means that much of the species divergence is at the subspecies level. The endemism, combined with the small size of Bioko Island means that the few endemic species, such as Bioko Speirops (*Zosterops brunneus*), are highly range-restricted and thus vulnerable to extinction (IUCN 2018). Focusing on diverse, tropical communities of birds allows us to add microbiome data to both widespread and range restricted branches on the avian tree of life.

Here, we characterize the gut microbiome of 92 individuals from 74 Equatoguinean bird species using 16S rRNA gene sequencing. We define the bacterial taxonomic units in the dataset in four different ways to determine the impact of these methods on biological conclusions. The first two ways use a genetic similarity cutoff to define Operational Taxonomic Units (OTUs) at 97% and 99% similarity; the third way is the 99% OTU dataset with all singletons removed (hereafter "99%OTU(NS)" for No Singletons). The fourth way uses unique sequences as the OTUs, a technique also known as Amplicon Sequence Variants (ASVs, (Callahan, McMurdie, & Holmes, 2017)). We then estimate and compare patterns in bacterial alpha and beta diversity. To investigate the evolution of microbiomes as a species level trait across avian species, we use phylogenetic comparative methods. We assess the phylogenetic signal of various microbiome traits (e.g., relative abundance

of common bacterial phyla, alpha diversity estimates) and assess the fit of four distinct evolutionary models on the microbiome as a trait of the hosts. Based on the results of previous avian microbiome work, we expect to see both high intraspecific (e.g., Hird et al. 2014, Ganz et al., 2017; Hird, Ganz, Eisen, & Boyce, 2018) and high interspecific variation (e.g., Hird et al. 2015, Waite and Taylor 2014) in the microbiomes. We further expect that microbiome community composition will correlate to host taxonomy and ecology (e.g., Hird et al. 2015). Our null hypothesis is that microbiomes will fit a neutral model of evolution (i.e., the Brownian motion model).

METHODS

Sample Collection

Birds were collected from eight sites in Equatorial Guinea from 5 January 2016 to 2 February 2016 as part of a general collecting trip with the Louisiana State University Museum of Natural Science (LSUMNS, n=92, Fig. 1; Table 1). Intestinal samples ranged from ~2cm to ~8cm. The 92 individuals belonged to 74 species, 55 genera, 30 families and 7 orders. Mist nets were used to catch birds, which were then euthanized by rapid cardiac compression (Engilis, Engilis, & Paul-Murphy, 2018) and immediately prepared as museum specimens; fieldwork was conducted under an LSU IACUC protocol 15-036. Intestinal tracts were tied off, removed and stored in 100% ethanol within approximately 15 minutes of death. Specimens were collected and transported under permits provided by the The Dirección General de Protección y Guardería Forestal and the Universidad Nacional de Guinea Ecuatorial. Samples were stored in -20°C freezers upon arrival at LSUMNS and remained frozen until they were thawed for extraction.

DNA Extraction and Sequencing

DNA was extracted from the luminal contents of the intestinal tract using Qiagen Power Fecal kits (Qiagen, Hilden, Germany) following manufacturer's protocols. Negative controls were included with every set of extractions. When a sample allowed (i.e., was large enough and in good enough condition to subdivide), the contiguous intestine was physically divided into adjacent subsamples; these samples are labeled with the suffixes A, B, C, D, representative of their physical relationship to one another (i.e., A was the first piece, B second, C third, D fourth).

Microbial communities from 142 samples were sequenced at the Microbial Analysis, Resources and Services (MARS) facility at the University of Connecticut. Quant-iT PicoGreen kit was used to quantify DNA concentrations. Amplification of the V4 region of the 16S rRNA gene was done using 30 ng of extracted DNA

as template. The V4 region was amplified using 515F and 806R with Illumina adapters and dual barcodes (Caporaso et al., 2011; Kozich, Westcott, Baxter, Highlander, & Schloss, 2013). The PCR reaction consisted of 95˚C for 3.5 minutes, 30 cycles of 30 s at 95.0°C, 30 s at 50.0°C and 90 s at 72.0°C, followed by final extension at 72.0°C for 10 minutes. PCR products were normalized based on the concentration of DNA from 250-400 bp then pooled. The pooled PCR products were cleaned using the Mag-Bind RxnPure Plus (Omega Bio-tek) according to the manufacturer's protocol. The cleaned pool was sequenced on the MiSeq using v2 2x250 base pair kit (Illumina, Inc). PCR replicates were sequenced at random for a subset of the samples to assess sequencing variation; these samples are labeled with the suffixes X, Y. Negative extraction and PCR controls were sequenced with the samples.

Sequence Analyses: Mothur

Raw sequences were processed and analyzed in Mothur version 1.39.5 (Schloss et al., 2009) according to the standard MiSeq protocol (available at http://www.mothur.org/wiki/MiSeq_SOP). Quality control, trimming and de-noising were performed following (Kozich et al., 2013). All sequences were aligned to the Silva reference alignment database version 128 (Quast et al., 2013), using the align.seqs command and the default settings (kmer searching and Needleman-Wunsch pairwise alignment). Database choice can influence results (Balvočiūtė & Huson, 2017) and Silva was chosen because it is updated regularly, manually curated and is predominantly based on ribosomal RNA sequence data. Sequences were filtered so that they overlapped with no overhangs and were clustered into Operational Taxonomic Units (OTUs) at 97% or 99% sequence similarity using the opticlust algorithm (Westcott & Schloss, 2017). The 99%OTU(NS) dataset is the 99%OTU dataset with all singletons removed. Taxonomic identification of all OTUs (up to the genus-level) was assigned according to current taxonomy (Silva v128), using the classify.seqs command and the default settings (using a k-nearest neighbor consensus and the wang method). Chimeric sequences were detected using the VSEARCH algorithm and removed from the dataset based on abundance of unique sequences within each sample. Sequences that were not identified as Bacteria and those that were identified as mitochondria and chloroplasts were removed from analysis. Additional analyses (described below) were conducted in R (R Core Team, 2013) using the packages phyloseq (McMurdie & Holmes, 2013) and vegan (Dixon, 2003; Oksanen et al., 2010).

Sequence Analyses: DADA2

Raw sequences were processed through DADA2 according to the DADA2 tutorial 1.2 in R (Callahan et al., 2016). DADA2 is an R package that estimates ASVs from sequence data and is error aware because it includes quality data when deciding whether to include each sequence in the dataset. Filtering was completed according

to the DADA2 protocol based on the number of ambiguous bases, a minimum quality score, and the expected errors in a read. The filtered fastq files were dereplicated and output as unique sequences with their corresponding abundance. Denoising and merging were completed according to the DADA2 pipeline tutorial, while error parameters removed sequencing errors (Callahan et al., 2016). The command removeBimeraDenovo, using the "consensus" method, detected chimeric sequences, which were then removed. All sequences were aligned to the current Silva reference alignment database (Silva v128), the same as was used for the Mothur analysis. R packages DECIPHER (Wright, 2016) and phangorn (Schliep, 2011) were used to create an alignment and a phylogenetic tree of the ASVs. Sequences that were identified as mitochondria or chloroplasts were removed from analysis. Additional analyses (described below) were conducted in R, using the packages phyloseq (McMurdie & Holmes, 2013) and vegan (Dixon, 2003; Oksanen et al., 2010). To assess the genera present in the samples, we used the ASV dataset and the tax_glom function in phyloseq to merge all ASVs assigned to the same genus.

Identifying contaminants

The R package Decontam (Davis, Proctor, Holmes, Relman, & Callahan, 2018) was run on all four datasets to remove likely contaminant sequences. Decontam uses a statistical model and a dataset's negative controls to identify potential contaminants. Decontam conducted prevalence-based contaminant identification on the phyloseq objects created for each dataset according to a stringent threshold set at 0.5, which identifies as contaminants all sequences that are more prevalent in the negative controls than in the samples. The number of OTUs that fell below the threshold were: 83 of 9249 (97%OTU), 117 of 22737 (99%OTU), 86 of 12126 OTUs (99%OTU(NS)), and 12 of 4233 (ASV). The OTUs/ASVs identified as contaminants were removed from the dataset and subsequent analyses.

Alpha diversity

After rarefying all samples to 2000 reads, we estimated the number of bacterial taxa found in the samples using phyloseq (Callahan et al., 2016) in R. We calculated three metrics: raw richness (Observed OTUs/ASVs), Shannon diversity, and the Simpson index. Shannon diversity takes abundance and evenness into account while the Simpson index uses richness and evenness.

Beta diversity

To maximize the number of samples retained in this analysis, samples were rarefied to 1550 sequences / sample. This kept 135 of the 142 samples in the 97%OTU, 99%OTU and 99%OTU(NS) datasets and 132

samples from the ASV dataset. Pairwise distance matrices using the phylogenetic distance metric UniFrac (Lozupone & Knight, 2005) were created in Mothur for the 97%OTU, 99%OTU and 99%OTU(NS) datasets using the unifrac.unweighted and unifrac.weighted commands. We then conducted non-metric multidimensional scaling (NMDS) on the distance matrices in vegan using the metaMDS function. For the ASV dataset, we calculated the UniFrac distance matrix and conducted the NMDS ordination in phyloseq, using ordinate. Distance matrices using the non-phylogenetic Bray-Curtis distance metric were also calculated and NMDS was conducted for all four datasets in phyloseq.

To determine the significance of metadata associated with each sample, PERMANOVA (Adonis) tests were performed using the vegan package in R. We tested bird order, bird genus, sampling locality, a larger geographic sampling area (province), and diet. The variables were tested in all datasets using three distance metrics (Bray-Curtis, weighted UniFrac, and unweighted UniFrac).

Sample Vs. PCR Replicates

To compare the consistency of sample replicates to the consistency of PCR replicates, we calculated (pairwise) Jaccard Index of bacterial phyla between adjacent sample replicates and PCR replicates. Jaccard Index is a presence/absence metric that calculates the intersection of a set over the union. A Jaccard Index of 1 indicates perfect coincidence between two sets and 0 means no taxa were shared. A Wilcoxon test was used to assess significance.

Phylogenetic comparative methods

To test the evolutionary signal of the microbiome, we used a species-level phylogeny from birdtree.org using the "all species" tree on the Ericson backbone (Ericson et al., 2006; Jetz et al., 2014; Jetz, Thomas, Joy, Hartmann, & Mooers, 2012). Some taxonomic names have recently changed and we include these changes on Table 1, with birdtree.org names in parentheses behind current taxonomic classifications. Microbiome traits were averaged across individuals and replicates belonging to the same species. Although this means some species included more data than others, it made use of all available data and avoided the bias introduced should we randomly prune all species to a single individual. The host traits we tested were the relative abundances of the five most abundant bacterial phyla, the average alpha diversity (observed taxa) for all datasets, two metrics of beta diversity (the coordinates of the NMDS ordination calculated above at the 1550 sequences/sample rarefaction level) and bird mass. Bird mass was measured in the field using an Ohaus HH320 electronic scale and then averaged across individuals belonging to the same species. This morphological trait was included as a

proof of concept for the phylogenetic comparative methods, since morphological traits, and specifically bird size, have been analyzed with phylogenetic comparative methods and provide an expectation for how a trait has evolved on our phylogeny.

We used the R package geiger (Harmon, Weir, Brock, Glor, & Challenger, 2008; Pennell et al., 2014) to test the model fit of the Brownian motion, Ornstein-Uhlenbeck (single optimum), Early Burst, and White Noise models and compared them using weighted AIC. We used the fitContinuous() function to fit the models to the data; the aicw() function calculated the Akaike weights from the AIC scores computed for each model. We also calculated phylogenetic signal ($\boldsymbol{\lambda}$) (Pagel, 1999) of the traits using fitContinuous(); a $\boldsymbol{\lambda}$ value close to 0 indicates no phylogenetic signal whereas a value close to 1 indicates the trait is consistent with neutral evolution along the phylogeny.

All statistical analyses were conducted in R version 1.1.423 (R Core Team 2013). Code used for all analyses can be found in Supplemental Materials.

RESULTS

Microbiome composition

Firmicutes (average relative abundance 57%) and Proteobacteria (average relative abundance 23%) were the most abundant phyla among samples and all datasets (Fig. 2). A common microbiome of 19 phyla was shared among all four datasets; an additional seven phyla were found in the three OTU-based datasets and one phylum was unique to ASVs (Fig. 2). Four phyla were found in >50% of all samples: Proteobacteria, Firmicutes, Actinobacteria and Bacteroidetes. Seven genera were found in more than one quarter of all samples (Table 2). *Escherichia/Shigella* was found in the greatest number of samples (121 / 142) and *Catellicoccus* had the greatest abundance from within 91 samples.

Alpha diversity

A raw richness estimate showed distinctly lower levels of alpha diversity in the ASV dataset, as compared to the Mothur-derived datasets (Fig. 3). At lower taxonomic ranks, Mothur-generated datasets and DADA2 deviated from one another (Fig. 3). The OTU-based datasets all contained more classes, orders, families and genera compared to DADA2 (Fig. 3). Shannon diversity and Simpson index contained similar patterns (Figure S1).

Beta diversity

Bray-Curtis, unweighted UniFrac, and weighted UniFrac distances were calculated between all pairwise sample comparisons in all datasets (all samples rarefied to 1550 sequences) and NMDS ordination was conducted on the matrices (Fig. S2-S4). PERMANOVA (Adonis) tests were run to quantify the effect size and significance of the metadata (all exact values are in Table S1). Bird order, genus, sampling locality, sampling region, diet, and sex were included as variables for all datasets and distance metrics (Fig. 4B). Across all datasets, bird taxonomy was highly significant (p=0.001 for all tests) and the variance explained (R^2) was largely consistent: between 6 and 13% for bird order and 43-70% for bird genus. The R^2 values for sampling locality were similar to that of bird order (7-11%), diet was slightly lower (1-4%). Sampling region (i.e., Bioko vs. mainland) and sex were the least significant variables and had R^2 values between 0 and 3%. An additional Bray-Curtis distance matrix was calculated for the four datasets at a rarefaction level of 10,000 sequences to assess the robustness of the data; the NMDS ordination is shown in Fig. 4. At 10,000 sequences/sample, bird order, bird genus, diet, locality and sex were all significant at p<0.005 for all four distance metrics with R^2 values similar to the levels of the 1,550 sequences/sample rarefied data. Sampling region was not significant for any distance metric (p>0.12).

Sample Vs. PCR Replicates

The Jaccard Indices for the PCR replicates (seven samples) varied from 0.33 to 1 across the four datasets; the sample replicates (39 samples) varied from 0.166 to 1 (Fig. 5). None of the four datasets had significantly different means of the distributions of Jaccard Indices (p<0.05), although the ASV dataset had a p-value of 0.062.

Phylogenetic Comparative Methods

We calculated phylogenetic signal ($\boldsymbol{\lambda}$, Fig. 6) and compared four evolutionary models using weighted AICs (Brownian motion, Ornstein-Uhlenbeck, Early Burst, and White Noise, Fig. 6). Results for lambda, as well as the model comparisons, were largely consistent across all datasets, meaning if $\boldsymbol{\lambda}$=0 for one dataset, $\boldsymbol{\lambda}$ = 0 for all four datasets. Three of the phyla (Actinobacteria, Proteobacteria, Spirochaetes) and alpha diversity metric (observed OTUs) had $\boldsymbol{\lambda}$ = 0. Firmicutes $\boldsymbol{\lambda}$ ~ 0.42 and Bacteroidetes $\boldsymbol{\lambda}$ ~ 0.76. Verrucomicrobia $\boldsymbol{\lambda}$ = 1, as did the first NMDS coordinate in the ASV dataset. The beta diversity metrics were more variable across the datasets than the taxonomic and alpha diversity metrics; NMDS 1 averaged ~.67 (with ASV dataset ~1) and NMDS_2 averaged ~0.2 (with 97%OTU and ASV ~0). Bird body mass had $\boldsymbol{\lambda}$ of 1, as did the dataset simulated under

Brownian motion. The White Noise simulation, as well as the randomized White Noise data and the randomized Brownian motion data had $\lambda = 0$.

Regarding model selection, for every microbiome trait, the Ornstein-Uhlenbeck and White Noise models shared all the weight, whereas no model had >0 weight on the Brownian motion or Early Burst models (Fig. 6). The one exception to this was the Verrucomicrobia, which had 100% support for the Early Burst model across all datasets. Analysis of the morphological character, bird mass, supported the Brownian motion model (60%) and the Early Burst and Ornstein-Uhlenbeck models each at ~20%. When we simulated data under a Brownian motion model, Brownian motion received 45% of the weight, the Ornstein-Uhlenbeck model 39% and Early Burst 16%. When the Brownian motion data were randomized, the White Noise model received 73% of the weight and OU received the remaining 26%. These values were nearly identical to when we simulated data under a White Noise model and when those data were randomized (Fig. 6).

DISCUSSION

The Equatoguinean Bird Intestinal Microbiome

A general goal of evolutionary biology is to describe biodiversity; 16S rRNA surveys of host-associated microbiomes provide valuable information about the distribution of microbial life. Birds play a critical role in terrestrial and aquatic ecosystems across the globe yet very few of the >10,000 species have anything known about their microbiomes. Birds are the largest clade of volant vertebrates, contain immense ecological and morphological diversity (especially with respect to their diets and gastrointestinal tracts), and are capable of ecologically connecting geographically distant locations through migration and dispersal (Bauer & Hoye, 2014). The avifauna of Equatorial Guinea is representative of the afrotropical region, with some endemic species and subspecies, especially on Bioko. Their diets are diverse and can contain large amounts of insects and fruit and while some species (not sampled here) undertake intra-tropical migrations, most species do not migrate. The Equatoguinean bird intestinal microbiome is dominated by Proteobacteria and Firmicutes, similar to other bird microbiome studies (Hird et al., 2015; Waite & Taylor, 2014). An additional 17 phyla were found in all of the datasets (Fig. 2E) – a common, or "core", microbiome that includes many phyla commonly found in the guts of birds (Grond et al., 2019b; Hird et al., 2015; Waite & Taylor, 2014) and mammals (Human Microbiome Project Consortium, 2012; McKenzie et al., 2017). Notably, this study contained a greater proportion of Firmicutes than some other bird microbiome surveys. A major feature of this study is the remoteness of the sampling sites that necessitated the use of 100% ethanol and a period of no refrigeration for preserving the samples. Several preservation methods, including ethanol, have been shown to affect the relative

abundances of certain taxa and are frequently in favor of increased Firmicutes (Song et al., 2016). Therefore, microbiome differences between this study and those with non-ethanol treated samples may be due -- at least in part -- to sample storage differences. The two bacterial genera detected in greater than half of the samples were *Escherichia/Shigella* and *Catellicoccus*. *Escherichia/Shigella* is a common gut microbe in many taxa Katouli, 2010), including birds. *Catellicoccus* were first described in 2006 following isolation from decomposing sea mammals (Lawson, Collins, Falsen, & Foster, 2006) and have subsequently been discovered to be a significant member of the gull fecal microbiome (Koskey, Fisher, Traudt, Newton, & McLellan, 2014; Lu, Santo Domingo, Lamendella, Edge, & Hill, 2008). *Enterococcus*, *Staphylococcus*, *Lactobacillus*, *Enterobacter*, *Clostridium* and *Lactococcus* were detected in at least one quarter of the samples and are all common gut microbes in birds (e.g., Waite & Taylor, 2014; Grond, Perreau, et al., 2019; Liu et al., 2019; van Dongen et al., 2013).

Bird taxonomy, diet and locality had similar amounts of variation explained (R^2) and significance (p) across the four datasets, with the highest percent explained by bird taxonomy; this is another commonality with multispecies microbiome surveys (e.g., Hird et al., 2015). Regardless of OTU picking method, the dataset contained thousands of "species level" bacterial taxa. While many of these taxa were low abundance, and there is the risk of false positives in any dataset, this is a high level of raw richness. Alpha diversity patterns were consistent across the four datasets; the Shannon diversity of many of the birds was lower than the average Shannon diversity of both wild and captive mammals (McKenzie et al., 2017). Again, this may be due to a high level of low abundance taxa. Sex of the bird was not significant in ten of the 12 tests in the full dataset but was significant ($p<0.005$) for the dataset at the higher level of rarefaction. This implies that sex may have an effect on the microbiome and requires relatively more data to detect. This is supported by its low explanatory power across all datasets ($R^2<3\%$). Sex-based differences in avian microbiome shared across species may just be difficult to detect, e.g., differences may be temporally restricted (e.g., during breeding season) or require a large sample size.

Sample Vs. PCR Replicates

Replicate samples from a single intestine were sequenced to elucidate the microvariation that occurs within the intestine. Variation in microbiota between discrete organs of the gastrointestinal tract has been demonstrated in diverse taxa e.g., in snakes (Colston, Noonan, & Jackson, 2015), in humans (Costello et al., 2009), in geese (Drovetski et al., 2018), in alligators (Keenan, Engel, & Elsey, 2013), and in teleost fish (Pratte, Besson, Hollman, & Stewart, 2018). However, few studies have looked at the variation within a single organ. The

replicates reveal clear taxonomic differences between adjacent sample replicates (Fig. 5) but the overall difference between adjacent sample replicates was not significantly different from the differences between PCR replicates. This contrasts with previous research which found PCR replicates to be more consistent than sample replicates (Poretsky, Rodriguez-R, Luo, Tsementzi, & Konstantinidis, 2014). Clearly some sample replicates varied more than others, e.g., see *Muscicapa infuscata* on Fig. 5, and the lowest Jaccard Index was 0.166 (indicating low overlap between the two samples). An NMDS ordination of Bray-Curtis distances for just the replicates shows that pairs tend to be relatively close together (Fig. S5). Standardization of sample collection from identical localities within a gastrointestinal organ may be difficult to impossible, especially under field conditions, but efforts should be made to ensure as consistent as possible sampling within regions of the gut. Using the Jaccard Index of bacterial phyla was a coarse comparison meant to detect the most significant differences between samples; future studies may delve further into the microvariation within body site microbiomes.

Phylogenetic Comparative Methods

The first bird trait that we analyzed was a morphological host trait (body mass), which is expected to contain phylogenetic signal and thus, acted as a positive control for our study. Harmon et al. (2010) found that, for bird size, the Early Burst model is best supported when analyses are conducted *across* bird orders whereas the Brownian motion model is supported *within* avian orders. Our dataset is largely Passeriformes (81 of 92 specimens collected) but included seven orders in total, so the breakdown of support for the four models is consistent with our expectations (Fig. 6): roughly 60% of the AIC weight was on the Brownian motion model, with ~20% on Early Burst and ~20% on Ornstein-Uhlenbeck. Regarding the microbiome traits, the Ornstein-Uhlenbeck and White Noise models received 100% of the support for all the microbiome traits, except the phylum Verrucomicrobia (discussed below). When we simulated a pure Brownian motion model on our phylogenetic tree, support for Brownian motion (~45%) and Ornstein-Uhlenbeck (~38%) were roughly equal, with Early Burst receiving ~17% and White Noise receiving 0%. This was not entirely unexpected, as Ornstein-Uhlenbeck models can receive support under certain conditions, including on “small” trees, or less than 100 taxa (Cooper, Thomas, Venditti, Meade, & Freckleton, 2016). Notably, White Noise was not supported in the Brownian simulation but when we randomized the leaves of the tree, White Noise received ~73% of the support, with Ornstein-Uhlenbeck getting the remaining 26%; these were the same values as the White Noise simulation and the randomized White Noise simulation. Thus, model support near 73% White Noise/26% Ornstein-Uhlenbeck cannot be distinguished from pure White Noise. Several of the traits, in many of the datasets, were near this value (e.g., Proteobacteria, Bacteroidetes). It may be the case that all datasets that

share support between the White Noise and the Ornstein-Uhlenbeck models are indicative of White Noise, or lack of phylogenetic signal. The beta diversity traits – the first two axes of the NMDS plots from all samples rarefied to 1550 sequences – show more variability across the four datasets than the other microbiome traits. We hypothesize this is due to the nature of the metric: ordinations attempt to recreate accurate distances between samples in two-dimensional space from multidimensional data and are subject to variation.

λ indicates how transformed the phylogeny has to be to fit the trait distribution: a value of 0 means the tree has to be reduced to a star phylogeny to fit the data (because the data contain no phylogenetic signal) and a value of 1 means the tree fits the data (because the data have been neutrally evolving along the phylogeny). Values between 0 and 1 indicate some amount of scaling of the tree to fit the data. Our **λ** values ranged from 0 to 0.7; bird mass and simulated Brownian motion both had a **λ** of 1. We interpret this to mean that Brownian motion does not fit our data. At a minimum, our collective results support that the microbiome is not a neutrally evolving trait in birds, because there was no support of the Brownian motion model for any trait. Additionally, we believe the relative abundance of bacterial phyla, as estimated by 16S rRNA sequencing, is consistent with a White Noise model, in which microbiome traits are effectively a random distribution of possible values and contain little to no phylogenetic signal.

The Verrucomicrobia trait received 100% support for the Early Burst model from all four datasets, which contrasted starkly with the rest of the traits. This is likely because Verrucomicrobia was found in the four replicates from a single Cuculiformes specimen at a relative abundance between 7 and 42% but not above 1% in any of the other samples. Additionally, Cuculiformes is one of the earlier branching lineages on the phylogeny and for Verrucomicrobia to be high in Cuculiformes and low everywhere else is consistent with the Early Burst model. It must be noted that this result may be a false positive. Some of our bird taxa, like the Cuculiformes, are represented by a single individual and this greatly limits the robustness of the results. We cannot say whether the individual accurately represents the species or if it is anomalous (e.g., if it were sick). General collecting trips are stochastic in nature and we did not target specific bird species. Analyses with larger sample sizes will be possible with targeted sampling and higher throughput methods of microbiome sampling, perhaps fecal samples or cloacal swabs. Additionally, the use of ASVs for OTU estimation allows datasets to be combined (Callahan, McMurdie, & Holmes, 2017) and future meta-analyses of the avian microbiome may confirm or refute the evolution of Verrucomicrobia (and other microbiome traits) in the avian microbiome. Verrucomicrobia contains important gut commensals in humans (Fujio-Vejar et al., 2017; Collado, Derrien,

Isolauri, de Vos, & Salminen, 2007)) so, while it may be true that Verrucomicrobia are evolving under different pressures from the other phyla in our dataset, we urge caution with that interpretation.

There are some caveats to our use of phylogenetic comparative methods. For phylogenetic comparative methods to be useful, the trait under investigation must be heritable. Our understanding of the heritability of the microbiome is maturing, couched within host-microbe symbiosis, and varies across host taxa (van Opstal & Bordenstein, 2015). The heritability of the avian microbiome, and especially in these species, is unknown but there are several pieces of evidence to support that the microbiome may be at least partially heritable. First, there is a negative correlation between genetic distance and microbiome similarity in multiple bird species (e.g., Banks, Cary & Hogg, 2009; Pearce, Hoover, Jennings, Nevitt, & Docherty, 2017), indicating that host genetics exert some control on the microbiome. Second, the members of the microbiome are not a random sample of the environment and this selection, whether direct or indirect, is a property that may be vertically transmitted. Both vertical transfer of microbes and genetic control of microbiomes imply heritability (Kreisinger et al., 2017). Lastly, the data we have collected, 16S rRNA gene sequencing and relative abundance of bacterial taxa, can be thought of as the first pass of data for understanding the microbiome of a host species or clade. Taxonomic composition is more variable than functional capability in many microbial systems (Burke, Steinberg, Rusch, Kjelleberg, & Thomas, 2011; Human Microbiome Project Consortium, 2012), including avian microbiomes (Oakley et al., 2014), which is likely due to functional redundancy (Louca et al., 2018). Additionally, sequencing a single gene and assessing the relative abundance of the data is compositional by definition and limited in its ability to describe and compare communities because changes in relative abundance cannot be used to infer changes in absolute abundance (Gloor, Macklaim, Pawlowsky-Glahn, & Egozcue, 2017). Gathering absolute abundance of microbes, via quantitative PCR or flow cytometry, and functional data, via metatranscriptomics or metabolomics, for example, may provide more relevant data to the question of evolutionary signal in the microbiome.

Comparison of taxonomic units for wild microbiome studies

The definition of taxonomic units is a fundamental step in any ecological study; microbiome studies frequently rely on sequence-based datasets and must partition the sequences into biological units to analyze the data. Our results show that similar biological conclusions can be drawn from the four bacterial taxonomic units employed: 97%OTU, 99%OTU, 99%OTU(NS) and ASVs. Bacterial composition (Fig. 2), alpha diversity (Fig. 3, S1), beta diversity (Fig. 4), significance of the metadata (Fig. 4), and support of evolutionary models (Fig. 6) were similar across the four datasets. The relative patterns in alpha diversity are quite similar across the four

datasets, indicating that despite differences in raw counts of taxa, their relationships in the context of their communities is consistent. The beta diversity results were also consistent (Fig. 4B), showing bird taxonomy as the highest correlate to microbiome diversity and largely agreeing on which metadata were significant correlates. This is despite the fact that we used different programs to calculate the UniFrac distances which was done to recreate the most likely data generation practices of empirical scientists and to test the robustness of the analyses. Qualitatively, there were very few significant differences in interpretation of the data between the taxonomic units; a result that has been found elsewhere (Nearing, Douglas, Comeau, & Langille, 2018).

There were, however, patterns worth noting in the quantitative differences. Total taxa decreased from 97%OTU to 99%OTU to 99%OTU(NS) to ASVs for taxonomic levels class through genus (Fig. 3). Only the 97%OTU and 99%OTU datasets contained the exact same phyla (Fig. 2); the datasets that excluded singletons, 99%OTU(NS) and ASVs thus contained less alpha diversity, but it seems reasonable that the diversity that was "lost" was low abundance taxa or false positives. There was a noticeable decrease in the number of taxa between the percent similarity OTU datasets and the ASVs. On the surface, this is counterintuitive, as the percent similarity OTU datasets should be lumping diversity into fewer units than the unique sequence-based method. However, DADA2 implements strict quality control and removes all singletons from the dataset, which eliminates a large number of ASVs. It may also be true that the datasets with higher quantitative diversity have a higher false positive rate. It seems almost certain that this method is conservative and may remove real biological diversity from the dataset; however, there is no way to distinguish error from rare, but true, taxa in this scenario. The necessity of applying some rules and cutoffs to biological data means that a balance will always need to be struck between splitting and lumping. The fact that relative results and biological conclusions were consistent across datasets is reassuring but improvements in the detection of false positives and a better understanding of the "rare biosphere" will enhance broad scale biodiversity studies.

ACKNOWLEDGMENTS

The authors would like to thank Alexandra Alexiev for help with the DNA extractions. Luke Powell, Jacob Cooper, Jared Wolfe, Kristin Brzeski, Amancio Etinguë, and Tristan Davis provided invaluable assistance and companionship during field work in Equatorial Guinea. The Bioko Biodiversity Protection Program provided logistical support on Bioko Island. Donna Dittmann assisted with specimen permitting and transportation. Funding for field work was provided by the Louisiana State University Museum of Natural Science. Kirsten Grond and five anonymous reviewers provided thoughtful comments on the manuscript that improved it greatly. SMH would also like to thank the UC Davis Chancellor's Postdoctoral Fellowship and the University

of Connecticut for support. OJ is supported by the National Science Foundation Graduate Research Fellowship under grant no. DGE-1247192.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723.

Al-Asmakh, M., Stukenborg, J.-B., Reda, A., Anuar, F., Strand, M.-L., Hedin, L., … Söder, O. (2014). The gut microbiota and developmental programming of the testis in mice. PloS One, 9(8), e103809.

Alberdi, A., Aizpurua, O., Bohmann, K., Zepeda-Mendoza, M. L., & Gilbert, M. T. P. (2016). Do Vertebrate Gut Metagenomes Confer Rapid Ecological Adaptation? Trends in Ecology & Evolution, 31(9), 689–699.

Balvočiūtė, M., & Huson, D. H. (2017). SILVA, RDP, Greengenes, NCBI and OTT - how do these taxonomies compare? BMC Genomics, 18(Suppl 2), 114.

Banks, J. C., Cary, S. C., & Hogg, I. D. (2009). The phylogeography of Adelie penguin faecal flora. Environmental Microbiology, 11(3), 577–588.

Bauer, S., & Hoye, B. J. (2014). Migratory animals couple biodiversity and ecosystem functioning worldwide. Science, 344(6179), 1242552.

Benson, A. K. (2016). The gut microbiome-an emerging complex trait. Nature Genetics, 48(11), 1301–1302.

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.

Borrow, N., & Demey, R. (2014). Birds of western Africa; Second Edition. Princeton University Press, New Jersey, USA.

Brandl, H. B., van Dongen, W. F. D., Darolová, A., Krištofík, J., Majtan, J., & Hoi, H. (2014). Composition of Bacterial Assemblages in Different Components of Reed Warbler Nests and a Possible Role of Egg Incubation in Pathogen Regulation. PloS One, 9(12), e114861.

Brooks, A. W., Kohl, K. D., Brucker, R. M., van Opstal, E. J., & Bordenstein, S. R. (2016). Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. PLoS Biology, 14(11), e2000225.

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., & Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. Proceedings of the National Academy of Sciences of the United States of America, 108(34), 14288–14293.

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. The ISME Journal, 11(12), 2639–2643.

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), 581–583.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., … Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences of the United States of America, 108 (Supplement 1), 4516–4522.

Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. Nature Reviews Genetics, 13(4), 260–270.

Collado, M. C., Derrien, M., Isolauri, E., de Vos, W. M., & Salminen, S. (2007). Intestinal integrity and Akkermansia muciniphila, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. Applied and Environmental Microbiology, 73(23), 7767–7770.

Colston, T. J., Noonan, B. P., & Jackson, C. R. (2015). Phylogenetic Analysis of Bacterial Communities in Different Regions of the Gastrointestinal Tract of Agkistrodon piscivorus, the Cottonmouth Snake. PloS One, 10(6), e0128793–13.

Colston, T. J., & Jackson, C. R. (2016). Microbiome evolution along divergent branches of the vertebrate tree of life: what is known and unknown. Molecular Ecology, 25(16), 3776–3800.

Cooper, N., Thomas, G. H., Venditti, C., Meade, A., & Freckleton, R. P. (2016). A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. Biological Journal of the Linnean Society. Linnean Society of London, 118(1), 64–77.

Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., & Knight, R. (2009). Bacterial community variation in human body habitats across space and time. Science, 326(5960), 1694–1697.

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome, 6(1), 226.

Dewar, M. L., Arnould, J. P. Y., Krause, L., Dann, P., & Smith, S. C. (2014). Interspecific variations in the faecal microbiota of Procellariiform seabirds. FEMS Microbiology Ecology, 89(1), 47–55.

Diaz Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., … Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. Proceedings of the National Academy of Sciences of the United States of America, 108(7), 3047–3052.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. Journal of Vegetation Science: Official Organ of the International Association for Vegetation Science, 14(6), 927–930.

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., & Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proceedings of the National Academy of Sciences of the United States of America, 107(26), 11971–11975.

Drovetski, S. V., O'Mahoney, M., Ransome, E. J., Matterson, K. O., Lim, H. C., Chesser, R. T., & Graves, G. R. (2018). Spatial Organization of the Gastrointestinal Microbiota in Urban Canada Geese. Scientific Reports, 8(1), 3713.Edgar, R. C. (2017). Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. PeerJ, 5, e3889.

Engel, P., & Moran, N. A. (2013). The gut microbiota of insects - diversity in structure and function. FEMS Microbiology Reviews, 37(5), 699–735.

Engilis, A., Jr., Engilis, I. E., & Paul-Murphy, J. (2018). Rapid cardiac compression: An effective method of avian euthanasia. The Condor, 120(3), 617–621.

Ericson, P. G. P., Anderson, C. L., Britton, T., Elzanowski, A., Johansson, U. S., Källersjö, M., … Mayr, G. (2006). Diversification of Neoaves: integration of molecular sequence data and fossils. Biology Letters, 2(4), 543–547.

Felsenstein, J. (1985). Phylogenies and the Comparative Method. The American Naturalist, 125(1), 1–15.

Fujio-Vejar, S., Vasquez, Y., Morales, P., Magne, F., Vera-Wolf, P., Ugalde, J. A., … Gotteland, M. (2017). The Gut Microbiota of Healthy Chilean Subjects Reveals a High Abundance of the Phylum Verrucomicrobia. Frontiers in Microbiology, 8, 1221.

Ganz, H. H., Doroud, L., Firl, A. J., Hird, S. M., Eisen, J. A., & Boyce, W. M. (2017). Community-Level Differences in the Microbiome of Healthy Wild Mallards and Those Infected by Influenza A Viruses. mSystems, 2(1). doi: 10.1128/mSystems.00188-16.

Givens, C. E., Ransom, B., Bano, N., & Hollibaugh, J. T. (2015). Comparison of the gut microbiomes of 12 bony fish and 3 shark species. Marine Ecology Progress Series, 518, 209–223.

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome Datasets Are Compositional: And This Is Not Optional. Frontiers in Microbiology, 8, 2224.

Godoy-Vitorino, F., Goldfarb, K. C., Karaoz, U., Leal, S., Garcia-Amado, M. A., Hugenholtz, P., … Dominguez-Bello, M. G. (2012). Comparative analyses of foregut and hindgut bacterial communities in hoatzins and cows. The ISME Journal, 6(3), 531–541.

Godoy-Vitorino, F., Ley, R. E., Gao, Z., Pei, Z., Ortiz-Zuazaga, H., Pericchi, L. R., … Domínguez-Bello, M. G. (2008). Bacterial community in the crop of the hoatzin, a neotropical folivorous flying bird. Applied and Environmental Microbiology, 74(19), 5905–5912.

Grond, K., Lanctot, R. B., Jumpponen, A., & Sandercock, B. K. (2017). Recruitment and establishment of the gut microbiome in arctic shorebirds. FEMS Microbiology Ecology, 93(12). https://doi.org/10.1093/femsec/fix142

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.

Grond, K., Perreau, J. M., Loo, W. T., Spring, A. J., Cavanaugh, C. M., & Hird, S. M. (2019a). Longitudinal microbiome profiling reveals impermanence of probiotic bacteria in domestic pigeons. PloS One, 14(6), e0217804.

Grond, K., Santo Domingo, J. W., Lanctot, R. B., Jumpponen, A., Bentzen, R. L., Boldenow, M. L., … Sandercock, B. K. (2019b). Composition and Drivers of Gut Microbial Communities in Arctic-Breeding Shorebirds. Frontiers in Microbiology, 10, 2258.

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.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2008). GEIGER: investigating evolutionary radiations. Bioinformatics , 24(1), 129–131.

He, Y., Caporaso, J. G., Jiang, X.-T., Sheng, H.-F., Huse, S. M., Rideout, J. R., … Zhou, H.-W. (2015). Stability of operational taxonomic units: an important but neglected property for analyzing microbial diversity. Microbiome, 3, 20.

Hird, S. M. (2017). Evolutionary Biology Needs Wild Microbiomes. Frontiers in Microbiology, 8, 725.

Hird, S. M., Carstens, B. C., Cardiff, S. W., Dittmann, D. L., & Brumfield, R. T. (2014). Sampling locality is more detectable than taxonomy or ecology in the gut microbiota of the brood-parasitic Brown-headed Cowbird (Molothrus ater). PeerJ, 2(10), e321.

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.

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

Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. Nature, 486(7402), 207–214.

Ibáñez-Álamo, J. D., Ruiz-Rodríguez, M., & Soler, J. J. (2014). The mucous covering of fecal sacs prevents birds from infection with enteric bacteria. Journal of Avian Biology, 45(4), 354–358.

IUCN 2018. The IUCN Red List of Threatened Species. Version 2018. http://www.iucnredlist.org.

Jacob, S., Parthuisot, N., Vallat, A., Ramon-Portugal, F., Helfenstein, F., & Heeb, P. (2015). Microbiome affects egg carotenoid investment, nestling development and adult oxidative costs of reproduction in Great tits. Functional Ecology, 29(8), 1048–1058.

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.

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.

Katouli, M. (2010). Population structure of gut Escherichia coli and its role in development of extra-intestinal infections. Iranian Journal of Microbiology, 2(2), 59–72.

Keenan, S. W., Engel, A. S., & Elsey, R. M. (2013). The alligator gut microbiome and implications for archosaur symbioses. Scientific Reports, 3, 1–7.

Kilgas, P., Saag, P., Mägi, M., Tilgar, V., & Mänd, R. (2012). Plumage bacterial load increases during nest-building in a passerine bird. Journal of Ornithology / DO-G, 153(3), 833–838.

Kohl, K. D., Varner, J., Wilkening, J. L., & Dearing, M. D. (2018). Gut microbial communities of American pikas (Ochotona princeps): Evidence for phylosymbiosis and adaptations to novel diets. The Journal of Animal Ecology, 87(2), 323–330.

Koskey, A. M., Fisher, J. C., Traudt, M. F., Newton, R. J., & McLellan, S. L. (2014). Analysis of the gull fecal microbial community reveals the dominance of Catellicoccus marimammalium in relation to culturable Enterococci. Applied and Environmental Microbiology, 80(2), 757–765.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology, 79(17), 5112–5120.

Kreisinger, J., Kropáčková, L., Petrželková, A., Adámková, M., Tomášek, O., Martin, J.-F., … Albrecht, T. (2017). Temporal Stability and the Effect of Transgenerational Transfer on Fecal Microbiota Structure in a Long Distance Migratory Bird. Frontiers in Microbiology, 8, 50.

Kropáčková, L., Těšický, M., Albrecht, T., Kubovčiak, J., Čížková, D., Tomášek, O., … Kreisinger, J. (2017). Codiversification of gastrointestinal microbiota and phylogeny in passerines is not explained by ecological divergence. Molecular Ecology, 26(19), 5292–5304.

Kyle, G. Z., & Kyle, P. D. (2004). Rehabilitation and Conservation of Chimney Swifts. Driftwood Wildlife Association.

Lawson, P. A., Collins, M. D., Falsen, E., & Foster, G. (2006). Catellicoccus marimammalium gen. nov., sp. nov., a novel Gram-positive, catalase-negative, coccus-shaped bacterium from porpoise and grey seal. International Journal of Systematic and Evolutionary Microbiology, 56(Pt 2), 429–432.

Ley, R. E., Peterson, D. A., & Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell, 124(4), 837–848.

Liu, H., Chen, Z., Gao, G., Sun, C., Li, Y., & Zhu, Y. (2019). Characterization and comparison of gut microbiomes in nine species of parrots in captivity. Symbiosis , 78(3), 241–250.

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'Connor, M. I., … Parfrey, L. W. (2018). Function and functional redundancy in microbial systems. Nature Ecology & Evolution, 2(6), 936–943.

Lozupone, C., & Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Applied and Environmental Microbiology, 71(12), 8228–8235.

Lu, J., Santo Domingo, J. W., Lamendella, R., Edge, T., & Hill, S. (2008). Phylogenetic diversity and molecular detection of bacteria in gull feces. Applied and Environmental Microbiology, 74(13), 3969–3976.

Mazel, F., Davis, K. M., Loudon, A., Kwong, W. K., Groussin, M., & Parfrey, L. W. (2018). Is Host Filtering the Main Driver of Phylosymbiosis across the Tree of Life? mSystems, 3(5). doi: 10.1128/mSystems.00097-18.

McKenzie, V. J., Song, S. J., Delsuc, F., Prest, T. L., Oliverio, A. M., Korpita, T. M., … Knight, R. (2017). The Effects of Captivity on the Mammalian Gut Microbiome. Integrative and Comparative Biology, 57(4), 690–704.

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.

Nearing, J. T., Douglas, G. M., Comeau, A. M., & Langille, M. G. I. (2018). Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. PeerJ, 6, e5364.

Oakley, B. B., Lillehoj, H. S., Kogut, M. H., Kim, W. K., Maurer, J. J., Pedroso, A., … Cox, N. A. (2014). The chicken gastrointestinal microbiome. FEMS Microbiology Letters, 360(2), 100–112.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'hara, R. B., Simpson, G. L., … Wagner, H. (2010). vegan: Community Ecology Package. R package version 1.17-2. Http://cran. R-Project. Org>. Acesso Em, 23, 2010.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. Nature, 401(6756), 877–884.

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.

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.

Pennell, M. W., & Harmon, L. J. (2013). An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. Annals of the New York Academy of Sciences, 1289, 90–105.

Pollock, F. J., McMinds, R., Smith, S., Bourne, D. G., Willis, B. L., Medina, M., … Zaneveld, J. R. (2018). Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. Nature Communications, 9(1), 4921.

Poretsky, R., Rodriguez-R, L. M., Luo, C., Tsementzi, D., & Konstantinidis, K. T. (2014). Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. PloS One, 9(4), e93827.

Pratte, Z. A., Besson, M., Hollman, R. D., & Stewart, F. J. (2018). The Gills of Reef Fish Support a Distinct Microbiome Influenced by Host-Specific Factors. Applied and Environmental Microbiology, 84(9). doi: 10.1128/AEM.00063-18

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41(Database issue), D590–D596.

R Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Roggenbuck, M., Schnell, I. B., Blom, N., Bælum, J., Bertelsen, M. F., Sicheritz-Pontén, T., … Hansen, L. H. (2014). The microbiome of New World vultures. Nature Communications, 5, 5498.

Rohlfs, R. V., Harrigan, P., & Nielsen, R. (2014). Modeling gene expression evolution with an extended Ornstein-Uhlenbeck process accounting for within-species variation. Molecular Biology and Evolution, 31(1), 201–211.

Rubio, L. A., Brenes, A., Setién, I., de la Asunción, G., Durán, N., & Cutuli, M. T. (1998). Lactobacilli counts in crop, ileum and caecum of growing broiler chickens fed on practical diets containing whole or dehulled sweet lupin (Lupinus angustifolius) seed meal. British Poultry Science, 39(3), 354–359.

Schliep, K. P. (2011). phangorn: phylogenetic analysis in R. Bioinformatics , 27(4), 592–593.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., … Weber, C. F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology, 75(23), 7537–7541.

Song, S. J., Amir, A., Metcalf, J. L., Amato, K. R., Xu, Z. Z., Humphrey, G., & Knight, R. (2016). Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. mSystems, 1(3). https://doi.org/10.1128/mSystems.00021-16

Spor, A., Koren, O., & Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. Nature Reviews. Microbiology, 9(4), 279–290.

Strickland, H. E. (1844). XLVI.—Descriptions of several new or imperfectly-defined Genera and Species of Birds. Annals and Magazine of Natural History, 13(86), 409–421.

Teyssier, A., Lens, L., Matthysen, E., & White, J. (2018). Dynamics of Gut Microbiota Diversity During the Early Development of an Avian Host: Evidence From a Cross-Foster Experiment. Frontiers in Microbiology, 9, 1524.

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the Theory of the Brownian Motion. Physics Review, 36(5), 823–841.

Van Dongen, W. F. D., White, J., Brandl, H. B., Moodley, Y., Merkling, T., Leclaire, S., … Wagner, R. H. (2013). Age-related differences in the cloacal microbiota of a wild bird species. BMC Ecology, 13, 11.

van Opstal, E. J., & Bordenstein, S. R. (2015). Rethinking heritability of the microbiome. Science, 349(6253), 1172–1173.

Waite, D. W., & Taylor, M. W. (2014). Characterizing the avian gut microbiota: membership, driving influences, and potential function. Frontiers in Microbiology. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4032936/

Wang, J., Brelsfoard, C., Wu, Y., & Aksoy, S. (2013). Intercommunity effects on microbiome and GpSGHV density regulation in tsetse flies. Journal of Invertebrate Pathology, 112 Suppl, S32–S39.

Westcott, S. L., & Schloss, P. D. (2017). OptiClust, an Improved Method for Assigning Amplicon-Based Sequence Data to Operational Taxonomic Units. mSphere, 2(2). https://doi.org/10.1128/mSphereDirect.00073-17

Whipps, J. M., Lewis, K., & Cooke, R. C. (1988). Mycoparasitism and plant disease control. Fungi in Biological Control Systems, 161–187.

Wienemann, T., Schmitt-Wagner, D., Meuser, K., Segelbacher, G., Schink, B., Brune, A., & Berthold, P. (2011). The bacterial microbiota in the ceca of Capercaillie (Tetrao urogallus) differs between wild and captive birds. Systematic and Applied Microbiology, 34(7), 542–551.

Wright, E. S. (2016). Using DECIPHER v2. 0 to analyze big biological sequence data in R. The R Journal, 8(1). Retrieved from https://journal.r-project.org/archive/2016-1/wright.pdf

Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., … Farnleitner, A. H. (2019). Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. Nature Communications, 10(1), 2200.

DATA ACCESSIBILITY

The data that support the findings of this study are openly available in the Dryad Digital Repository at https://doi.org/10.5061/dryad.jsxksn05j.

AUTHOR CONTRIBUTIONS

DC designed the project, performed the research, analyzed data, wrote and revised the manuscript.

OJ performed the research, collected data, revised manuscript.

RT performed the research, collected data, revised manuscript.

SH designed the project, performed the research, collected data, analyzed data, wrote and revised the manuscript.

FIGURES

Figure 1: Overview of bird hosts and sampling localities in Equatorial Guinea. Representative species from each order are shown with the key. Apodiformes (*Apus affinis*), Charadriiformes (*Actitis hypoleucos*), Columbiformes (*Turtur tympanistra*), Coraciiformes (*Corythornis leucogaster*), Cuculiformes (*Cercococcyx olivinus*), Passeriformes (left: *Platysteira chalybea*, right: *Mandingoa nitidula*), Piciformes (*Pogoniulus bilineatus*). Pie charts are sized relative to sample size.

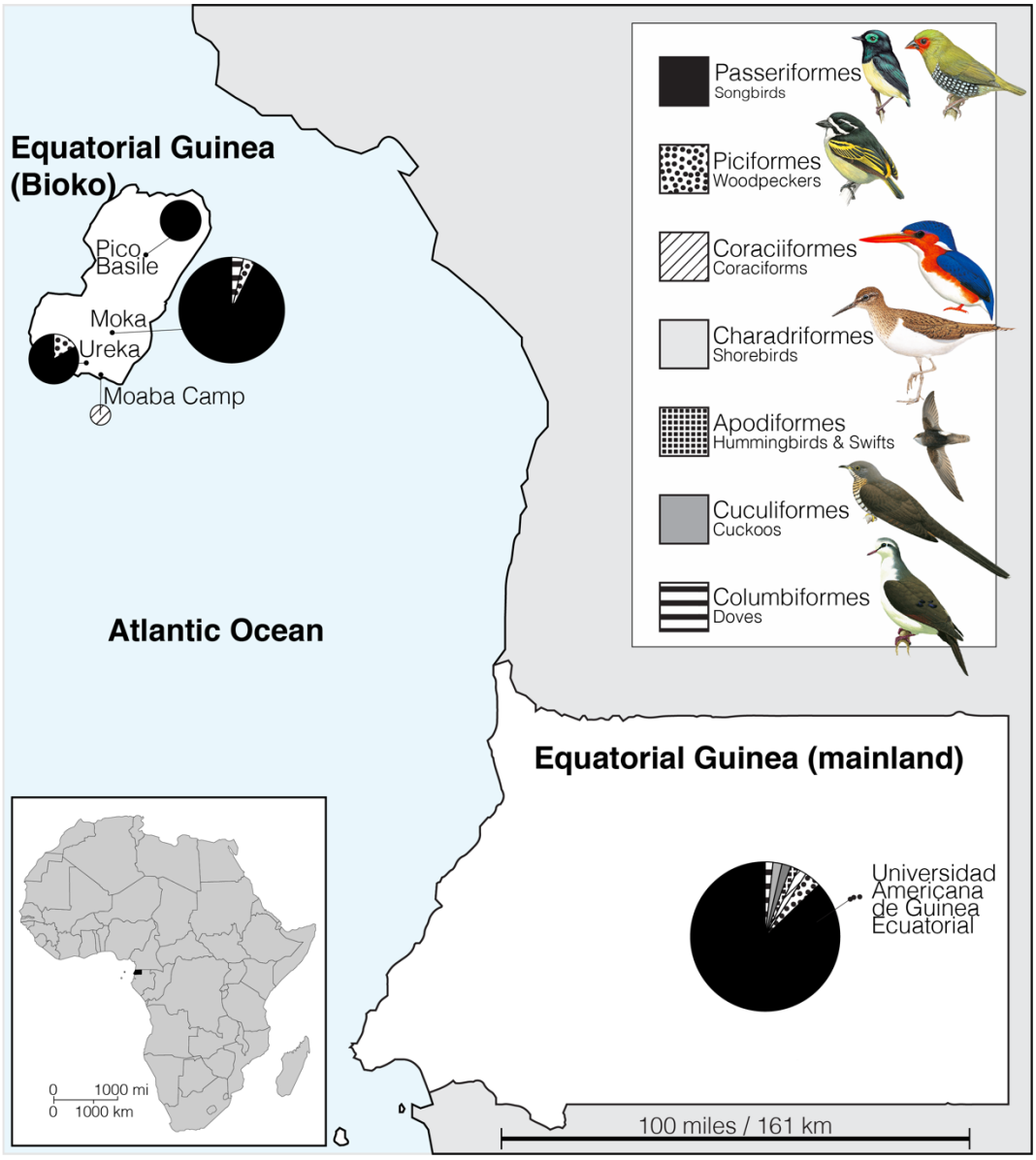

Figure 2: Relative abundance of bacterial phyla in all samples across the four datasets, (A) 97% OTUs, (B) 99% OTUs, (C) 99%NS OTUs (no singletons), (D) ASVs. (E) The overlap of microbiota found in the datasets.

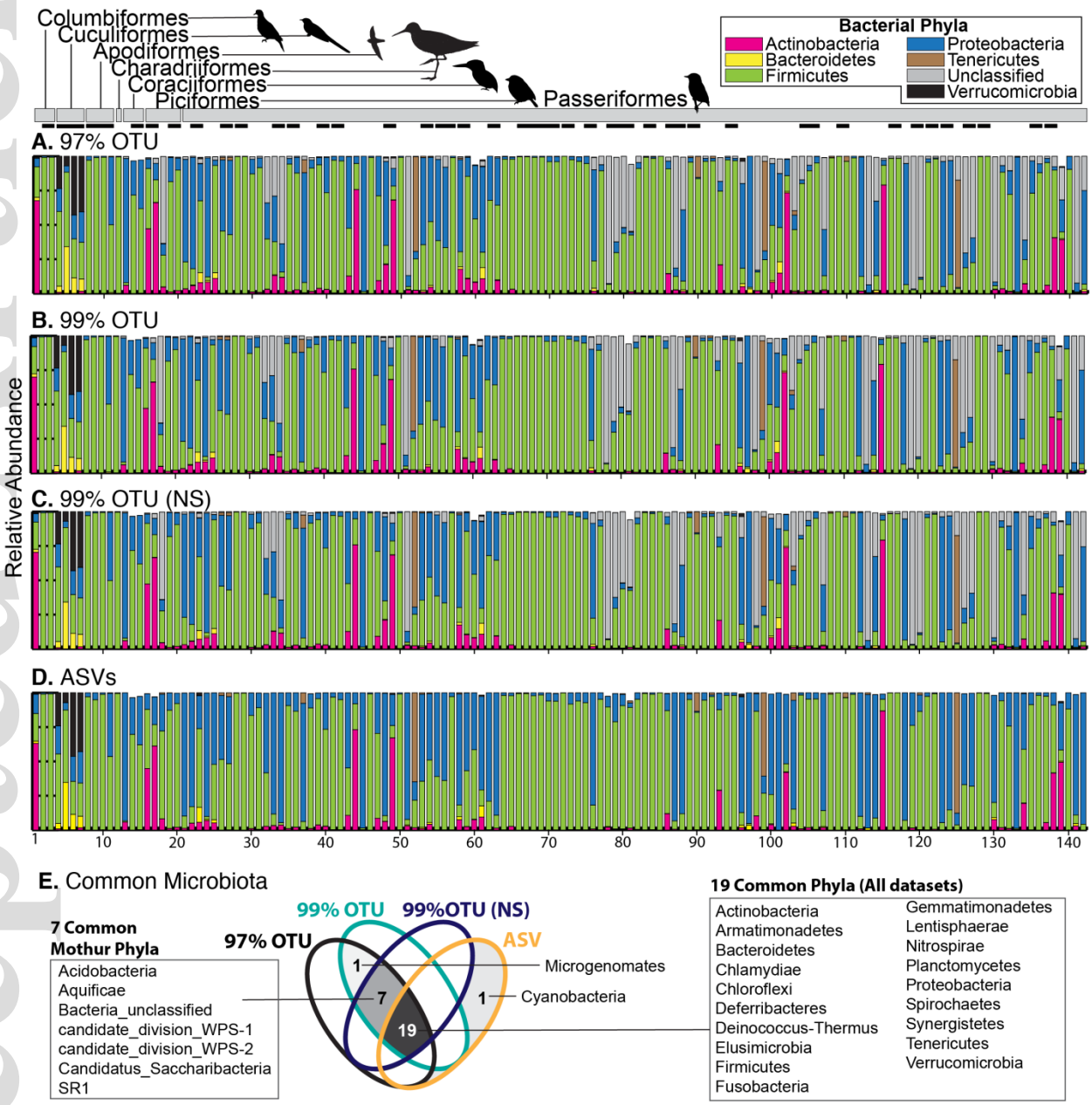

Figure 3: Alpha diversity across datasets (rarefied to 2000 reads). (A) Shannon Index for all samples, sorted by bird order and colored by dietary preference. (B) Number of taxa observed at different ranks across all datasets; datasets not rarefied to show total comparison of beginning datasets except for bar chart labeled OTUs (rarefied), which was rarefied to 2000 reads/sample.

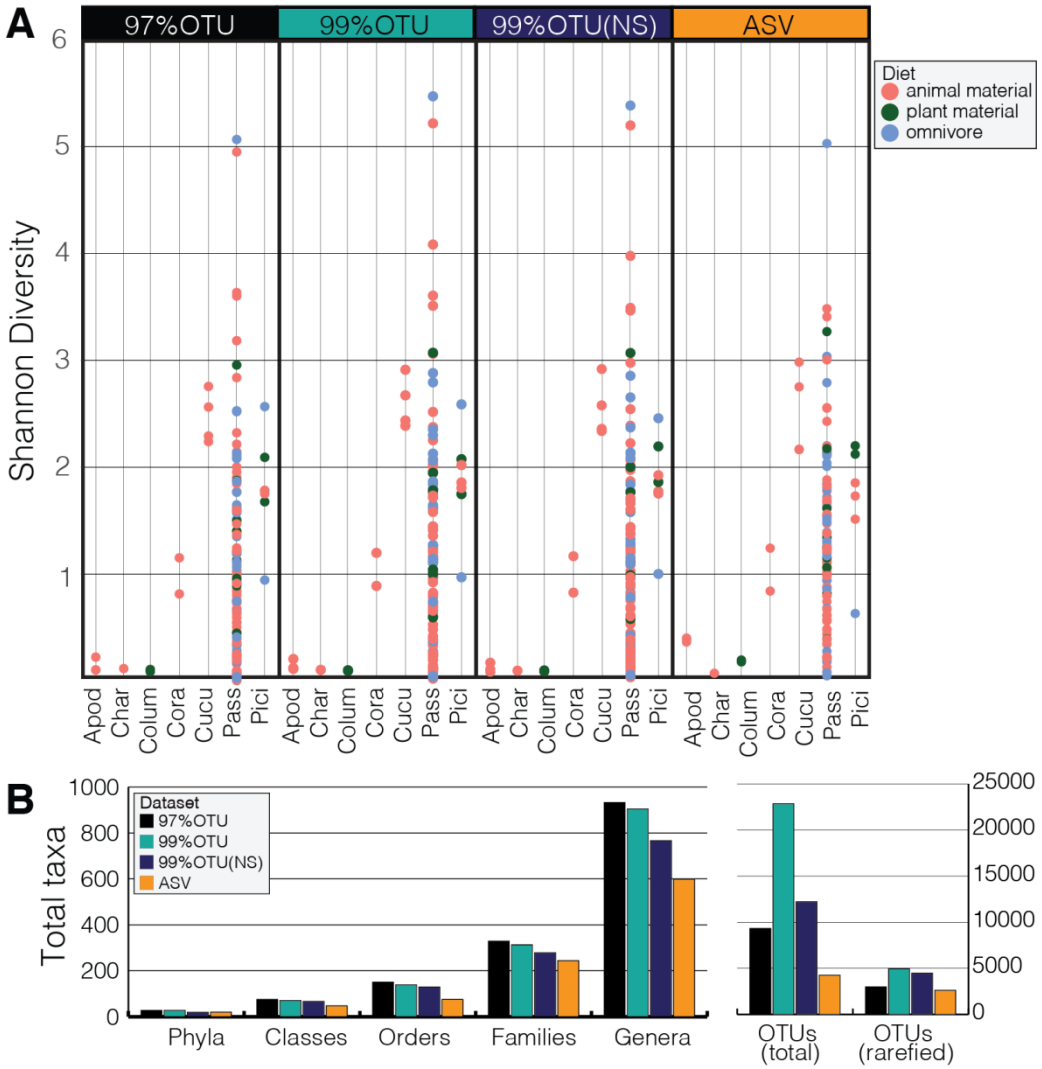

Figure 4: Representative beta diversity across datasets. (A) Ordination of Bray-Curtis distances in all four datasets colored by host order and rarefied to 10000 reads; effect size and significance of bird order, estimated by PERMANOVA (Adonis), shown on plots. (B) Effect size and significance of all categorical metadata for all datasets and distance metrics, rarefied to 1550 sequences.

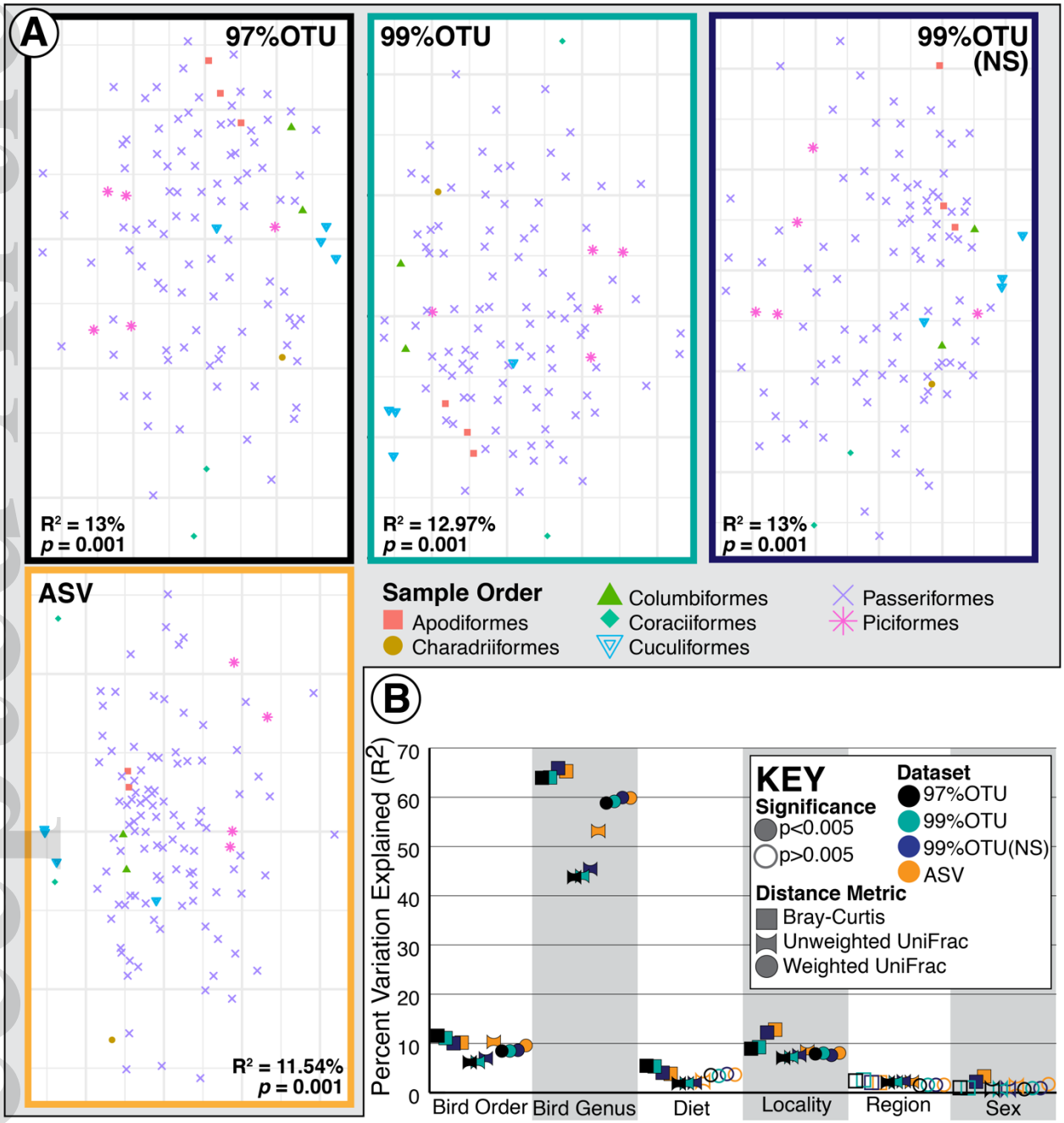

Figure 5: Comparison of sample replicates and PCR replicates. (A) Taxonomic composition of replicate samples at the level of bacterial phyla (for ASV data only). (B) Jaccard Indices for sample and PCR replicates across all four datasets (97OTU, 99OTU, 99OTU(NS), ASV); *p*-values refer to the Wilcoxon non-parametric test for difference in the means.

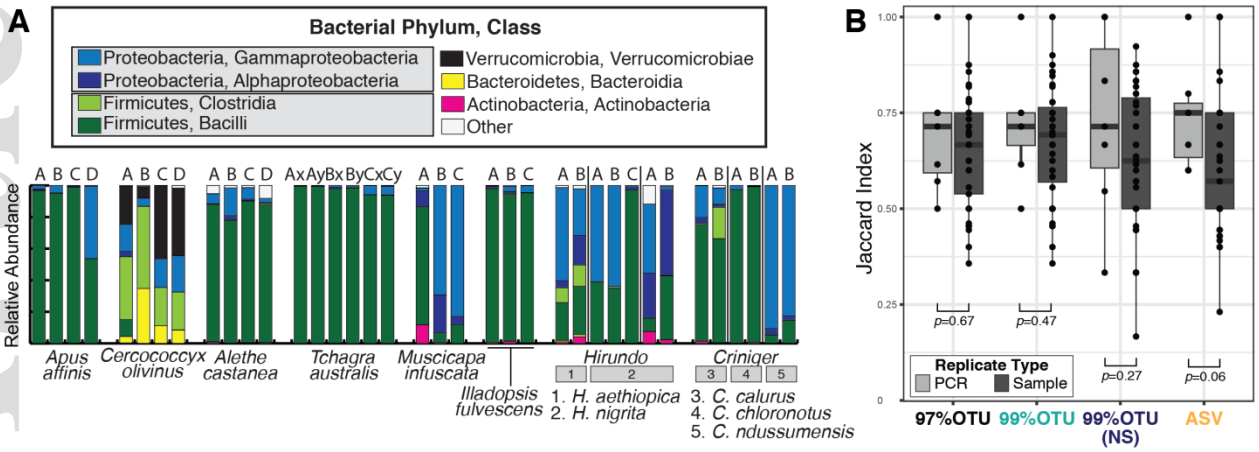

Figure 6: Phylogenetic comparative methods. (A) Host phylogeny colored by the average relative abundance of the phylum Firmicutes. (B) Weighted AICs for the four evolutionary models and estimates of lambda across datasets for the relative abundances of six abundant bacterial phyla (Proteobacteria, Firmicutes, Actinobacteria, Verrucomicrobia, Spirochaetes), an overall measure of alpha diversity ("Observed OTUs") and two beta diversity metrics (coordinates on NMDS axes). Bird weight is included as a proof of concept; simulated data included a simulation of Brownian motion ("Simulated BM"), which was also subsequently randomized ("Simulated BM randomized") and a simulated White Noise dataset ("Simulated WN"), which was subsequently randomized ("Simulated WN randomized"). Lambda estimates for all datasets are overlaid as red circles.

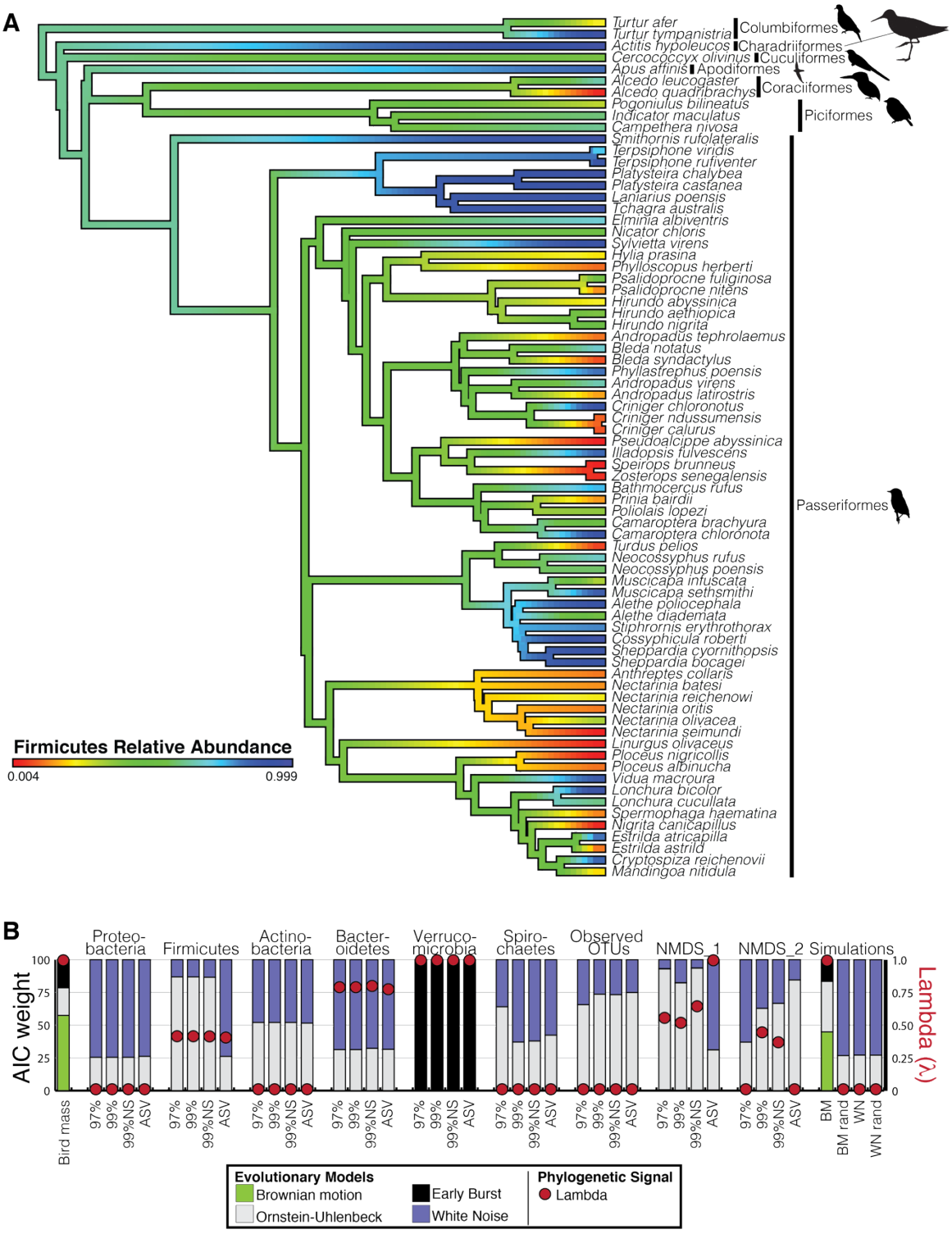

Table 1. Sample information for the 92 individuals in this study. Fig.2 refers to the sample placement on Fig.2 in the manuscript. SampleID is the unique name given the sample, which is a combination of the first three letters of the genus, the first three letters of the species and the last three digits of the LSUMNS B-number given upon deposition to the museum (the full B-number for each specimen is 90XXX). The number of sample replicates (Sample Reps) and PCR replicates (PCR Reps) indicates which individuals had more than one sample taken. Other data are bird order (-iformes has been removed from each name), family, genus, species, sampling locality (see Fig. 1), sampling province, diet and mass (g). Bird taxonomy follows current eBird taxonomy. Genus/species names in parentheses refer to names on Fig. 5 that come from birdtree.org.

Fig.2	SampleID	Sample Reps	PCR Reps	Order	Family	Genus	Species	Locality[1]	Province	Diet[2]	Sex	Mass (g)[3]
1	Turafe469	1		Columb.	Columbidae	*Turtur*	*afer*	UA de GE	Wele-Nzas	O	M	52.1
2	Turtym568	1	2	Columb.	Columbidae	*Turtur*	*tympanistra*	Moka	Bioko	PM	M	67.5
4	Ceroli502	4		Cucul.	Cuculidae	*Cercococcyx*	*olivinus*	UA de GE	Wele-Nzas	AM	F	49.3
8	Apuaff412	4		Apod.	Apodidae	*Apus*	*affinis*	UA de GE	Wele-Nzas	AM	F	22.6
12	Acthyp485	1		Charadri.	Scolopacidae	*Actitis*	*hypoleucos*	UA de GE	Wele-Nzas	AM	M	39.2
13	Alcqua402	1		Coraci.	Alcedinidae	*Alcedo*	*quadribrachys*	UA de GE	Wele-Nzas	AM	M	29.2
14	Corleu653	2		Coraci.	Alcedinidae	*Corythornis (Alcedo)*	*leucogaster*	Moaba Camp	Bioko	AM	M	18.6
16	Indmac481	2		Pic.	Indicatoridae	*Indicator*	*maculatus*	UA de GE	Wele-Nzas	O	F	54.9
18	Pogbil620	2		Pic.	Lybiidae	*Pogoniulus*	*bilineatus*	Moka	Bioko	PM	M	12.2
20	Camniv473	1		Pic.	Picidae	*Campethera*	*nivosa*	UA de GE	Wele-Nzas	AM	M	32.6
21	Camniv655	2		Pic.	Picidae	*Campethera*	*nivosa*	Ureka	Bioko	AM	M	36.4
23	Hylpra468	1		Passer.	Cettiidae	*Hylia*	*prasina*	UA de GE	Wele-Nzas	AM	M	9.5
24	Hylpra472	2		Passer.	Cettiidae	*Hylia*	*prasina*	UA de GE	Wele-Nzas	AM	M	11.1
26	Hylpra545	1		Passer.	Cettiidae	*Hylia*	*prasina*	Moka	Bioko	AM	M	14.5
27	Batruf444	1		Passer.	Cisticolidae	*Bathmocercus*	*rufus*	UA de GE	Wele-Nzas	AM	M	15.8
28	Cambra414	2		Passer.	Cisticolidae	*Camaroptera*	*brachyura*	UA de GE	Wele-Nzas	AM	M	13.7
30	Camchl475	2		Passer.	Cisticolidae	*Camaroptera*	*chloronota*	UA de GE	Wele-Nzas	AM	M	11
32	Camchl574	1	2	Passer.	Cisticolidae	*Camaroptera*	*chloronota*	Moka	Bioko	AM	F	10.9
34	Pollop618	1		Passer.	Cisticolidae	*Poliolais*	*lopezi*	Moka	Bioko	AM	M	12.6
35	Pribai435	2		Passer.	Cisticolidae	*Prinia*	*bairdii*	UA de GE	Wele-Nzas	AM	M	15.7
37	Cryrei531	2		Passer.	Estrildidae	*Cryptospiza*	*reichenovii*	Moka	Bioko	PM	M	13.1
39	Estast533	1		Passer.	Estrildidae	*Estrilda*	*astrild*	Moka	Bioko	O	M	7.7
40	Estatr405	1		Passer.	Estrildidae	*Estrilda*	*atricapilla*	UA de GE	Wele-Nzas	O	F	7.1
41	Estatr417	2		Passer.	Estrildidae	*Estrilda*	*atricapilla*	UA de GE	Wele-Nzas	O	F	7.1
43	Estatr418	2		Passer.	Estrildidae	*Estrilda*	*atricapilla*	UA de GE	Wele-Nzas	O	M	7.3
45	Mannit476	1		Passer.	Estrildidae	*Mandingoa*	*nitidula*	UA de GE	Wele-Nzas	PM	M	10.8
46	Nigcan591	1		Passer.	Estrildidae	*Nigrita*	*canicapillus*	Moka	Bioko	PM	F	18
47	Spebic427	1		Passer.	Estrildidae	*Spermestes (Lonchura)*	*bicolor*	UA de GE	Wele-Nzas	PM	F	9.3
48	Specuc423	1		Passer.	Estrildidae	*Spermestes (Lonchura)*	*cucullata*	UA de GE	Wele-Nzas	PM	F	NR
49	Spehae517	2		Passer.	Estrildidae	*Spermophaga*	*haematina*	UA de GE	Wele-Nzas	O	M	11.6
51	Smiruf474	1		Passer.	Eurylaimidae	*Smithornis*	*rufolateralis*	UA de GE	Wele-Nzas	AM	F	20.4
52	Linoli527	1		Passer.	Fringillidae	*Linurgus*	*olivaceus*	Pico Basile	Bioko	PM	M	20.7
53	Cecaby408	1		Passer.	Hirundinidae	*Cecropis (Hirundo)*	*abyssinica*	UA de GE	Wele-Nzas	O	M	19.7
54	Hiraet410	2		Passer.	Hirundinidae	*Hirundo*	*aethiopica*	UA de GE	Wele-Nzas	AM	F	13
56	Hirnig438	3		Passer.	Hirundinidae	*Hirundo*	*nigrita*	UA de GE	Wele-Nzas	AM	M	15.6
59	Hirnig439	2		Passer.	Hirundinidae	*Hirundo*	*nigrita*	UA de GE	Wele-Nzas	AM	M	16.7
61	Psaful650	1		Passer.	Hirundinidae	*Psalidoprocne*	*fuliginosa*	Moka	Bioko	PM	F	13.3
62	Psanit510	1		Passer.	Hirundinidae	*Psalidoprocne*	*nitens*	UA de GE	Wele-Nzas	AM	F	10.1
63	Sylvir401	1		Passer.	Macrosphenidae	*Sylvietta*	*virens*	UA de GE	Wele-Nzas	AM	F	7.9
64	Lanpoe578	1		Passer.	Malaconotidae	*Laniarius*	*poensis*	Moka	Bioko	AM	F	44.5
65	Tchaus501	3	2 (ea)	Passer.	Malaconotidae	*Tchagra*	*australis*	UA de GE	Wele-Nzas	AM	M	36.6
71	Terruf536	2		Passer.	Monarchidae	*Terpsiphone*	*rufiventer*	Moka	Bioko	AM	M	20
73	Tervir459	1		Passer.	Monarchidae	*Terpsiphone*	*viridis*	UA de GE	Wele-Nzas	O	F	13.7
74	Tervir470	2		Passer.	Monarchidae	*Terpsiphone*	*viridis*	UA de GE	Wele-Nzas	O	M	NR
76	Alecas432	1		Passer.	Muscicapidae	*Alethe*	*castanea (diademata)*	UA de GE	Wele-Nzas	AM	F	32
77	Alecas634	4		Passer.	Muscicapidae	*Alethe*	*castanea (diademata)*	Ureka	Bioko	AM	F	25.7
81	Chapol456	1		Passer.	Muscicapidae	*Chamaetylas (Alethe)*	*poliocephala*	UA de GE	Wele-Nzas	AM	M	31
82	Chapol535	2		Passer.	Muscicapidae	*Chamaetylas (Alethe)*	*poliocephala*	Moka	Bioko	AM	M	34.8
84	Cosrob544	1		Passer.	Muscicapidae	*Cossyphicula*	*roberti*	Moka	Bioko	O	M	17.6
85	Musinf413	3		Passer.	Muscicapidae	*Muscicapa*	*infuscata*	UA de GE	Wele-Nzas	O	F	19.4
88	Musset445	2		Passer.	Muscicapidae	*Muscicapa*	*sethsmithi*	UA de GE	Wele-Nzas	AM	F	8.2
90	Sheboc537	1		Passer.	Muscicapidae	*Sheppardia*	*bocagei*	Moka	Bioko	AM	F	16.9
91	Shecyo513	1		Passer.	Muscicapidae	*Sheppardia*	*cyornithopsis*	UA de GE	Wele-Nzas	AM	M	17.8
92	Stiery430	1		Passer.	Muscicapidae	*Stiphrornis*	*erythrothorax*	UA de GE	Wele-Nzas	AM	F	15.2
93	Stiery633	2		Passer.	Muscicapidae	*Stiphrornis*	*erythrothorax*	Ureka	Bioko	AM	M	18.5
95	Hedcol505	1		Passer.	Nectariniidae	*Hedydipna (Anthodiaeta)*	*collaris*	UA de GE	Wele-Nzas	PM	F	6

96	Antsei584	1	Passer.	Nectariniidae	*Anthreptes (Nectarinia)*	*seimundi*	Moka	Bioko	PM	F	6.9
97	Cinbat453	1	Passer.	Nectariniidae	*Cinnyris (Nectarinia)*	*batesi*	UA de GE	Wele-Nzas	O	F	6
98	Cinrei524	1	Passer.	Nectariniidae	*Cinnyris (Nectarinia)*	*reichenowi*	Pico Basile	Bioko	O	M	8.9
99	Cyaoli443	1	Passer.	Nectariniidae	*Cyanomitra (Nectarinia)*	*olivacea*	UA de GE	Wele-Nzas	O	M	10
100	Cyaoli440	1	Passer.	Nectariniidae	*Cyanomitra (Nectarinia)*	*olivacea*	UA de GE	Wele-Nzas	O	M	8.9
101	Cyaoli466	1	Passer.	Nectariniidae	*Cyanomitra (Nectarinia)*	*olivacea*	UA de GE	Wele-Nzas	O	M	9.2
102	Cyaori548	1	Passer.	Nectariniidae	*Cyanomitra (Nectarinia)*	*oritis*	Moka	Bioko	O	F	10
103	Nicchl482	1	Passer.	Nicatoridae	*Nicator*	*chloris*	UA de GE	Wele-Nzas	O	M	52.4
104	Illful448	3	Passer.	Pellorneidae	*Illadopsis*	*fulvescens*	UA de GE	Wele-Nzas	AM	F	26.7
107	Phyher625	1	Passer.	Phylloscopidae	*Phylloscopus*	*herberti*	Moka	Bioko	AM	M	9
108	Placas419	1	Passer.	Platysteiridae	*Platysteira*	*castanea*	UA de GE	Wele-Nzas	O	M	12.9
109	Placas599	2	Passer.	Platysteiridae	*Platysteira*	*castanea*	Moka	Bioko	O	F	13.6
111	Placha530	1	Passer.	Platysteiridae	*Platysteira*	*chalybea*	Moka	Bioko	AM	M	10.2
112	Ploalb656	1	Passer.	Ploceidae	*Ploceus*	*albinucha*	Ureka	Bioko	O	M	26.2
113	Plonig575	1	Passer.	Ploceidae	*Ploceus*	*nigricollis*	Moka	Bioko	AM	M	28
114	Aritep580	1	Passer.	Pycnonotidae	*Arizelocichla (Andropadus)*	*tephrolaema (-us)*	Moka	Bioko	O	M	34.1
115	Blenot404	1	Passer.	Pycnonotidae	*Bleda*	*notatus*	UA de GE	Wele-Nzas	AM	F	34.8
116	Blenot640	2	Passer.	Pycnonotidae	*Bleda*	*notatus*	Ureka	Bioko	AM	F	40.3
118	Blesyn457	1	Passer.	Pycnonotidae	*Bleda*	*syndactyla*	UA de GE	Wele-Nzas	AM	F	46
119	Crical458	2	Passer.	Pycnonotidae	*Criniger*	*calurus*	UA de GE	Wele-Nzas	AM	M	27.9
121	Crichl492	2	Passer.	Pycnonotidae	*Criniger*	*chloronotus*	UA de GE	Wele-Nzas	O	F	38.6
123	Crindu461	2	Passer.	Pycnonotidae	*Criniger*	*ndussumensis*	UA de GE	Wele-Nzas	AM	M	27.4
125	Eurlat403	1	Passer.	Pycnonotidae	*Eurillas (Andropadus)*	*latirostris*	UA de GE	Wele-Nzas	O	M	25.8
126	Eurlat534	2	Passer.	Pycnonotidae	*Eurillas (Andropadus)*	*latirostris*	Moka	Bioko	O	F	31.5
128	Eurvir409	2	Passer.	Pycnonotidae	*Eurillas (Andropadus)*	*virens*	UA de GE	Wele-Nzas	O	F	19.6
130	Eurvir564	1	Passer.	Pycnonotidae	*Eurillas (Andropadus)*	*virens*	Moka	Bioko	O	M	25.4
131	Phypoe523	1	Passer.	Pycnonotidae	*Phyllastrephus*	*poensis*	Moka	Bioko	O	M	31.7
132	Elmalb557	1	Passer.	Stenostiridae	*Elminia*	*albiventris*	Moka	Bioko	AM	M	8.5
133	Sylaby539	1	Passer.	Sylviidae	*Sylvia (Pseudoalcippe)*	*abyssinica*	Pico Basile	Bioko	O	M	17.2
134	Neopoe446	1	Passer.	Turdidae	*Neocossyphus*	*poensis*	UA de GE	Wele-Nzas	AM	M	45
135	Neopoe643	2	Passer.	Turdidae	*Neocossyphus*	*poensis*	Ureka	Bioko	AM	F	50
137	Neoruf454	2	Passer.	Turdidae	*Neocossyphus*	*rufus*	UA de GE	Wele-Nzas	AM	M	59
139	Turpel554	1	Passer.	Turdidae	*Turdus*	*pelios*	Moka	Bioko	O	M	55.8
140	Vidmac411	1	Passer.	Viduidae	*Vidua*	*macroura*	UA de GE	Wele-Nzas	O	M	12.9
141	Zosbru658	1	Passer.	Zosteropidae	*Zosterops (Speirops)*	*brunneus*	Pico Basile	Bioko	O	M	15.8
142	Zossen561	1	Passer.	Zosteropidae	*Zosterops*	*senegalensis*	Moka	Bioko	AM	F	9.9

Table 2. Taxonomic and abundance information for the seven genera found in greater than one quarter of the samples.

Phylum	Class	Order	Family	Genus	Abundance	N
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Escherichia/Shigella*	10833	121
Firmicutes	Bacilli	Lactobacillales	Carnobacteriaceae	*Catellicoccus*	58246	91
Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	*Enterococcus*	8080	63
Firmicutes	Bacilli	Bacillales	Staphylococcaceae	*Staphylococcus*	2477	61
Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	*Lactobacillus*	10860	44
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Enterobacter*	1945	41
Firmicutes	Clostridia	Clostridiales	Peptostreptococcaceae	*Clostridium_XI*	7707	40
Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	*Lactococcus*	5249	38