Supplementary Materials: Composition and Potential Function of Fecal Bacterial Microbiota from Six Bird Species

Jose F. Garcia-Mazcorro, Cecilia Alanis-Lopez, Alicia G. Marroquin-Cardona and Jorge R. Kawas

1. The Case of Cyanobacteria

Cyanobacteria (formerly known as blue-green algae) is the name of a bacterial group mostly comprising microorganisms that obtain their energy through photosynthesis [1].

Here we briefly discuss the possibility of having autochthonous Cyanobacteria in the gut of humans and animals and the removal of Cyanobacteria from 16S rRNA gene datasets.

Chloroplasts were at some point free-living microorganisms related to Cyanobacteria [2] and therefore the 16S rRNA gene nucleotide composition of Cyanobacteria and chloroplasts display a high degree of similarity [3,4]. In order to minimize the presence of contaminants, most researchers choose to remove all Cyanobacteria 16S sequences from 16S sequencing surveys, even in situations where Cyanobacteria may be of interest for the scientific community. As mentioned above, Cyanobacteria mostly comprise photosynthetic bacteria; however, Cyanobacteria also comprises non-photosynthetic bacteria, for example Melainabacteria [5,6].

The Ribosomal Database Project (https://rdp.cme.msu.edu/) contains 26,471 16S sequences (last accessed on June 4th, 2020) from the Cyanobacteria/Chloroplast phylum (i.e. for RDP they represent the same group). The use of the greengenes database in our study revealed information about Cyanobacteria and two more unrelated taxa: Chlorobi [7] and Chloroflexi [8]. However, in our study only Cyanobacteria showed measurable amounts (Chlorobi and Chloroflexi were present in only one sample, respectively, at <0.01% relative abundance). These thoughts are important because one paper related to the obese human gut microbiota published by Ley et al. [9] showed evidence to suggest the presence of a "deep-branching clade of the Cyanobacteria in the guts of mice and other animals". To explain this, the authors proposed that "this group may represent descendants of non-photosynthetic ancestral cyanobacteria that have adapted to life in animal gastrointestinal tracts". Interestingly, there are publications that have not removed Cyanobacteria from the sequencing results [10,11] but the reasons behind this decision were not discussed.

We used the unfiltered OTU table from the open OTU picking approach to calculate the relative abundance of all bacterial groups, including Cyanobacteria. Figure S1 shows the relative abundance of 16S rRNA gene sequences from Cyanobacteria in all the six bird species. We reasoned that if these sequences were truly dietary contaminants, then one would expect similar abundances in birds consuming similar diets. However, this was not the case. For instance, samples from *N. hollandicus* (NH) and *P. krameri* had similar diets, yet NH showed very little abundance of Cyanobacteria.
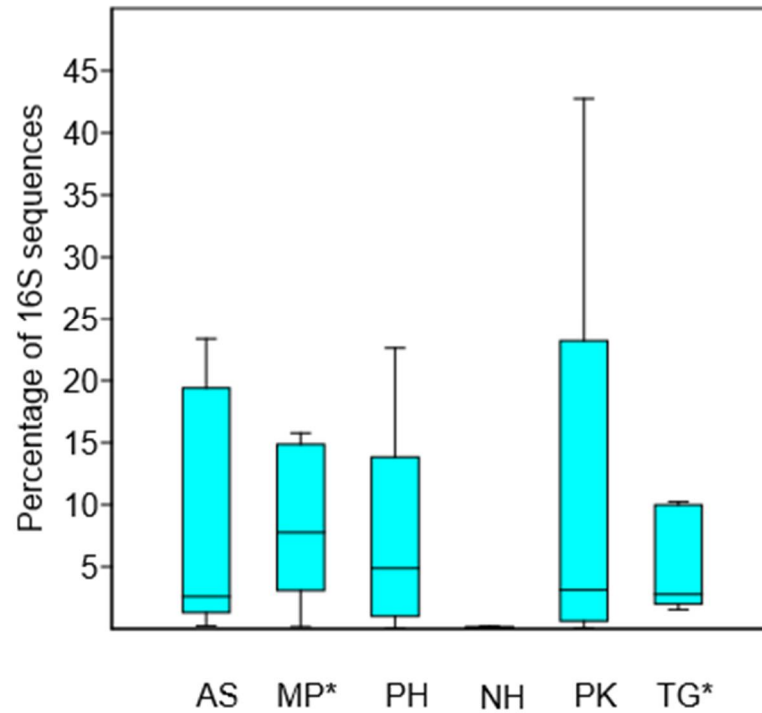
Figure S1. Box plots showing relative proportions of 16S reads from Cyanobacteria among all six bird species. Samples from Passeriformes are highlighted (*). AS: *Agaporni* spp., MP: *Mimus polyglottos*, PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameri*, TG: *Taeniopygia guttata*. Although the observed difference in the abundance of Cyanobacteria did not reach statistical significance in a Kruskal-Wallis test (p = 0.07), multiple comparisons using the Mann-Whitney test revealed significant differences between NH and AS (p = 0.01), between NH and MP (p = 0.02), and between NH and TG (p = 0.01).

2. LEfSe Analyses

The linear discriminant analysis (LDA) effect size (LEfSe) was used to find organisms that consistently explain the differences between microbial communities among the bird species, as explained in the main text. In this study, each bird species harbored specific bacterial communities (Figure S2).
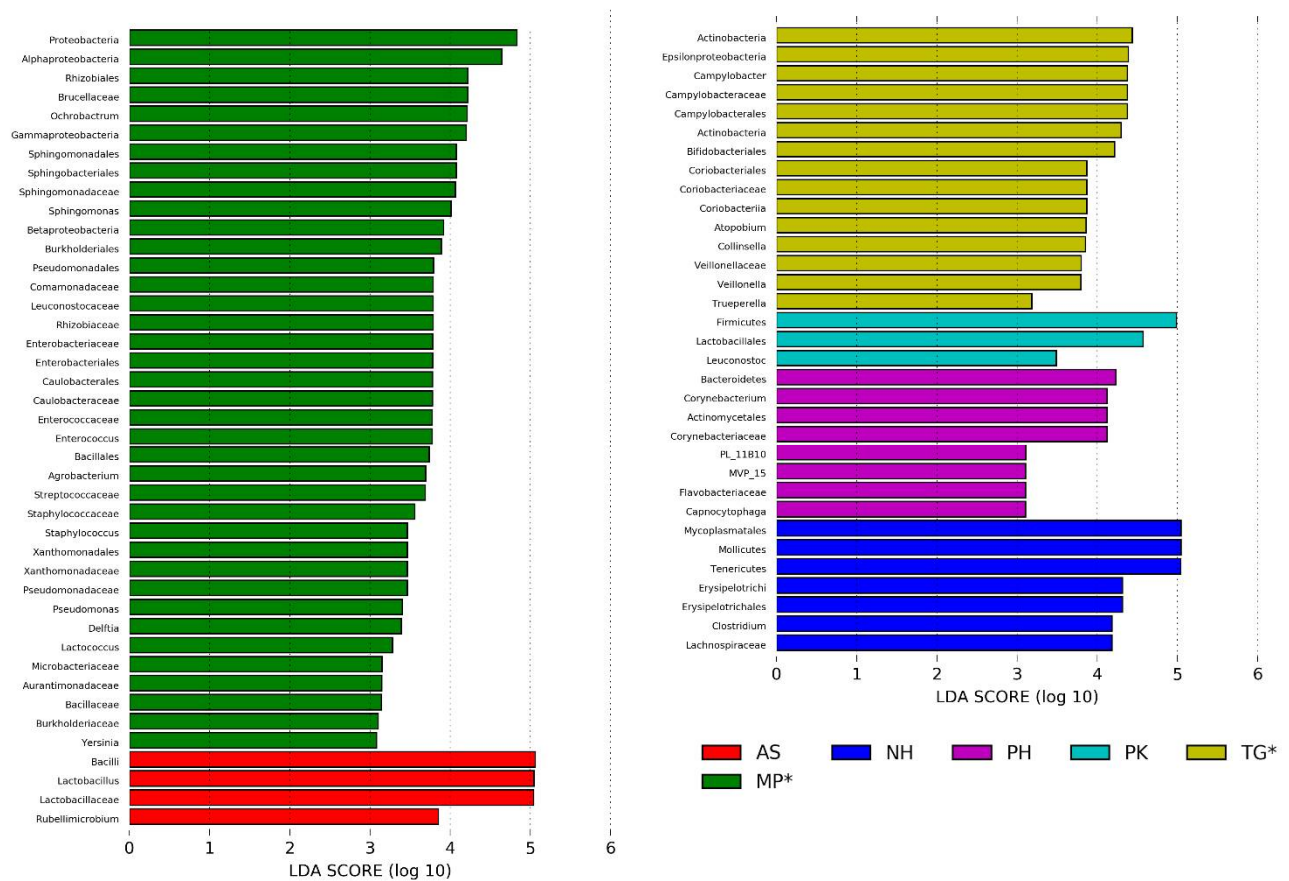
Figure S2. Plot showing LEfSe results. Bars at the right show bacterial groups that were significantly higher and bars at the left show bacterial groups that were significantly lower. As the reader can appreciate, in this study there were no bacterial groups that were lower in any given bird species. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
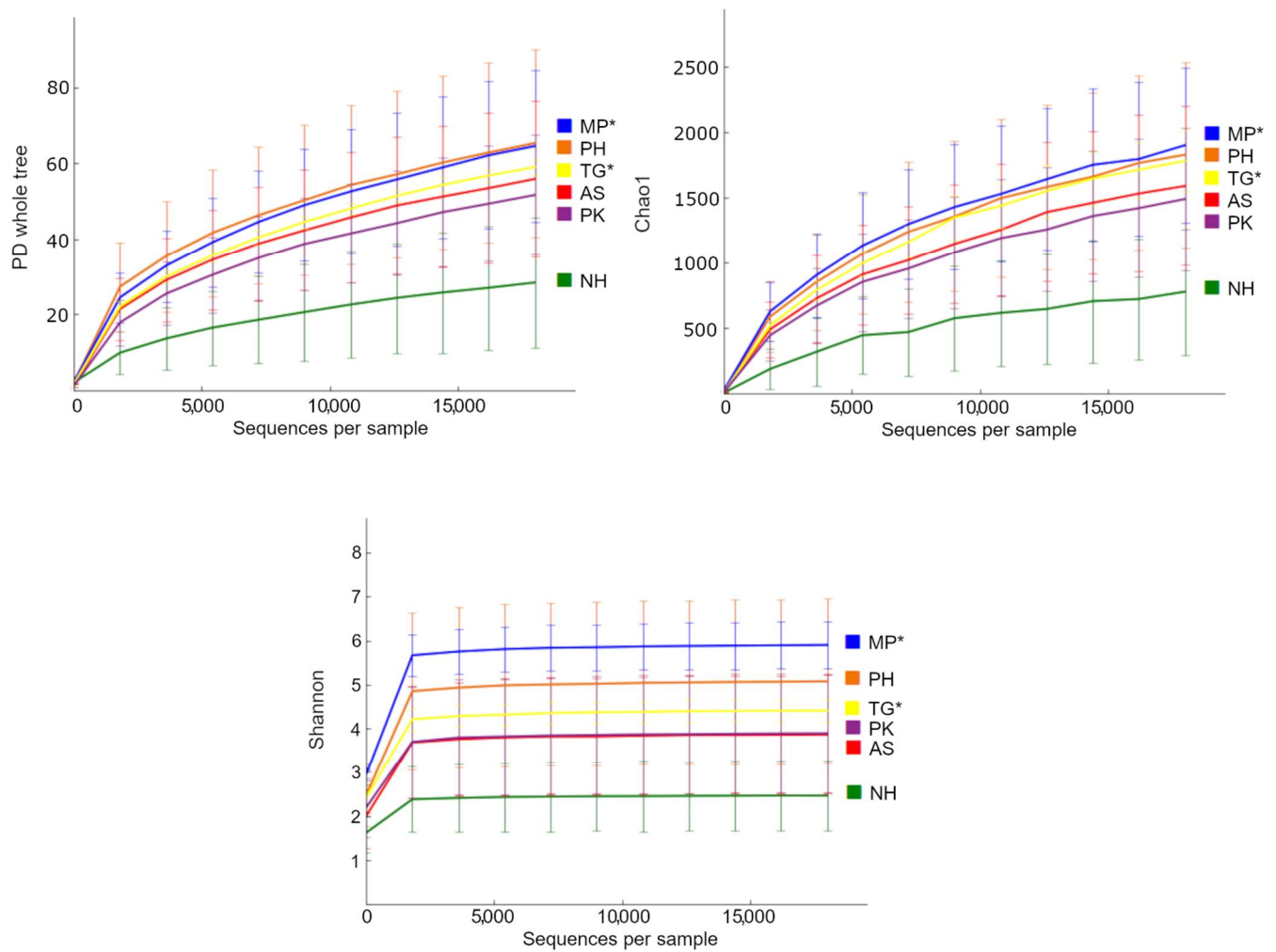
3. Alpha Diversity

Figure S3. Rarefaction plots of Phylogenetic Diversity (PD) whole tree, Chao1, and Shannon diversity indexes. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, TG: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameri*.

4. Bootstrapped Trees

Figure S4. Bootstrapped tree using both weighted UniFrac distances. A total of 18,000 sequences were used in each jackknifed subset. Red is for 75–100% bootstrapped support, yellow for 50–75%, green for 25–50%, and blue for <25% support. The bar represents dissimilarity. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, TG: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameri*.
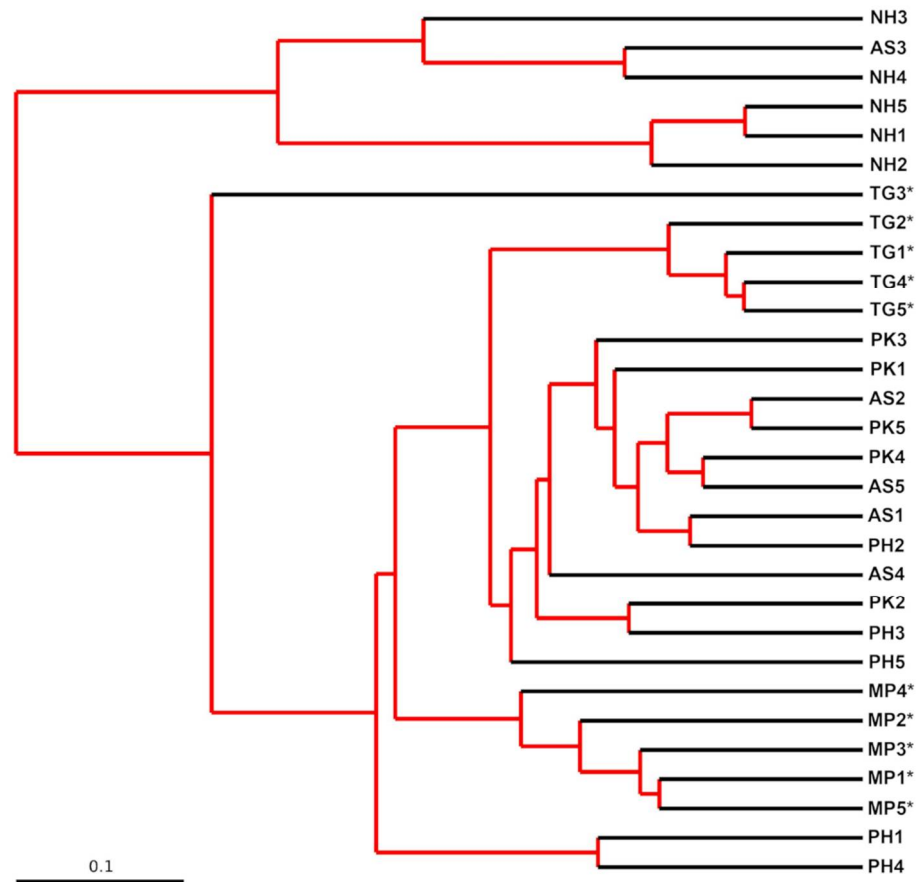
Figure S5. Bootstrapped tree using unweighted UniFrac distances. A total of 18,000 sequences were used in each jackknifed subset. Red is for 75–100% bootstrapped support, yellow for 50–75%, green for 25–50%, and blue for <25% support (note that the clustering was stronger using weighted distances). The bar represents dissimilarity. Samples from Passeriformes are highlighted MP: *Mimus polyglottos*, TG: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameri*.
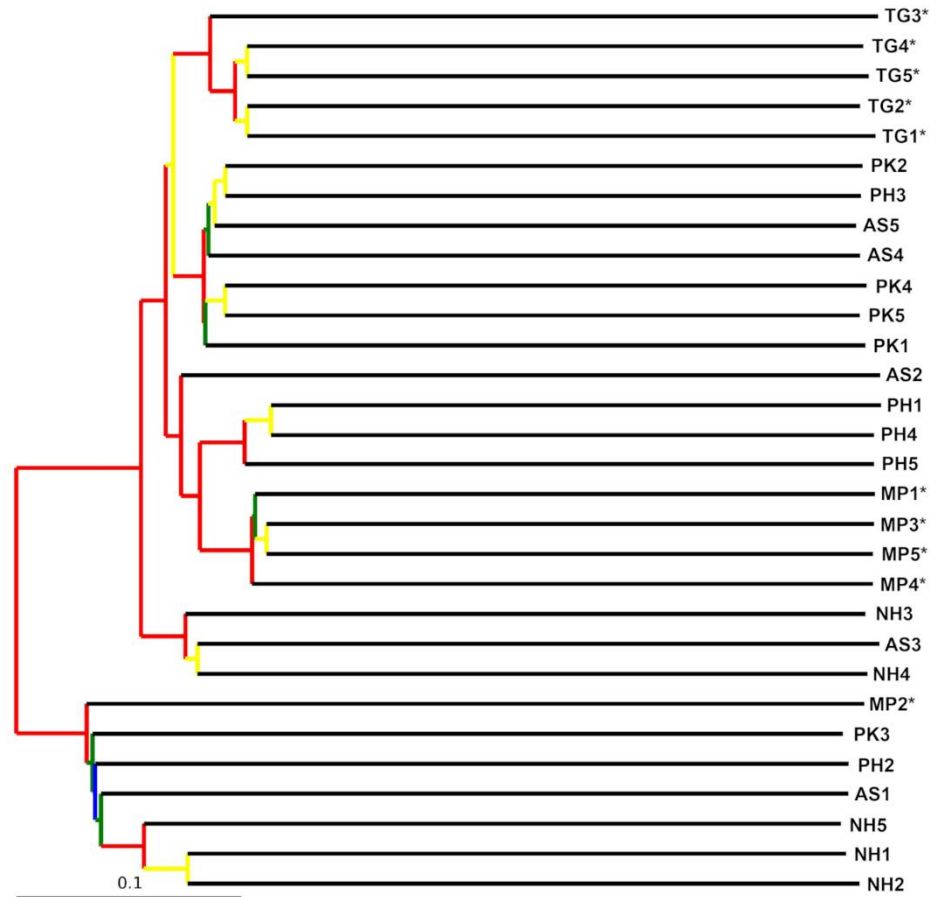
5. Differences in UniFrac Distances

Figure S6. Box plots of weighted (**a**) and unweighted (**b**) UniFrac distances between sample groupings. Each individual plot represents the difference between each bird species (labelled in *x* axis) and all the rest (for example, the boxes in the first plot in A represent the difference in weighted UniFrac distances between MP and TG, MP and AS, MP and PH, MP and NH, and MP and PK). Different letters indicate statistical significance difference (*p* < 0.05, Bonferroni corrected). MP: *Mimus polyglottos*, TG: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameri*.
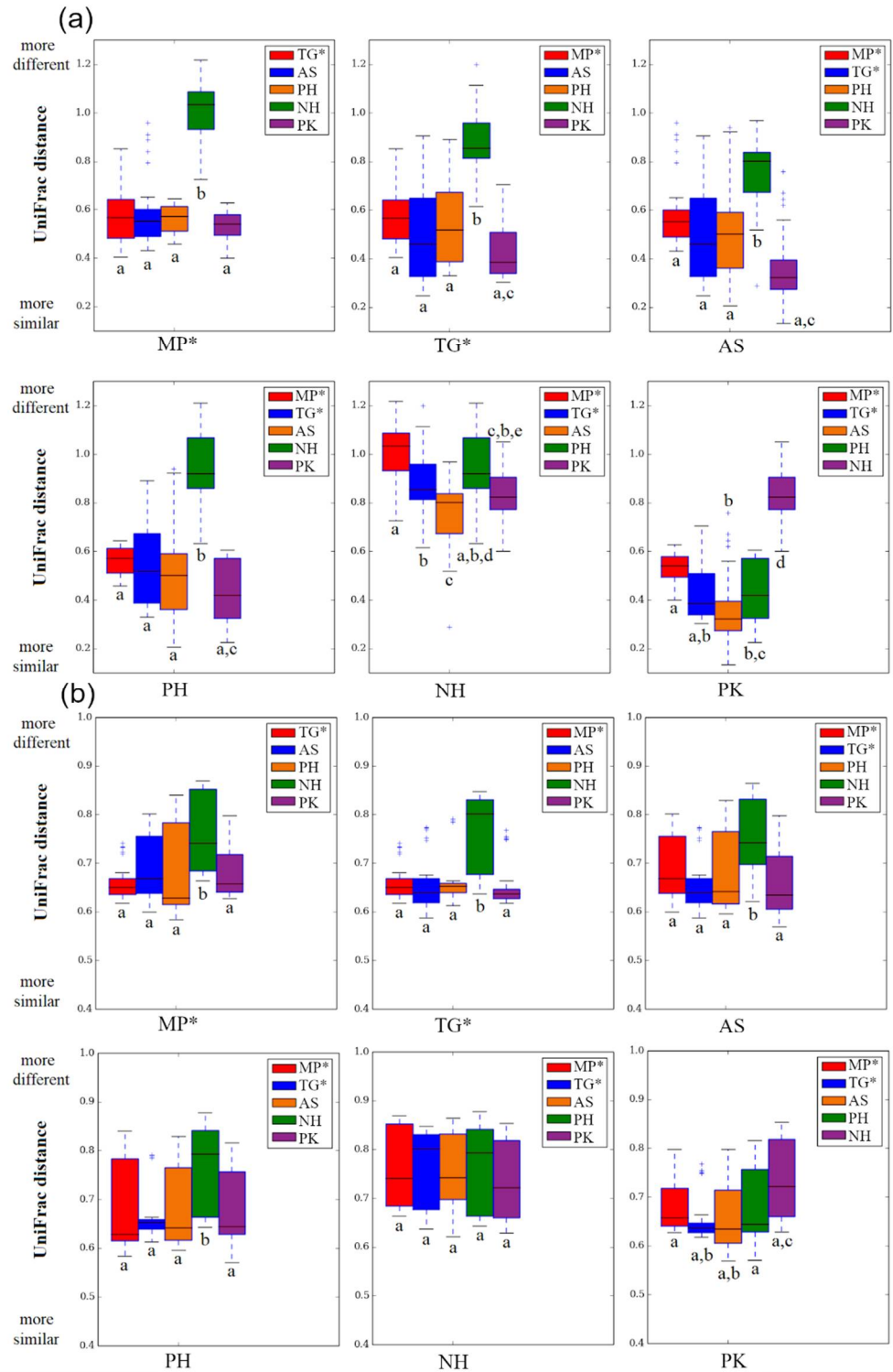
6. PICRUSt Analyses

Phylogenetic Investigation of Microbial Communities by Reconstruction of Unobserved States (PICRUSt) is a computational approach that was introduced by Langille et al. [12] in 2013 to predict the functional composition of a metagenome using marker gene data (e.g. 16S gene) and a database of reference genomes. PICRUSt was used by Waite and Taylor (2014) to predict the potential function of the gut microbiota from domestic (e.g. chickens) and wild (e.g. hoatzins) bird species. In this work, the authors mentioned that the predictions were consistent with the known state of avian microbiology but warned that caution must be taken in interpreting these predictions. For instance, in this study the feature "Carbohydrate metabolism", which Waite and Taylor (2014) emphasized in their discussion, ranged from 11.6% (MP samples) to 12.2% (PK samples), yet these values are similar to our results on cockatiels and budgerigars [13]. Table S1 shows a summary of the comparison of all features that showed statistical significance.

Table S1. Summary of those PICRUSt features with the lowest adjusted p values.

Feature	MP *	TG *	AS	PH	NH	PK	*P* Value
Non-homologous end-joining	↑				↓		1.62×10^{-9}
Caprolactam degradation	↑					↓	1.52×10^{-5}
Geraniol degradation	↑				↓		1.52×10^{-5}
Bladder cancer	↑				↓		1.64×10^{-5}
Lysine degradation	↑		↓				3.86×10^{-5}
Valine, leucine and isoleucine degradation	↑		↓				5.93×10^{-5}
Styrene degradation	↑				↓		7.96×10^{-5}
Tryptophan metabolism	↑					↓	8.90×10^{-5}
Amyotrophic lateral sclerosis	↑					↓	9.37×10^{-5}
Chagas disease	↑				↓		0.0001
Phenylalanine metabolism	↑		↓				0.0002
Prion diseases	↑				↓		0.0002
Type I diabetes mellitus	↓		↑				0.0002
African trypanosomiasis	↑				↓		0.0002
Renin-angiotensin system	↑				↓		0.0004
Hypertrophic cardiomyopathy	↑		↓				0.0006
Prenyltransferases	↓		↑				0.0006
Fatty acid metabolism	↑				↓		0.0008
Amino acid related enzymes	↓				↑		0.0009
Limonene and pinene degradation	↑					↓	0.002
Flagellar assembly	↑		↓				0.002
Glyoxylate and dicarboxylate metabolism	↑					↓	0.002
Drug metabolism - cytochrome P450	↑				↓		0.002
Biosynthesis of unsaturated fatty acids	↑				↓		0.004
Metabolism of xenobiotics by cytochrome P450	↑				↓		0.005
Homologous recombination	↓				↑		0.006
Glycosphingolipid biosynthesis - ganglio series	↑					↓	0.001

Feature	MP	TG	AS	PH	NH	PK	p-value
Beta-Alanine metabolism	↑				↓		0.001
One carbon pool by folate	↓				↑		0.01
Pertussis	↑				↓		0.01
Mismatch repair	↓				↑		0.02
Cellular antigens	↑				↓		0.02
Ribosome	↓				↑		0.02
P53 signaling pathway	↑					↓	0.02
Toxoplasmosis	↑					↓	0.02
Colorectal cancer	↑					↓	0.02
Small cell lung cancer	↑					↓	0.02
Viral myocarditis	↑					↓	0.02
Influenza A	↑					↓	0.02
Circadian rhythm -plant	↑				↓		0.02
Peptidoglycan biosynthesis	↓				↑		0.02
Protein export	↓				↑		0.02
DNA repair and recombination proteins	↓				↑		0.02
Chlorocyclohexane and chlorobenzene degradation	↑				↓		0.03
Aminobenzoate degradation	↑				↓		0.03
Terpenoid backbone biosynthesis	↓	↑					0.03
DNA replication proteins	↓				↑		0.03
Bacterial secretion system		↑				↓	0.03
Ascorbate and aldarate metabolism	↑	↓					0.03
Zeatin biosynthesis	↓				↑		0.04
Methane metabolism	↓				↑		0.04
Biosynthesis of siderophore group nonribosomal peptides	↑				↓		0.05

Samples that showed the highest (↑) and lowest (↓) proportion of genes related with each feature are marked for easier visualization. Features from samples of *M. polyglottos* (MP) showing the highest proportion of most features are shaded for better visualization. Passeriformes are highlighted (*). MP: *Mimus polyglottos*, TG: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameri*.

7. BugBase Results

BugBase is an algorithm that predicts organism-level coverage of functional pathways as well as biologically interpretable phenotypes such gram staining and pathogenic potential, within complex microbiomes. To our knowledge this study is the first using this tool in bird's microbiomes. In this study, we used the OTU table from the closed OTU picking approach for upload into BugBase (https://bugbase.cs.umn.edu/). Among other things, the results offered useful insights into the potential differences among bird species with regards to the OTU contributions to aerobic, anaerobic and facultatively anaerobic bacteria, bacteria with the potential to form biofilms, gram-positive and gram-negative bacteria, bacteria with mobile elements, potentially pathogenic and stress tolerant bacteria (Figures. S6–S14). However, these results are predictions only and therefore should be interpreted cautiously unless they are supported by other analyses such as in vitro phenotyping.
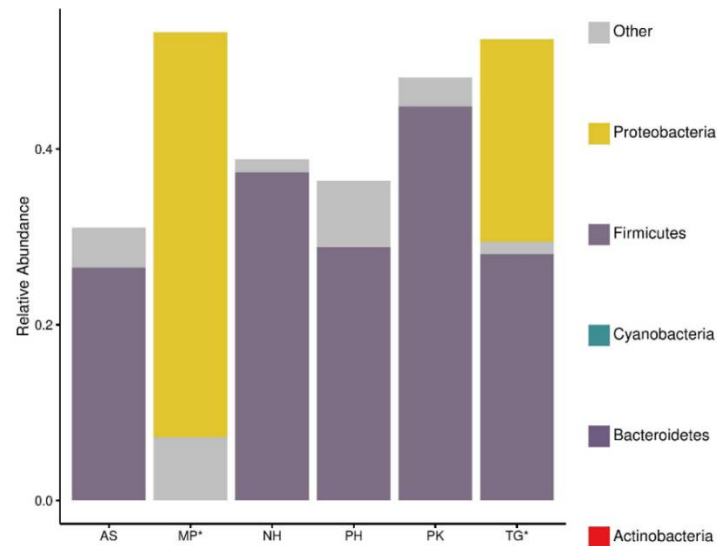
Figure S7. OTU contributions from the four more abundant phyla and others for aerobic bacteria accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
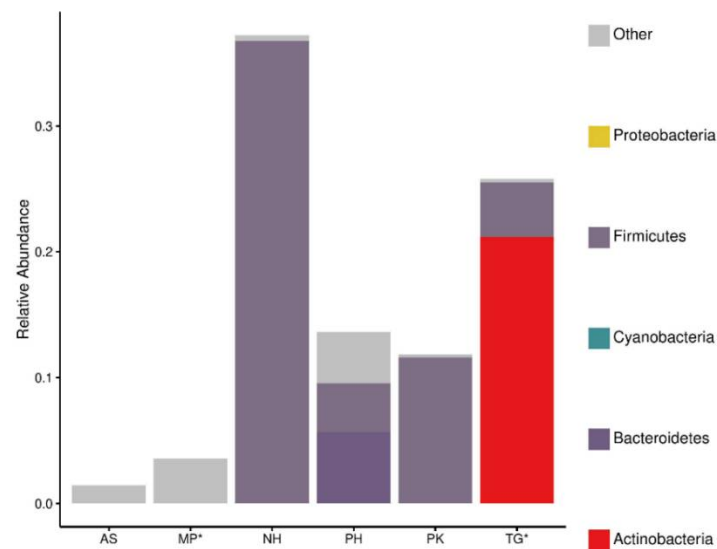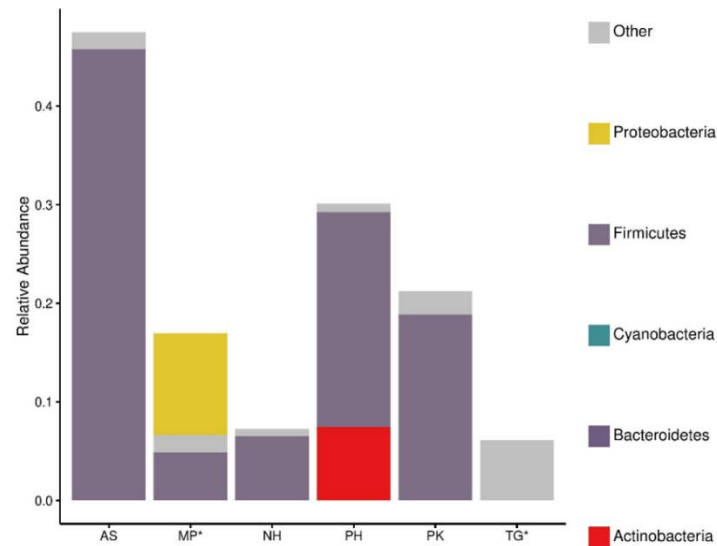

Figure S8. OTU contributions from the four more abundant phyla and others for anaerobic bacteria accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.

Figure S9. OTU contributions from the four more abundant phyla and others for facultative anaerobic bacteria accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
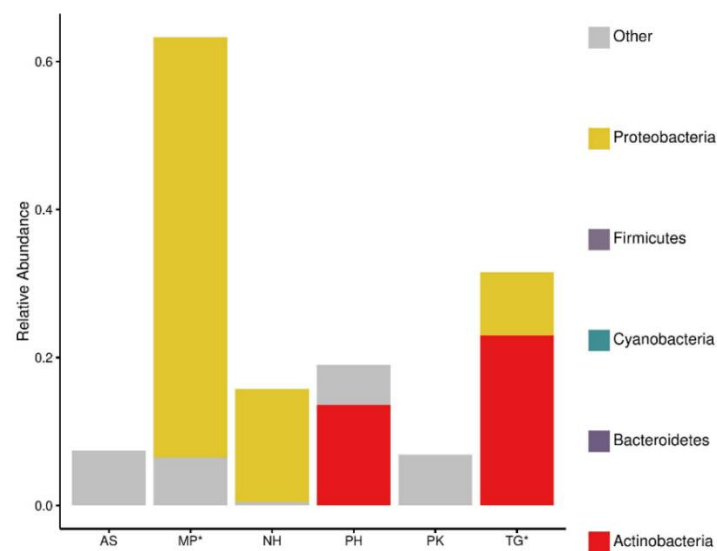

Figure S10. OTU contributions from the four more abundant phyla and others for bacteria with potential to form biofilms accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
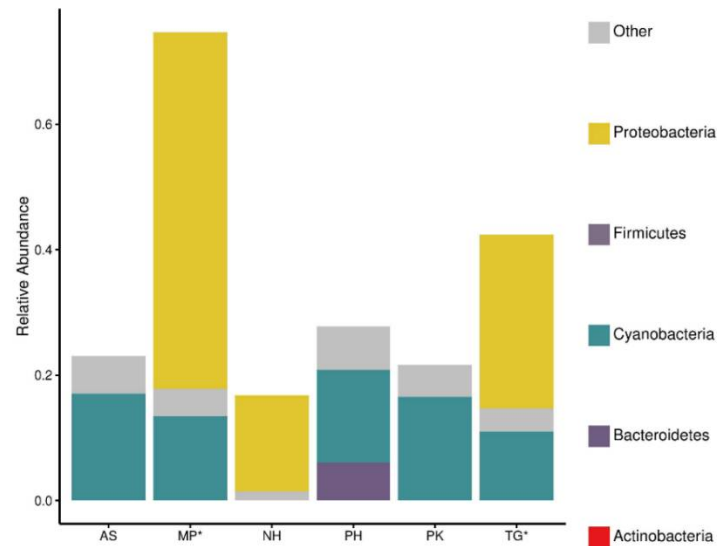
Figure S11. OTU contributions from the four more abundant phyla and others for gram-negative bacteria accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
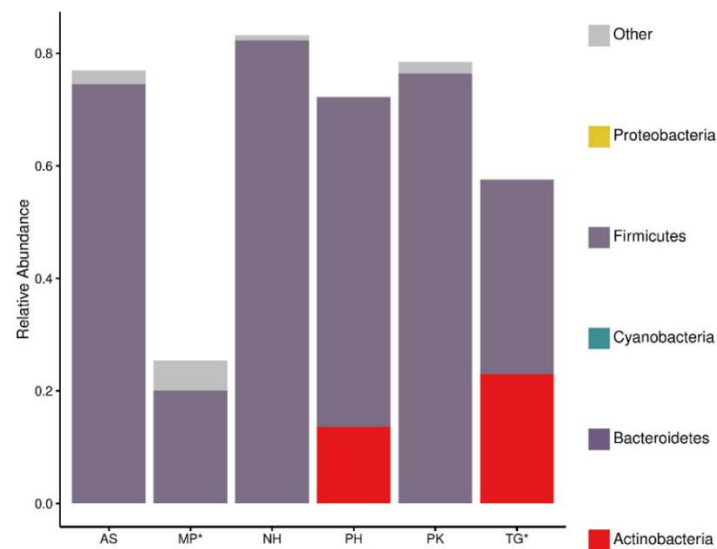

Figure S12. OTU contributions from the four more abundant phyla and others for gram-positive bacteria accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
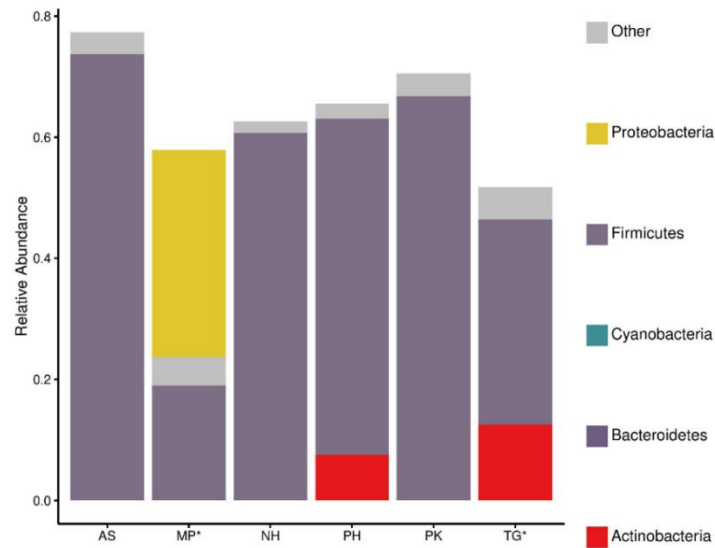
Figure 13. OTU contributions from the four more abundant phyla and others for bacteria with mobile elements accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.

Figure S14. OTU contributions from the four more abundant phyla and others for potentially pathogenic bacteria accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
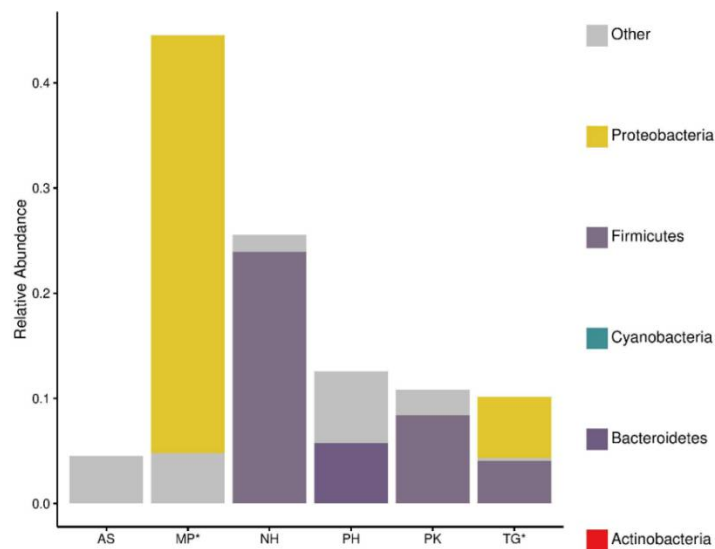
Figure S15. OTU contributions from the four more abundant phyla and others for stress tolerant bacteria accordingly to BugBase analyses. Samples from Passeriformes are highlighted (*). MP: *Mimus polyglottos*, Tg: *Taeniopygia guttata*, AS: *Agaporni* spp., PH: *Psephotus haematonotus*, NH: *Nymphicus hollandicus*, PK: *Psittacula krameria*.
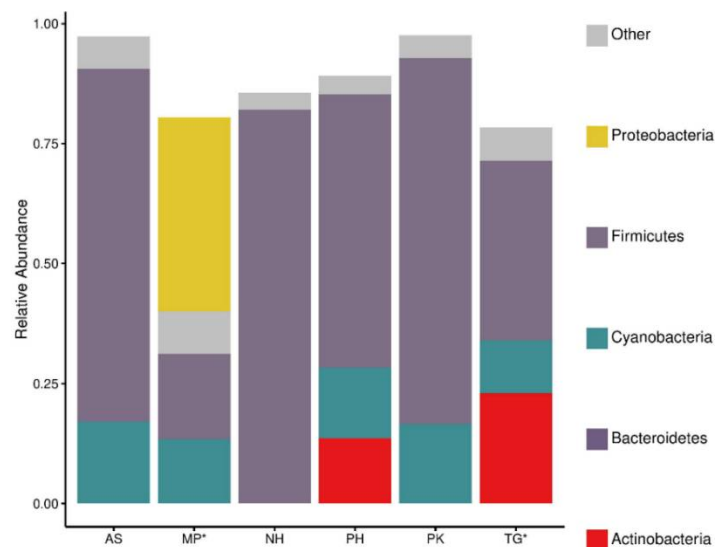
References

1. Alvarenga, D.O.; Fiore, M.F.; Varani, A.M. A metagenomic approach to cyanobacterial genomics. *Front. Microbiol.* **2017**, *8*, 809.
2. Archibald, J.M. Genomic perspectives on the birth and spread of plastids. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10147–10153.
3. Markov, A.V.; Zakharov, I.A. Evolution of gene orders in genomes of cyanobacteria. *Genetika* **2009**, *45*, 1036–1047.
4. Ochoa de Alda, J.A.G.; Esteban, R.; Diago, M.L.; Houmard, J. The plastid ancestor originated among one of the major cyanobacterial lineages. *Nature Comm.* **2014**, *5*, 4937.
5. Di Rienzi, S.C.; Sharon, I.; Wrighton, K.C.; Koren, O.; Hug, L.A.; Thomas, B.C.; Goodrich, J.K.; Bell, J.T.; Spector, T.D.; Banfield, J.F.; Ley, R.E. The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *eLife* **2013**, *2*, e01102.
6. Utami, Y.D.; Kuwahara, H.; Murakami, T.; Morikawa, T.; Sugaya, K.; Kihara, K.; Yuki, M.; Lo, N.; Deevong, P.; Hasin, S.; Boonriam, W.; Inoue, T.; Yamada, A.; Ohkuma, M.; Hongoh, Y. Phylogenetic diversity and single-cell genome analysis of "Melainabacteria", a non-photosynthetic Cyanobacterial group, in the termite gut. *Microbes Environ.* **2018**, *33*, 50–57
7. Hiras, J.; Wu, Y-W.; Eichorst, S.A.; Simmons, B.A.; Singer, S.W. Refining the phylum Chlorobi by resolving the phylogeny and metabolic potential of the representative of a deeply branching, uncultivated lineage. *The ISME J.* **2015**, *10*, 833–845.
8. Campbell, A.G.; Schwientek, P.; Vishnivetskaya, T.; Woyke, T.; Levy, S.; Beall, C.J.; Griffen, A.; Leys, E.; Podar, M. Diversity and genomic insights into the uncultured Chloroflexi from the human microbiota. *Environ. Microbiol.* **2014**, *16*, 2635–2643.
9. Ley, R.E.; Bäckhed, F.; Turnbaugh, P.; Lozupone, C.A.; Knight, R.D.; Gordon, J.I. Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11070–11075.
10. García-Amado, M.A.; Shin, H.; Sanz, V.; Lentino, M.; Martínez, L.M.; Contreras, M.; Michelangeli, F.; Domínguez-Bello, M.G. Comparison of gizzard and intestinal microbiota of wild neotropical birds. *PLoS ONE* **2018**, *13*, e0194857.
11. Hird, S.M.; Sanchez, C.; Carstens, B.C.; Brumfield, R.T. Comparative gut microbiota of 59 neotropical bird species. *Front Microbiol.* **2015**, *6*, 1403.
12. Langille, M.G.I.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.A.; Clemente, J.C.; Burkepile, D.E.; Vega Thurber, R.L.; Knight, R.; Beiko, R.G.; Huttenhower, C. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **2013**, *31*, 814–821.
13. Garcia-Mazcorro, J.F.; Castillo-Carranza, S.A.; Guard, B.; Gomez-Vazquez, J.P.; Dowd, S.E.; Brightsmith, D.J. Comprehensive molecular characterization of bacterial communities in feces of pet birds using 16S marker sequencing. *Microb. Ecol.* **2017**, *73*, 224–235.