ORIGINAL ARTICLE

WILEY MicrobiologyOpen
Open Access

Comparative metagenomics of the gut microbiota in wild greylag geese (*Anser anser*) and ruddy shelducks (*Tadorna ferruginea*)

Wen Wang[1] (ID) | Sisi Zheng[2] | Laixing Li[2] | Yongsheng Yang[2] | Yingbao Liu[3] | Aizhen Wang[4] | Kirill Sharshov[5] | Yao Li[4]

[1]State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xi'ning, Qinghai, China

[2]Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xi'ning, Qinghai, China

[3]College of Life Science, Yangtze University, Jingzhou, Hubei, China

[4]College of Eco-Environmental Engineering, Qinghai University, Xi'ning, Qinghai, China

[5]Research Institute of Experimental and Clinical Medicine, Novosibirsk, Russia

Correspondence
Wen Wang, State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xi'ning, Qinghai, China.
Email: 007CELL@163.com
and
Laixing Li, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xi'ning, Qinghai, China.
Email: lxli@nwipb.cas.cn

Abstract

Gut microbiome contributes to host health by maintaining homeostasis, increasing digestive efficiency, and facilitating the development of immune system. Wild greylag geese (*Anser anser*) and ruddy shelducks (*Tadorna ferruginea*), migrating along the central Asian flyway, appear to be one of the most popular species in the rare birds rearing industries of China. However, the structure and function of the gut microbial communities associated with these two bird species remain poorly understood. Here, for the first time, we compared gut metagenomes from greylag geese to ruddy shelducks and investigated the similarities and differences between these two bird species in detail. Taxonomic classifications revealed the top three bacterial phyla, *Firmicutes*, *Proteobacteria*, and *Fusobacteria*, in both greylag geese and ruddy shelducks. Furthermore, between the two species, 12 bacterial genera were found to be more abundant in ruddy shelducks and 41 genera were significantly higher in greylag geese. A total of 613 genera (approximately 70%) were found to be present in both groups. Metabolic categories related to carbohydrate metabolism, metabolism of cofactors and vitamins, lipid metabolism, amino acid metabolism, and glycan biosynthesis and metabolism were significantly more abundant in ruddy shelducks, while greylag geese were enriched in nucleotide metabolism and energy metabolism. The herbivorous greylag geese gut microbiota harbored more carbohydrate-active enzymes than omnivorous ruddy shelducks. In our study, a range of antibiotic resistance categories were also identified in the gut microbiota of greylag geese and ruddy shelducks. In addition to providing a better understanding of the composition and function of wild birds gut microbiome, this comparative study provides reference values of the artificial domestication of these birds.

KEYWORDS
antibiotic resistance genes, carbohydrate-active enzymes, greylag geese, gut metagenomes, ruddy shelducks

1 | INTRODUCTION

Animals are colonized by rich and complex communities of microorganisms, both externally (e.g., on skin and feathers) and internally (e.g., in the gastrointestinal and reproductive tracts) (Colston & Jackson, 2016; McFall-Ngai et al., 2013). Advances in next-generation sequencing and bioinformatic technologies permit the study of these microorganisms, their genes, and their metabolites (termed the microbiome) at an unprecedented scale (Eisen, 2015; Jovel et al., 2016). Trillions of microbes inhabit the gastrointestinal tract of animals, forming a dynamic ecological community within the gut, which is termed the "gut microbiome" (Ley et al., 2008; Lloyd-Price, Abu-Ali, & Huttenhower, 2016). A wealth of studies have shown that gut microbiome plays an important role in several fundamental and crucial processes in humans and other animal hosts, such as development (Malmuthuge, Griebel, & Guan, 2015), immune homeostasis (Ahern, Faith, & Gordon, 2014), nutrient assimilation (Kau, Ahern, Griffin, Goodman, & Gordon, 2011), vitamins synthesis and sterols metabolism (O'Mahony, Clarke, Borre, Dinan, & Cryan, 2015), and diseases (e.g., obesity, diabetes, and cancer) (Kinross, Darzi, & Nicholson, 2011; Lee & Hase, 2014). Given these important findings, many gut microbiome projects have been launched in several countries (Pylro, Mui, Rodrigues, Andreote, & Roesch, 2016; Stulberg et al., 2016). However, most of these projects and studies describe the microbiota of humans and some mammalian animals; a major gap identified was that there was no project for the vast majority of ecologically relevant taxa, birds.

Compared to other mammalian vertebrates, several characteristics make birds some of the most interesting and useful models for studying the gut microbiome. First, bird brood parasites lay their eggs in the nests of appropriate brood hosts, thus offering a unique and powerful model to investigate the influence of genetic and environmental factors on the colonizing process of gut microbiota (Hird, Carstens, Cardiff, Dittmann, & Brumfield, 2014). Second, whereas mammals acquire important maternal microbes during the birth process, many birds regurgitate food to their young, providing a mode of vertical transmission of gut microbiome across generations (Godoy-Vitorino et al., 2010; Putignani, Del Chierico, Petrucca, Vernocchi, & Dallapiccola, 2014). Third, birds possess a cloaca, serving dual functions for excretion and sexual copulation. Thus, the gastrointestinal tract microbiota of birds provide another avenue for exploring the potential exchange of components of the endogenous microbiome during reproduction (Kreisinger, Cizkova, Kropackova, & Albrecht, 2015).

Currently, there are two main strategies for the analysis of gut microbiome using next-generation sequencing, shotgun metagenomics, and 16S rRNA gene sequencing. The metagenomics approach, in which all the DNA fragments in a sample are sequenced rather than only 16S rRNA amplicons, results in greater in-depth coverage and more informative sequencing datasets (Turaev & Rattei, 2016). Analyses of these datasets will help to elucidate the composition of microbial communities and are valuable resources for identifying carbohydrate-active enzymes and antibiotics resistance genes present in gut communities. To the best of our knowledge, only very limited metagenomic analyses of the functional aspects of avian gut microbiota have been reported (Danzeisen, Kim, Isaacson, Tu, & Johnson, 2011; Godoy-Vitorino et al., 2012; Lu, Santo Domingo, & Shanks, 2007; Wang, Zheng, et al., 2017; Wang, Song, et al., 2017). A meta-analysis of these studies showed that the gut microbiota of birds were dominated by four major phyla, *Firmicutes*, *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes* (Waite & Taylor, 2014, 2015).

In this study, metagenomic sequencing was performed to compare the gut microbial compositions and functions of two bird species, greylag geese (*Anser anser*) and ruddy shelducks (*Tadorna ferruginea*). These birds belong to the same family (Anatidae) and are two ecologically and economically important waterfowl. greylag geese and ruddy shelducks have worldwide distributions and migrate along the central Asian flyway between their breeding and wintering areas (Takekawa et al., 2013). In addition, these two species are artificially reared in several provinces of China to meet market demands (e.g., meat and eggs) and for conservation purposes. The results of this study provide a deeper exploration of the gut microbiomes of wild geese and ducks and may provide useful information for the further application of probiotic strains (isolated from the feces of wild geese and ducks) in the artificial rearing of these birds.

2 | MATERIALS AND METHODS

2.1 | Ethics statement

This study conformed to the guidelines for the care and use of experimental animals established by the Ministry of Science and Technology of the People's Republic of China (Approval number: 2006-398). The research protocol was reviewed and approved by the Ethical Committee of Qinghai University. In this study, only feces of greylag geese and ruddy shelducks were collected for relevant molecular studies. No direct capture or hunting involved.

2.2 | Sampling

Fecal samples were taken from three wild greylag geese (in text abbreviation: GG group) and three wild ruddy shelducks (in text abbreviation: RSD group) from the Gengga-hai Lake (N36°11′59.8″ E100°05′39.7″, elevation 2,800 m) on the northeastern Qinghai-Tibet Plateau, Qinghai Province, China. We chose the sampling sites utilized by these birds as stopover in autumn during migration and waited when they were foraging in the farmlands where only a single species was present in the morning. Fresh fecal material was collected and stored in sterile tubes. The fecal samples were collected at a minimum distance interval of five meters to ensure that all fresh droppings were expelled from different individuals. All samples were transported to the laboratory using a −20°C portable freezer and stored at −80°C until further treatment.

2.3 | DNA extraction and metagenomic sequencing

Metagenomic DNA was isolated from approximately 1 g of fecal sample, using the E.Z.N.A.® stool DNA Kit (Omega Bio-tek, Norcross, GA, USA) following the manufacturer's instruction. DNA concentration and quality were assessed by Qubit fluorometer and agarose gel electrophoresis, respectively. The high-quality DNA was then used to create an Illumina DNA library and sequenced using Illumina NovaSeq (2 × 125 bp) (Illumina, USA) platform at Novogene Bioinformatics Technology Co. Ltd (Beijing, China).

2.4 | Sequence analyses and metagenome assembly

To obtain the clean data for subsequent analysis, the raw data from the Illumina NovaSeq sequencing platform were processed using Readfq (V8, https://github.com/cjfields/readfq). The specific processing steps were as follows: (a) removed reads which contain low-quality bases (default quality threshold value less than or equal to 38) above a certain portion (default length of 40 bp); (b) removed reads in which the N base reached a certain percentage (default length of 10 bp); (c) removed reads which shared the overlap above a certain portion with Adapter (default length of 15 bp). To filter the reads that were of host origin, clean data were then blast against the host database using SoapAligner software (soap2.21, http://soap. genomics.org.cn/soapaligner.html). The parameters were as follows: identity greater than or equal to 90%, -l 30, -v 7, -M 4, -m 200, -x 400 (Law et al., 2013). Then, the high-quality reads of each sample were assembled by the SOAPdenovo software (V2.04, http://soap. genomics.org.cn/soapdenovo.html) (Luo et al., 2012), with the parameters -d 1, -M 3, -R, -u, -F, -K 55 (Qin et al., 2014). After de novo assembly for each sample independently, all reads that not used from all samples were combined and performed mixed assembly in order to maximize the usage of data. Subsequently, we broke the assembled Scaffolds from N connection and obtained the Scaftigs. At last, the fragments longer than 500 bp in all of Scaftigs were used for further analysis.

2.5 | Gene prediction and construction of the nonredundant gene set

We used MetaGeneMark (V2.10, http://topaz.gatech.edu/ GeneMark/) to predict ORFs from the Scaftigs assembled from each sample as well as the Scaftigs from the mixed assembly. Then, the ORFs with length <100 bp were filtered out. For the predicted ORFs, CD-HIT software (V4.5.8, http://www.bioinformatics.org/cd-hit) was used to reduce sequence redundancy (Fu, Niu, Zhu, Wu, & Li, 2012) and the unique initial gene catalogue (the genes here refer to the nucleotide sequences coded by unique and continuous genes) was obtained (Sunagawa et al., 2015). To obtain the gene catalogue (Unigenes) eventually used for subsequently analysis, the clean data of each sample were mapped to the unique initial gene catalogue using SoapAligner (soap 2.21). Based on the number of mapped reads and the length of gene, the abundance information of each Unigene in each sample was statistically analyzed.

2.6 | Gene taxonomic prediction

DIAMOND software (V0.7.9, https://github.com/bbuchfink/diamond/) was used to blast the unigenes to the sequences of bacteria, fungi, archaea, and viruses which were all extracted from the NR database (version: 20161115, https://www.ncbi.nlm.nih.gov/) of NCBI with the parameter of blastp, -e 1e-5 (Buchfink, Xie, & Huson, 2015). To identify bacterial taxa, the lowest common ancestor (LCA) algorithm in MEGAN software was used (Huson, Mitra, Ruscheweyh, Weber, & Schuster, 2011). The exhibition of PCA (R ade4 package, version 2.15.3) (Avershina, Trine, & Knut, 2013) and NMDS (R vegan package, version 2.15.3) (Rivas et al., 2013) decrease-dimension analyses was based on the abundance of each taxonomic hierarchy. Analysis of similarities (ANOSIM) was performed using the vegan package in R (version 2.15.3). Metastats analyses were used to look for the different species between groups. Venn diagram analyses at the genus level were performed using VennDiagram package in the software R (2.15.3).

2.7 | Functional gene annotation

Functional annotation of metagenomes was conducted using DIAMOND software (V0.7.9) to blast unigenes to KEGG database (version 201609, http://www.kegg.jp/kegg/) (Kanehisa et al., 2014), CAZy database (version 20150704, http://www.cazy.org/) (Cantarel et al., 2009), and the Comprehensive Antibiotic Resistance Database (https://card.mcmaster.ca/) with the parameter setting of blastp, evalue ≤ 1e - 30 (Jia et al., 2017). For each sequence's blast result, the best Blast Hit was used for subsequent analysis. The relative abundance of each functional hierarchy equaled the sum of relative abundance annotated to that functional level. Based on the abundance of each hierarchy, heat map of hierarchy cluster and Metastats analysis were performed. Permutation test between groups was used in Metastats analysis, and the p value was corrected by controlling the false discovery rate using the Benjamini–Hochberg method (White, Nagarajan, & Pop, 2009).

3 | RESULTS

3.1 | Summary of the metagenomic datasets

A total of six metagenomes were sequenced using an Illumina's NovaSeq platform (2 × 150 bp). The output data encompassed a total of 38,306.65 Mb of raw reads, with an average of 6,384.44 Mb per sample (Supporting Information Table S1). Then, these sequenced raw reads were filtered to obtain clean reads by removing the short and low-quality reads, adaptors and any eukaryotic sequences. As a result, a total of 36,473.30 Mb of clean reads were generated (Supporting Information Table S1) for further clear assembly and annotation analyses. The de novo assembly of these clean

reads resulted in a total of 320.93 Mb of scaftigs (with the total length ranging from 14.49 to 117.78 Mb), with an N50 of 1150.67 bp (Supporting Information Table S1). Based on these scaftigs, a total of 310,560 unigenes with an average length of 550.27 bp and an average GC content of 42.93% were predicted using the gene-finding algorithm MetaGeneMark (Supporting Information Table S1). These unigenes were then used for taxonomic analysis and functional annotation, and the results of which were summarized in Supporting Information Table S2.

3.2 | Comparison of the composition of the metagenome-based microbial communities in greylag geese and ruddy shelducks

To investigate the gut microbiome compositions of both greylag geese and ruddy shelducks, based on the BLAST results, approximately 150,882.57 unigenes were further classified into phyla and down to species level using the MEGAN software (Supporting Information Table S2). The top four phyla in greylag geese were *Firmicutes* (31.64%), *Proteobacteria* (11.51%), *Tenericutes* (6.34%), and *Fusobacteria* (0.96%) (Figure 1a and Supporting Information Table S3). In the ruddy shelducks group, *Firmicutes* also held the overwhelming predominance, with the average relative abundance of 39.99%, followed by *Bacteroidetes* (15.66%), *Proteobacteria* (3.93%), and *Fusobacteria* (3.40%) (Figure 1a and Supporting Information Table S3). Sequences that could not be classified into any known groups and that were detected with low abundance were grouped as "others." The proportion of these unclassified sequences varied between 11.19% and 83.09% among the different samples (Figure 1a and Supporting Information Table S3). This indicated that a great number of unknown bacteria were present in the guts of these two species of birds. The most abundant genus was *Streptococcus* (7.25%), followed by *Escherichia* (5.72%), *Mycoplasma* (5.34%), *Romboutsia* (5.27%), and *Staphylococcus* (2.60%) in greylag geese (Figure 1b and Supporting Information Table S4). In contrast, the dominant microbial genera in ruddy shelducks were *Enterococcus* (16.70%), *Bacteroides* (10.87%),
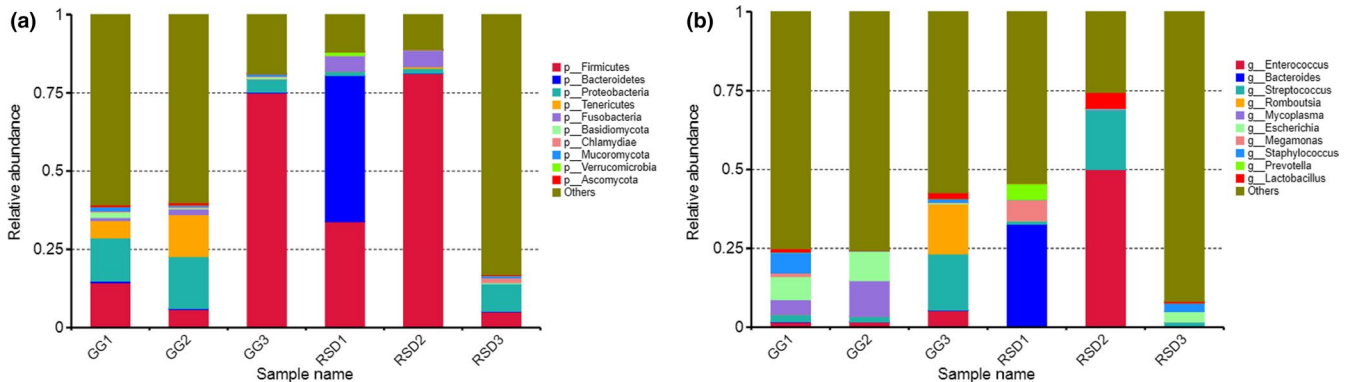

FIGURE 1 Taxonomic profiles of the microbial communities at the phylum level (a) and genus level (b) in each sample. GG refers to the greylag geese group; RSD refers to the ruddy shelducks group
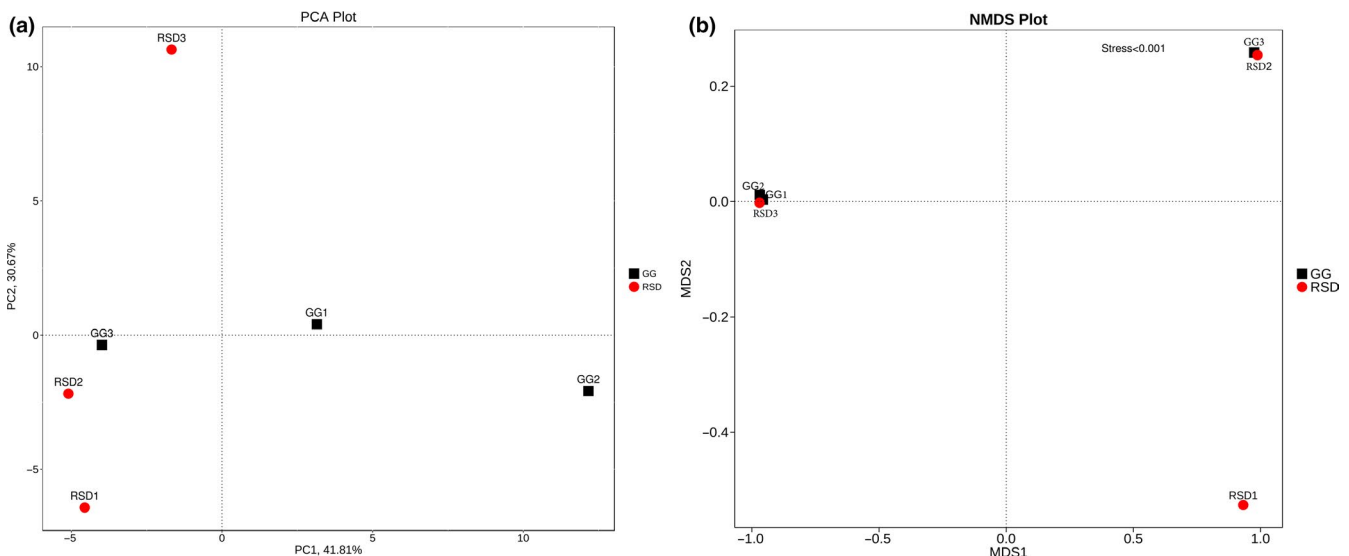

FIGURE 2 The beta diversity results of PCA plot (a) and NMDS plot (b) indicating the microbial phyla distribution between the groups. GG refers to the greylag geese group samples; RSD refers to the ruddy shelducks group samples

Streptococcus (7.17%), *Megamonas* (2.22%), and *Lactobacillus* (1.85) (Figure 1b and Supporting Information Table S4).

Differences in bacterial community composition between the two groups were estimated using principal component analysis (PCA) and nonmetric multidimensional scaling (NMDS) (Figure 2). At the phylum level, greylag geese populations were observed to demonstrate a visually distinct profile from the ruddy shelducks. These dissimilarities in microbial communities were further confirmed by performing genus-level PCA and NMDS calculations (Supporting
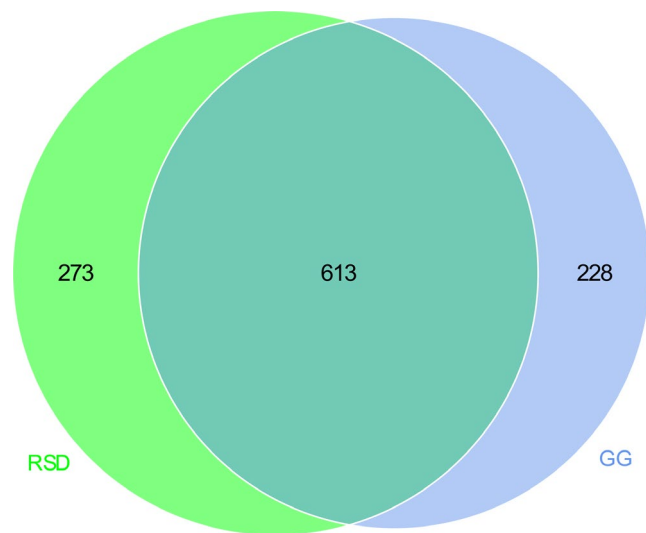
Information Figure S1). Furthermore, analysis of similarities (ANOSIM) was performed to determine whether the total variation in the gut microbiome was due to the differences within or between groups. The results showed that differences observed within group outweighed that occurring between groups (Supporting Information Figure S2). These interindividual variances may be partially due to the unequal and relatively small sample size of this study.

Metastats analyses were performed to detect differentially abundant genera between the two groups (Supporting Information Table S5). A total of 12 genera were found to be more abundant in RSD group and 41 genera were significantly higher in GG group (Supporting Information Table S5). A Venn diagram showed that 613 genera were assigned to both groups, while 273 and 228 bacterial genera were only assigned to the RSD and GG groups, respectively (Figure 3).

3.3 | Functional profiling of the gut metagenome

Metagenome sequencing has the inherent advantage of allowing for an examination of gene content of the microbial populations and allows the direct inference of the metabolic capacity of these populations. To explore the overall functional profiles of the gut microbiome, a total of 141,279 out of 310,560 unigenes (45.49%) were identified by the KEGG database (Supporting Information Table S2). Of these unigenes, 82,371 (26.52%) could be assigned to 7,021 KEGG ortholog group (KOs) and 49,795 (16.03%) could be assigned to 384 KEGG pathways (Supporting Information Table S2). The unigenes matching to level 1 and level 2 KEGG functional categories were shown in Figure 4. In detail, the dominant functional categories

FIGURE 3 Venn diagrams showing the unique and shared microbial genera between Group GG and Group RSD. GG refers to the greylag geese group; RSD refers to the ruddy shelducks group
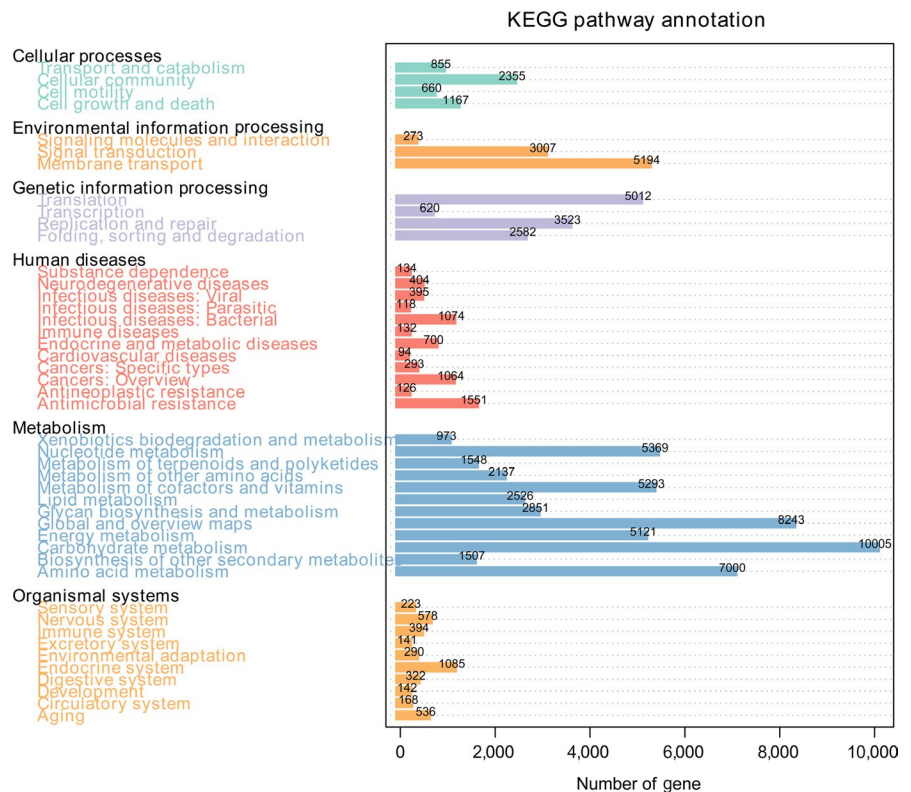

FIGURE 4 Summary of unigenes matched to each KEGG functional categories (level 1 and level 2) present in the gut metagenome datasets

identified included metabolism (59.89%), genetic information processing (13.37%), and environmental information processing (9.65%). The cumulative proportion of these three categories was above 82.91% for all assigned unigenes. This indicated that the metabolic potential of the gut microbiota related to these two species of birds was highly active. The proportions of KEGG categories (level 1) in each sample were shown in Figure 5.

3.4 | Comparison of functionality of the greylag geese and ruddy shelducks gut metagenomes

Due to differences in genetics and diet compositions, significantly different abundances of functional genes and pathways were expected to be observed in greylag geese and ruddy shelducks. A heat map of a hierarchical clustering analysis of the top 35 abundant KEGG functional categories (level 2) showed that increased metabolic categories involved in carbohydrate metabolism, metabolism of cofactors and vitamins, lipid metabolism, amino acid metabolism, and glycan biosynthesis and metabolism in RSD group, whereas the GG group were enriched in nucleotide metabolism and energy metabolism (Figure 6). Furthermore, among the detected 384 KEGG pathways (level 3), nine pathways were found to be significantly different between GG and RSD groups through a Metastats analysis (Supporting Information Table S6).

3.5 | Diversity profile of CAZymes

Carbohydrate-active enzymes (CAZymes), encoded by gut microbes, play a crucial role in the breakdown of complex dietary carbohydrates into components that can be absorbed by the host intestinal epithelium. To determine the gut CAZymes profiles, we performed a CAZymes analysis using the metagenomic data from both greylag geese and ruddy shelducks. In total, 9,060 putative genes were identified (Supporting Information Table S2). The majority of these genes identified were assigned to glycoside hydrolases (5,491, 57.78%). In the RSD group, a total of 210 CAZymes were identified, including five auxiliary activities (AAs), 43 carbohydrate binding modules (CBMs), 12 carbohydrate esterases (CEs), 89 glycoside hydrolases (GHs), 52 glycosyl transferases (GTs), and nine polysaccharide lyases (PLs). A total of 196 CAZymes were found in the GG group, including five AAs, 38 CBMs, 10 CEs, 78 GHs, 58 GTs, and seven PLs. The proportions of each CAZymes were compared using Metastats analysis. GH24, 99, 104 and GT33, 72, 73, 80 were found to be significantly higher in GG group ($p < 0.05$) than that in RSD group, while three CAZymes had higher proportions in RSD group ($p < 0.05$), including GH27, GT10, and CBM20.

3.6 | Antibiotic resistance profiles

The high concentrations of antibiotics used in clinical, agricultural, livestock, and poultry settings provide a strong selective pressure that favors the exchange of antibiotic resistance genes (ARGs) between pathogens and gut microbes. To investigate the ARGs present

in the gut microbiota of greylag geese and ruddy shelducks, the unigenes identified in the metagenome data were screened for antibiotic resistance factors using the Comprehensive Antibiotic Resistance Database (CARD). A total of 598 unigenes were annotated in CARD and 125 antibiotic resistance ontologies (AROs) were identified (Supporting Information Table S2). Of these detected AROs, the top 20 most abundant AROs in each sample were shown in Figure 7. A range of antibiotic resistance categories were also identified, including resistance to rifampin, mupirocin, novobiocin, pulvomycin, fosfomycin, norfloxacin, ciprofloxacin, acriflavine, pleuromutilin, amoxicillin, daptomycin, aminocoumarin, and polymyxin. These AROs were further analyzed for their microbial origin. In the GG group, approximately 50% of the AROs were highly enriched in phyla *Firmicutes* (32%), *Proteobacteria* (12%), and *Tenericutes* (6%) (Figure 8a). In the RSD group, approximately 60% of AROs were found to be enriched in phyla *Firmicutes* (40%), *Bacteroidetes* (16%), *Proteobacteria* (4%), and *Fusobacteria* (3%) (Figure 8b). These results indicate that different types of gut microbes contribute differently to the occurrence of antibiotic resistance genes.

4 | DISCUSSION

The results presented in this study represent the first gut metagenomic characterizations of both greylag geese and ruddy shelducks. To date, although several datasets are available for the gut metagenomes of domestic birds (e.g., chickens and ducks) (Day, Ballard, Duke, Scheffler, & Zsak, 2010; Delforno et al., 2017; Lu et al., 2007), relatively little information is available concerning the metagenomes of wild birds' gut microbiota (Godoy-Vitorino et al., 2008, 2012; Wang, Zheng, et al., 2017; Wang, Song, et al., 2017). Therefore, this study provides a significant contribution to the vast, yet little explored, field of wild birds' gut microbiome. Additionally, knowledge of the gut microbiomes of these two wild bird species could also provide a helpful reference for improving the rearing performance of wild birds in artificial rearing industries.

It was a logical choice to compare greylag geese samples with the ruddy shelducks' gut microbiota, because these two species of birds inhabit the same region and share part of the same food items. Based on the taxonomic classification, the gut microbiotas of greylag geese and ruddy shelducks differed substantially at the phylum and lower taxonomic levels. The reasons for these differences were not clear. Maybe the different genetic backgrounds were responsible for these differences. Support for a host genetic effect on the gut microbiome comes mostly from large comparative studies. For instance, Hird et al. sampled 116 intestines from 59 neotropical bird species and found that host taxonomic categories were most frequently able to significantly explain the most variations in the gut microbiotas of birds (Hird, Sánchez, Carstens, & Brumfield, 2015). In mammals, such studies had also shown that host taxonomy was strongly associated with gut microbiotas (Ley et al., 2008). Additionally, our study identified some common bacterial microbiota present in both greylag geese and ruddy shelducks. At
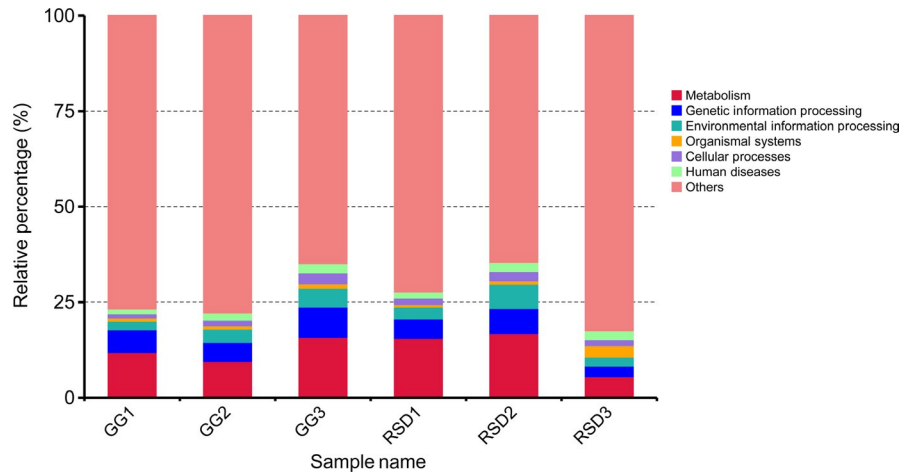
FIGURE 5 Relative abundance of different KEGG functional categories present in each sample. GG refers to the greylag geese group; RSD refers to the ruddy shelducks group
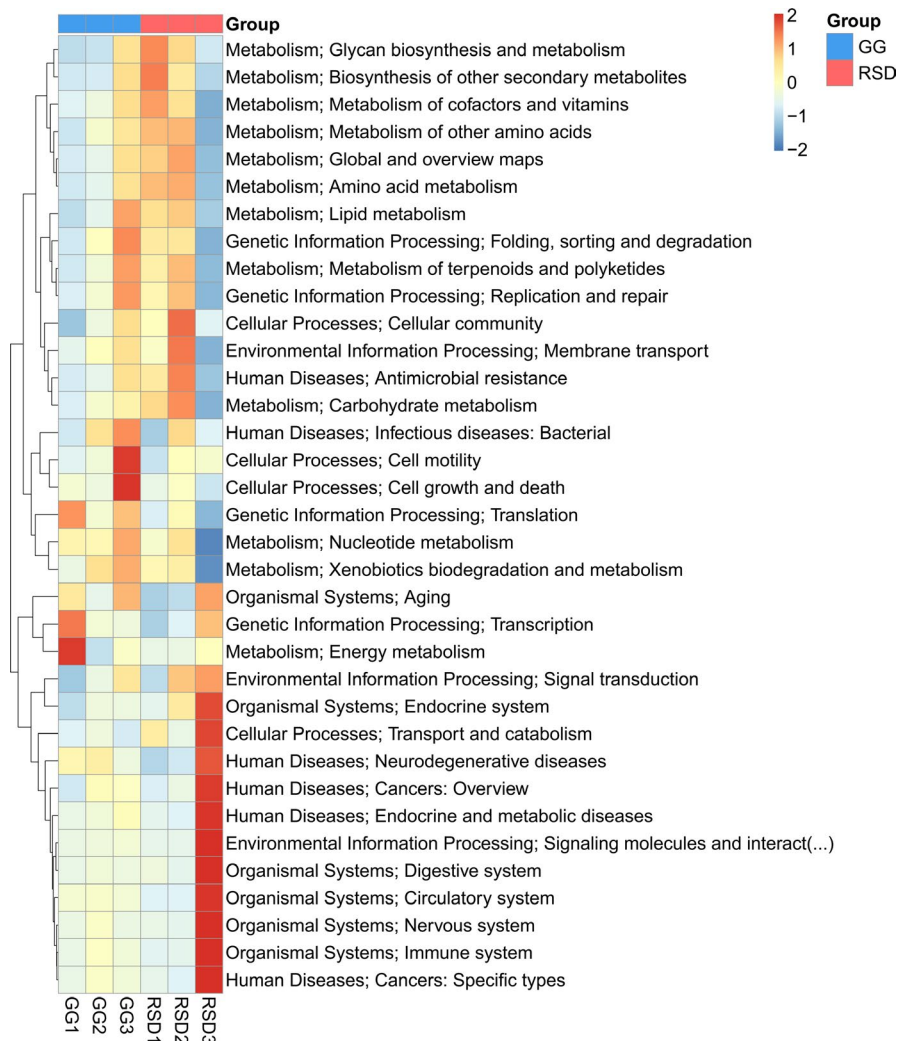

FIGURE 6 Heat map of hierarchical clustering analysis of the top 35 abundant KEGG functional categories (level 2) in Group GG and Group RSD. GG refers to the greylag geese group; RSD refers to the ruddy shelducks group
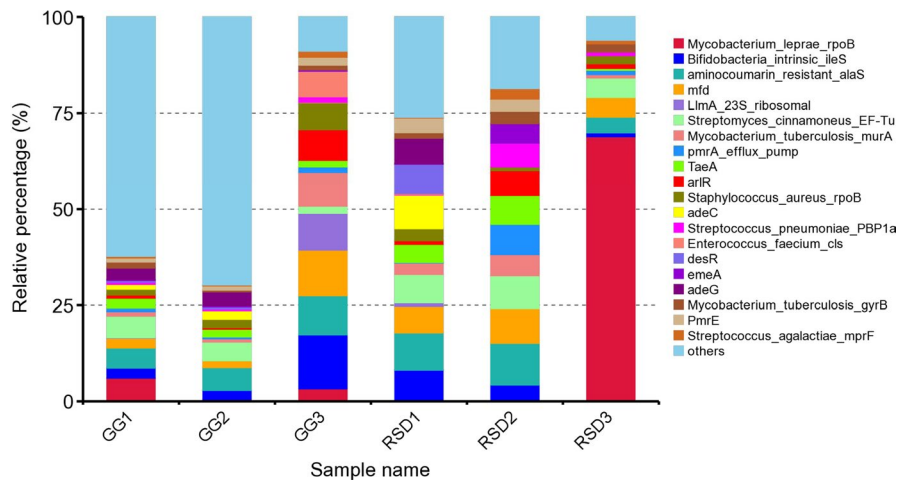
FIGURE 7 Relative proportion of the top 20 most abundant antibiotic resistance ontology in each sample. GG refers to the greylag geese group samples; RSD refers to the ruddy shelducks group samples
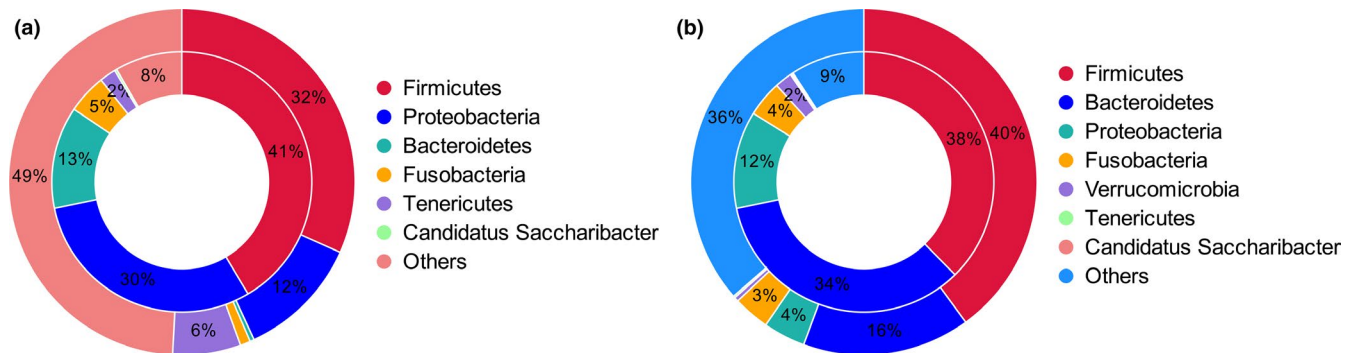

FIGURE 8 Circos plots representing alignment of the proportion of different antibiotic resistance ontology and microbial phyla in Group GG (a) and Group RSD (b). The inner-ring refers to the distribution of different antibiotic resistance ontology in corresponding microbial phyla. The outer-ring refers to the relative abundance of different phyla in each group. GG refers to the greylag geese group; RSD refers to the ruddy shelducks group

the phylum level, _Firmicutes_, _Proteobacteria_, and _Fusobacteria_ were found to be present at high abundance in each sample. _Firmicutes_ and _Proteobacteria_, the most widespread intestinal phyla, are commonly observed within gut environments of many birds (Waite & Taylor, 2014; Wang, Cao, Li, et al., 2016; Wang, Cao, Yang, et al., 2016; Wang, Zheng, et al., 2016). Members of these two phyla were frequently studied for their food digestion roles. For example, _Firmicutes_ members were associated with insoluble fiber degradation (Berry, 2016), and _Proteobacteria_ members were associated with cellulose activity (Reid, Addison, Macdonald, & Lloyd-Jones, 2011). A rich community of _Fusobacteria_ was frequently reported in the guts of carnivorous and omnivorous birds (Waite & Taylor, 2014). In our study, the guts of ruddy shelducks contained 3.4% _Fusobacteria_. As an omnivorous bird, ruddy shelducks feed mainly on fish, shrimps, crabs, aquatic plants, and cereals. The appearance of _Fusobacteria_ in the greylag geese gut microbiome is an interesting avenue for further study, since this species was considered as an herbivorous bird, consuming a diversity of foods that includes leaves, roots, and seeds (Olsen, 2015). At the genus level, the results showed that about 70% of genera were assigned to both greylag geese and ruddy shelducks groups. These common genera may be shaped by the overlapping food items between two groups. Diet is a factor that directly affects

gut microbial community composition (Zarrinpar, Chaix, Yooseph, & Panda, 2014). However, the fundamental characteristics of the relationship between the diets and the wild birds' gut microbiome are unknown. It should be of great research and practical application values to analyze these common genera to develop probiotics that may meet the demands of the artificial rearing industries for greylag geese and ruddy shelducks, helping to accelerate the domestication of these birds.

An Analysis of the overall functional profiles in the present study indicated that the gut microbes associated with these two bird species exhibited high metabolic activities. These results were consistent with earlier studies on the gut metagenomes of bar-headed geese (_Anser indicus_) (Wang, Zheng, et al., 2017; Wang, Song, et al., 2017). This high metabolic rate may be related to the energy consumption required to fulfill the demands of flight. Avian metabolism was reported to be approximately 60% higher than that those of most mammals (Scanes & Braun, 2012). The further comparison of the functional profiles of the datasets from greylag geese and ruddy shelducks revealed many remarkable differences, which suggested that not only the bacterial compositions but their functionalities were important. For example, functions related to carbohydrate metabolism, lipid metabolism, amino acid metabolism, and glycan

biosynthesis and metabolism were significantly more abundant in ruddy shelducks than that in greylag geese. It is reasonable to hypothesize that the microbiota of the omnivorous ruddy shelducks is probably specialized to degrade more diverse types of foods than that of the herbivorous greylag geese, which eats a more homogenous type of food. Several KEGG pathways (level 3) with different abundance dynamics in ruddy shelducks vs. greylag geese appear to be associated with the acclimation of microbiota to the nutrients. For example, we observed increases in the abundance of pentose and glucuronate interconversions [ko00040] in ruddy shelducks that may contribute to host glucose metabolism. We also observed increases in the abundance of pathways associated with limonene and pinene degradation [ko00903] and carotenoid biosynthesis [ko00906] in greylag geese that may contribute to the degradation and fermentation of plant material.

From these metagenome data, we also identified 9,060 putative carbohydrate-active genes. Conforming to expectations, the herbivorous greylag geese gut microbiota harbor more CAZymes than that of the omnivorous ruddy shelducks. These detected CAZymes would allow greylag geese and ruddy shelducks to make extensive use of plant material as a source of nutrients through the enzymatic activities of the gut microbes. The majority of the CAZymes genes identified were assigned to glycoside hydrolases in each sample. GHs are the most abundant enzymes used to break down polysaccharides into smaller products (Berlemont & Martiny, 2016). Of the total detected GHs, GH24 (predominant activity is lysozyme), GH99 (predominant activity is glycoprotein endo-α-1,2-mannosidase), and GH104 (predominant activity is peptidoglycan lytic transglycosylase) were found to be significantly increased in greylag geese group, while the proportion of GH27 (predominant activity is α-galactosidase) was significantly increased in ruddy shelducks group. These increased GHs in each group indicated an enrichment of different gut microbes that were specialized in utilization of diverse plant polysaccharides.

Due to unmonitored application of antibiotics, the widespread of antibiotic resistance genes (ARGs) and antibiotic resistant bacteria has become a great public concern (Wright, 2007). In our study, a range of antibiotic resistance categories were also identified in the gut microbiota of greylag geese and ruddy shelducks. The existence of these antibiotic resistance genes in wild birds supported the view that antibiotic resistance was naturally originated. Soil was increasingly recognized as a vast repository of antibiotic resistance genes (Forsberg et al., 2012). Consequently, it is difficult to find a bird that has never been exposed to antibiotic polluted environments. Based on the analysis of the origin of these antibiotic resistance genes, ARGs profiles were found to be correlated with microbial community compositions. For instance, the majority of ARGs in greylag geese originated from the dominant phyla *Firmicutes* and *Proteobacteria,* and in ruddy shelducks from the phyla *Firmicutes* and *Bacteroidetes*. These phyla, known as prevalent antibiotic producing bacteria, were also reported to be present in the gut microbiota of four songbird species (Carter et al., 2018). These results indicated that highly mobile wildlife, such as migratory birds, might enable the spread of antibiotic resistant bacteria when they are exposed to previously unexposed environments (Viana, Santamaría, & Figuerola, 2016). These results also indicated that manure management was important in wildfowl artificial rearing industries. Thus, it is necessary to reduce potential environmental contamination risks associated with antibiotic resistance from feces of these birds and the abuse of antibiotics.

The present study had several limitations that should be acknowledged. First, the relatively small sample size may reduce the accuracy of partial results. Second, because of feces used in this study were collected from wild birds, some physiological indexes of these birds remain unknown. Therefore, large variances existed between individuals. Lastly, large datasets under captive environments are needed to investigate potential mechanisms driving diets–microbiota interactions in the birds' gut.

In summary, we used metagenomics to gain an insight into both the compositional (profiles of microbiota) and the functional capabilities (KEGG functional categories, carbohydrate-active enzymes, and antibiotic resistance genes) of the gut microbiomes of two wild bird species, greylag geese and ruddy shelducks. By comparing the gut metagenomes of these two species, we also identified both substantial overlap and differences in microbial composition and function. Although it remains unclear to what extent these changes were determined by host genetics and/or diets, these results substantially increased our knowledge of the bird gut microbiome.

AUTHOR CONTRIBUTIONS

Wen Wang and Laixing Li participated in the design of the study, data analysis, and drafted the manuscript; Sisi Zheng and Yongsheng Yang participated in data analysis; Yingbao Liu, Aizhen Wang, Kirill Sharshov, and Yao Li participated in the design of the study and helped draft the manuscript. All authors gave final approval for publication.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DATA ACCESSIBILITY

All the raw metagenome sequences data reported in this paper have been deposited in the Genome Sequence Archive (Wang, Zheng,

et al., 2017; Wang, Song, et al., 2017) in BIG Data Center (BIG Data Center Members, 2018), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession numbers CRA000822 and CRA000823 that are publicly accessible at http://bigd.big.ac.cn/gsa.

ORCID

Wen Wang http://orcid.org/0000-0002-3156-7929

REFERENCES

Ahern, P. P., Faith, J. J., & Gordon, J. I. (2014). Mining the human gut microbiota for effector strains that shape the immune system. *Immunity*, *40*, 815–823. https://doi.org/10.1016/j.immuni.2014.05.012

Avershina, E., Trine, F., & Knut, R. (2013). De novo semi-alignment of 16S rRNA gene sequences for deep phylogenetic characterization of next generation sequencing data. *Microbes and Environments*, *28*, 211–216. https://doi.org/10.1264/jsme2.ME12157

Berlemont, R., & Martiny, A. C. (2016). Glycoside hydrolases across environmental microbial communities. *PLoS Computational Biology*, *12*, e1005300. https://doi.org/10.1371/journal.pcbi.1005300

Berry, D. (2016). The emerging view of Firmicutes as key fibre degraders in the human gut. *Environmental Microbiology*, *18*, 2081–2083. https://doi.org/10.1111/1462-2920.13225

BIG Data Center Members (2018). Database resources of the BIG data center in 2018. *Nucleic Acids Research*, *46*, D14–D20.

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*, 59–60. https://doi.org/10.1038/nmeth.3176

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The carbohydrate-active enzymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research*, *37*, D233–D238. https://doi.org/10.1093/nar/gkn663

Carter, D. L., Docherty, K. M., Gill, S. A., Baker, K., Teachout, J., & Vonhof, M. J. (2018). Antibiotic resistant bacteria are widespread in songbirds across rural and urban environments. *Science of the Total Environment*, *627*, 1234–1241. https://doi.org/10.1016/j.scitotenv.2018.01.343

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*, 3776–3800. https://doi.org/10.1111/mec.13730

Danzeisen, J. L., Kim, H. B., Isaacson, R. E., Tu, Z. J., & Johnson, T. J. (2011). Modulations of the chicken cecal microbiome and metagenome in response to anticoccidial and growth promoter treatment. *PLoS ONE*, *6*, e27949. https://doi.org/10.1371/journal.pone.0027949

Day, J. M., Ballard, L. L., Duke, M. V., Scheffler, B. E., & Zsak, L. (2010). Metagenomic analysis of the turkey gut RNA virus community. *Virology Journal*, *7*, 313. https://doi.org/10.1186/1743-422X-7-313

Delforno, T. P., Lacerda Júnior, G. V., Noronha, M. F., Sakamoto, I. K., Varesche, M. B. A., & Oliveira, V. M. (2017). Microbial diversity of a full-scale UASB reactor applied to poultry slaughterhouse wastewater treatment: Integration of 16S rRNA gene amplicon and shotgun metagenomic sequencing. *Microbiologyopen*, *6*, e00443. https://doi.org/10.1002/mbo3.443

Eisen, J. (2015). What does the term microbiome mean? And where did it come from? A bit of a surprise. *Winnower*, *2*, e142971.16196.

Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O., & Dantas, G. (2012). The shared antibiotic resistome of soil bacteria and human pathogens. *Science*, *337*, 1107–1111. https://doi.org/10.1126/science.1220761

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, *28*, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Godoy-Vitorino, F., Goldfarb, K., Brodie, E., Garcia-Amado, M. A., Michelangeli, F., & Dominguez-Bello, M. G. (2010). Developmental microbial ecology of the crop of the folivorous hoatzin. *ISME Journal*, *4*, 611–620. https://doi.org/10.1038/ismej.2009.147

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. *ISME Journal*, *6*, 531–541. https://doi.org/10.1038/ismej.2011.131

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 Environment Microbiology*, *74*, 5905–5912. https://doi.org/10.1128/AEM.00574-08

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*, e321. https://doi.org/10.7717/peerj.321

Hird, S. M., Sánchez, C., Carstens, B. C., & Brumfield, R. T. (2015). Comparative gut microbiota of 59 neotropical bird species. *Frontiers in Microbiology*, *6*, 1403.

Huson, D. H., Mitra, S., Ruscheweyh, H. J., Weber, N., & Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN4. *Genome Research*, *21*, 1552–1560. https://doi.org/10.1101/gr.120618.111

Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., ... McArthur, A. R. (2017). Card 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *45*, D566–D573. https://doi.org/10.1093/nar/gkw1004

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., ... Wong, G. K. (2016). Characterization of the gut microbiome using 16S or shotgun metagenomics. *Frontiers in Microbiology*, *7*, 459.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2014). Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Research*, *42*, D199–D205. https://doi.org/10.1093/nar/gkt1076

Kau, A. L., Ahern, P. P., Griffin, N. W., Goodman, A. L., & Gordon, J. I. (2011). Human nutrition, the gut microbiome and the immune system. *Nature*, *474*, 327–336. https://doi.org/10.1038/nature10213

Kinross, J. M., Darzi, A. W., & Nicholson, J. K. (2011). Gut microbiome-host interactions in health and disease. *Genome Medicine*, *3*, 14. https://doi.org/10.1186/gm228

Kreisinger, J., Cizkova, D., Kropackova, L., & Albrecht, T. (2015). Cloacal microbiome structure in a long-distance migratory bird assessed using deep 16sRNA pyrosequencing. *PLoS ONE*, *10*, e0137401. https://doi.org/10.1371/journal.pone.0137401

Law, J., Jovel, J., Patterson, J., Ford, G., O'keefe, S., Wang, W., ... Wong, G. K. (2013). Identification of hepatotropic viruses from plasma using deep sequencing: A next generation diagnostic tool. *PLoS ONE*, *8*, e60595. https://doi.org/10.1371/journal.pone.0060595

Lee, W. J., & Hase, K. (2014). Gut microbiota-generated metabolites in animal health and disease. *Nature Chemical Biology*, *10*, 416–424. https://doi.org/10.1038/nchembio.1535

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., ... Gordon, J. I. (2008). Evolution of mammals and their gut microbes. *Science*, *320*, 1647–1651. https://doi.org/10.1126/science.1155725

Lloyd-Price, J., Abu-Ali, G., & Huttenhower, C. (2016). The healthy human microbiome. *Genome Medicine*, *8*, 51. https://doi.org/10.1186/s13073-016-0307-y

Lu, J., Santo Domingo, J., & Shanks, O. C. (2007). Identification of chicken-specific fecal microbial sequences using a metagenomic approach. *Water Research*, *41*, 3561–3574. https://doi.org/10.1016/j.watres.2007.05.033

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., … Wang, J. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience*, *1*, 18. https://doi.org/10.1186/2047-217X-1-18

Malmuthuge, N., Griebel, P. J., & Guan, L. (2015). The gut microbiome and its potential role in the development and function of newborn calf gastrointestinal tract. *Frontiers in Veterinary Science*, *2*, 36.

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., … Wernegreen, J. J. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 3229–3236. https://doi.org/10.1073/pnas.1218525110

Olsen, A. M. (2015). Exceptional avian herbivores: Multiple transitions toward herbivory in the bird order Anseriformes and its correlation with body mass. *Ecology and Evolution*, *5*, 5016–5032. https://doi.org/10.1002/ece3.1787

O'Mahony, S. M., Clarke, G., Borre, Y. E., Dinan, T. G., & Cryan, J. F. (2015). Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behavioral Brain Research*, *277*, 32–48. https://doi.org/10.1016/j.bbr.2014.07.027

Putignani, L., Del Chierico, F., Petrucca, A., Vernocchi, P., & Dallapiccola, B. (2014). The human gut microbiota: A dynamic interplay with the host from birth to senescence settled during childhood. *Pediatric Research*, *76*, 2–10. https://doi.org/10.1038/pr.2014.49

Pylro, V. S., Mui, T. S., Rodrigues, J. L., Andreote, F. D., & Roesch, L. F. & Working Group Supporting the INCT Microbiome (2016). A step forward to empower global microbiome research through local leadership. *Trends in Microbiology*, *24*, 767–771. https://doi.org/10.1016/j.tim.2016.07.007

Qin, N., Yang, F., Li, A., Prifti, E., Chen, Y., Shao, L., … Li, L. (2014). Alterations of the human gut microbiome in liver cirrhosis. *Nature*, *513*, 59–64. https://doi.org/10.1038/nature13568

Reid, N. M., Addison, S. L., Macdonald, L. J., & Lloyd-Jones, G. (2011). Biodiversity of active and inactive bacteria in the gut flora of wood-feeding huhu beetle larvae (*Prionoplus reticularis*). *Applied and Environment Microbiology*, *77*, 7000–7006.

Rivas, M. N., Burton, O. T., Wise, P., Zhang, Y., Hobson, S. A., Lloret, M. G., … Chatila, T. A. (2013). A microbiota signature associated with experimental food allergy promotes allergic sensitization and anaphylaxis. *The Journal of Allergy and Clinical Immunology*, *131*, 201–212. https://doi.org/10.1016/j.jaci.2012.10.026

Scanes, C. G., & Braun, E. (2012). Avian metabolism: Its control and evolution. *Frontiers in Biology*, *8*, 134–159.

Stulberg, E., Fravel, D., Proctor, L. M., Murray, D. M., LoTempio, J., Chrisey, L., … Records, A. (2016). An assessment of US microbiome research. *Nature Microbiology*, *1*, 15015. https://doi.org/10.1038/nmicrobiol.2015.15

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., … Bork, P. (2015). Structure and function of the global ocean microbiome. *Science*, *348*, 1261359. https://doi.org/10.1126/science.1261359

Takekawa, J. Y., Prosser, D. J., Collins, B. M., Douglas, D. C., Perry, W. M., Yan, B., … Newman, S. H. (2013). Movements of wild ruddy shelducks in the Central Asian Flyway and their spatial relationship to outbreaks of highly pathogenic avian influenza H5N1. *Viruses*, *5*, 2129–2152. https://doi.org/10.3390/v5092129

Turaev, D., & Rattei, T. (2016). High definition for systems biology of microbial communities: Metagenomics gets genome-centric and strain-resolved. *Current Opinion in Biotechnology*, *39*, 174–181. https://doi.org/10.1016/j.copbio.2016.04.011

Viana, D. S., Santamaría, L., & Figuerola, J. (2016). Migratory birds as global dispersal vectors. *Trends in Ecology & Evolution*, *31*, 763–775. https://doi.org/10.1016/j.tree.2016.07.005

Waite, D. W., & Taylor, M. W. (2014). Characterizing the avian gut microbiota: Membership, driving influences, and potential function. *Frontiers in Microbiology*, *5*, 223.

Waite, D. W., & Taylor, M. W. (2015). Exploring the avian gut microbiota: Current trends and future directions. *Frontiers in Microbiology*, *6*, 673.

Wang, W., Cao, J., Li, J. R., Yang, F., Li, Z., & Li, L. X. (2016). Comparative analysis of the gastrointestinal microbial communities of bar-headed goose (*Anser indicus*) in different breeding patterns by high-throughput sequencing. *Microbiological Research*, *182*, 59–67. https://doi.org/10.1016/j.micres.2015.10.003

Wang, W., Cao, J., Yang, F., Wang, X., Zheng, S., Sharshov, K., & Li, L. (2016). High-throughput sequencing reveals the core gut microbiome of Bar-headed goose (*Anser indicus*) in different wintering areas in Tibet. *Microbiologyopen*, *5*, 287–295. https://doi.org/10.1002/mbo3.327

Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., … Zhao, W. (2017). GSA: Genome sequence archive. *Genomics Proteomics Bioinformatics*, *15*, 14–18. https://doi.org/10.1016/j.gpb.2017.01.001

Wang, W., Zheng, S., Sharshov, K., Cao, J., Sun, H., Yang, F., … Li, L. (2016). Distinctive gut microbial community structure in both the wild and farmed Swan goose (*Anser cygnoides*). *Journal of Basic Microbiology*, *56*, 1299–1307. https://doi.org/10.1002/jobm.201600155

Wang, W., Zheng, S., Sharshov, K., Sun, H., Yang, F., Wang, X., … Xiao, Z. (2017). Metagenomic profiling of gut microbial communities in both wild and artificially reared Bar-headed goose (*Anser indicus*). *MicrobiologyOpen*, *6*, e429.

White, J. R., Nagarajan, N., & Pop, M. (2009). Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Computational Biology*, *5*, e1000352.

Wright, G. D. (2007). The antibiotic resistome: The nexus of chemical and genetic diversity. *Nature Reviews Microbiology*, *5*, 175–186. https://doi.org/10.1038/nrmicro1614

Zarrinpar, A., Chaix, A., Yooseph, S., & Panda, S. (2014). Diet and feeding pattern affect the diurnal dynamics of the gut microbiome. *Cell Metabolism*, *20*, 1006–1017. https://doi.org/10.1016/j.cmet.2014.11.008

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.