Authors

Jonathan B Clayton, Gabriel A Al-Ghalith, Ha Thang Long, Bui Van Tuan, Francis Cabana, Hu Huang, Pajau Vangay, Tonya Ward, Vo Van Minh, Nguyen Ai Tam, Nguyen Tat Dat, Dominic A Travis, Michael P Murtaugh, Herbert Covert, Kenneth E Glander, Tilo Nadler, Barbara Toddes, John C M Sha, Randy Singer, Dan Knights, and Timothy J Johnson

SCIENTIFIC REPORTS

Associations Between Nutrition, Gut Microbiome, and Health in A Novel Nonhuman Primate Model

Jonathan B. Clayton [1,2,3], Gabriel A. Al-Ghalith[4], Ha Thang Long[2,3,5], Bui Van Tuan[2], Francis Cabana [3,6], Hu Huang[4], Pajau Vangay[4,7], Tonya Ward[7], Vo Van Minh[8], Nguyen Ai Tam[2], Nguyen Tat Dat[2], Dominic A. Travis[3,9], Michael P. Murtaugh[1], Herbert Covert[10], Kenneth E. Glander [3,11], Tilo Nadler[12], Barbara Toddes[13], John C. M. Sha[14], Randy Singer[1], Dan Knights[3,4,7,15] & Timothy J. Johnson[1,3,16]

Red-shanked doucs (*Pygathrix nemaeus*) are endangered, foregut-fermenting colobine primates which are difficult to maintain in captivity. There are critical gaps in our understanding of their natural lifestyle, including dietary habits such as consumption of leaves, unripe fruit, flowers, seeds, and other plant parts. There is also a lack of understanding of enteric adaptations, including their unique microflora. To address these knowledge gaps, we used the douc as a model to study relationships between gastrointestinal microbial community structure and lifestyle. We analyzed published fecal samples as well as detailed dietary history from doucs with four distinct lifestyles (wild, semi-wild, semi-captive, and captive) and determined gastrointestinal bacterial microbiome composition using 16S rRNA sequencing. A clear gradient of microbiome composition was revealed along an axis of natural lifestyle disruption, including significant associations with diet, biodiversity, and microbial function. We also identified potential microbial biomarkers of douc dysbiosis, including *Bacteroides* and *Prevotella*, which may be related to health. Our results suggest a gradient-like shift in captivity causes an attendant shift to severe gut dysbiosis, thereby resulting in gastrointestinal issues.

The primate gastrointestinal (GI) tract is home to trillions of bacteria that play major roles in digestion and metabolism, immune system development, pathogen resistance, and other important aspects of host health and behavior[1–4]. While there has been substantial progress in understanding the role microbial communities play in human health and disease[1–3], as well as numerous microbiome studies in non-primate animal models such as mice[5–9], somewhat less attention has been given to host-associated microbiomes and lifestyle, including diet, in nonhuman primates (NHPs). Developing a better understanding of the link between primate microbial

[1]Department of Veterinary and Biomedical Sciences, University of Minnesota, 1971 Commonwealth Avenue, Saint Paul, MN, 55108, USA. [2]GreenViet Biodiversity Conservation Center, K39/21 Thanh Vinh Street, Son Tra District, Da Nang, Vietnam. [3]Primate Microbiome Project, 6-124 MCB, 420 Washington Ave SE, Minneapolis, MN, 55455, USA. [4]Biomedical Informatics and Computational Biology, 200 Union St SE, University of Minnesota, Minneapolis, MN, 55455, USA. [5]Frankfurt Zoological Society, Bernhard-Grzimek-Allee 1, 60316, Frankfurt, Germany. [6]Wildlife Nutrition Centre, Wildlife Reserves Singapore, 80 Mandai Lake Road, 729826, Singapore, Singapore. [7]Biotechnology Institute, University of Minnesota, 1479 Gortner Avenue, Saint Paul, MN, 55108, USA. [8]Faculty of Biology and Environmental Science, The University of Da Nang - University of Science and Education, 459 Ton Duc Thang Street, Lien Chieu District, Da Nang, Vietnam. [9]Department of Veterinary Population Medicine, University of Minnesota, 1365 Gortner Avenue, 225 Veterinary Medical Center, Saint Paul, MN, 55108, USA. [10]Department of Anthropology, University of Colorado Boulder, 1350 Pleasant St, 233 UCB, Boulder, CO, 80309, USA. [11]Department of Evolutionary Anthropology, Duke University, 130 Science Drive, 104 Biological Sciences, Durham, NC, 27708, USA. [12]Endangered Primate Rescue Center, Cuc Phuong National Park, Nho Quan District, Ninh Binh Province, Vietnam. [13]Philadelphia Zoological Garden, 3400 West Girard Avenue, Philadelphia, PA, 19108, USA. [14]School of Sociology and Anthropology, Sun Yat-Sen University, Guangzhou, 510275, China. [15]Department of Computer Science and Engineering, University of Minnesota, 4-192 Keller Hall, 200 Union St SE, Minneapolis, MN, 55455, USA. [16]University of Minnesota, Mid-Central Research and Outreach Center, Willmar, Minnesota, USA. Correspondence and requests for materials should be addressed to T.J.J. (email: joh04207@umn.edu)

communities and lifestyle, including diet and health, is important not only in the context of primate ecology, but may also have profound implications for use of NHPs as model systems for lifestyle disruption and associated microbial changes.

One colobine primate species, the red-shanked douc (i.e., douc), may be of particular interest as a model organism. Because it performs both foregut fermentation and hindgut digestion[10,11], the douc shares digestive characteristics with both humans and ruminant livestock. From a conservation standpoint, it is endangered and fails to thrive in captivity[12–14]. From a health standpoint, this failure to thrive stems foremost from severe gastro-intestinal disease, which has been shown in other model organisms and humans to be directly associated with the gut microbiome[15–21]. Further, in contrast to human studies, the genetic background and environmental conditions of NHPs can be easily and directly ascertained or manipulated[22], which is critical as both environmental and genetic factors have been implicated in modulation of the microbiome[23–25]. Hence, the douc may serve as a model organism relevant to the domains of conservational and microbial ecology, and potentially inform both human and livestock health.

The douc is of particular conservational importance among primate species. The douc is listed as endangered by the IUCN[26], and recovery efforts to restore the douc population currently utilize conservation sanctuaries, anti-poaching laws, and captive breeding programs. While their digestive specializations have allowed them to thrive in their native habitat, the same specializations appear to challenge their survival in captivity. In fact, these primates are among the most difficult to keep in captivity and are rarely kept in zoological institutions[12–14]. Maintenance of doucs and other colobines as captive populations has been largely unsuccessful due in large part to an inadequate understanding of their nutritional requirements. They are highly susceptible to gastrointestinal (GI) disorders when maintained on commercially prepared diets in captivity. Improving their diet in captivity is challenging due to critical gaps in our understanding of local fibrous vegetation and the enteric microbial adaptations that facilitate efficient extraction of key nutrients. A deeper understanding the microbial properties and functions that underlie captivity-associated dysbiosis in these critically endangered non-human primates may positively impact our ability to intervene on their behalf. To what extent and in what capacity these findings generalize to human populations (who share a close genetic background) or livestock (who share many digestive characteristics) remains an important topic for future investigation. We hypothesized that specific and unique microbial subsets play a critical role in the utilization of fibrous vegetation with natural toxicants, and that captive doucs lack the microbiota to maintain optimal health due to inadequate dietary substrate. In order to better understand the link between lifestyle, gut microbial communities, and health, we examined the fecal microbiomes of four douc populations living four distinct lifestyles (wild, semi-wild, semi-captive, and captive).

Here, by comparing four different populations of the same species along a captivity/wildness (i.e., lifestyle) gradient, we sought to determine whether such a gradient in lifestyle also manifests in a gradient of gut microbiota and diet, and to assess whether any such trends corroborate health status. Furthermore, examining one species living under four different lifestyles allows one to examine the influence of environment independent of interspecific host variation on shaping gut microbial community structure. As microbes may act as indicators for health of the host[18], these results may allow for the development of predictive biomarkers to improve NHP health and management. As some microbial trends hold across species boundaries in other model systems[27–30], some biomarkers may translate to human and ruminant health as well, especially since doucs are closer genetic relatives to humans, and share digestive characteristics with ruminants.

Here, we focus on a subset of a rich dataset we collected over a two-year period across three countries, including a comprehensive sampling of the majority of doucs in captivity. These samples have been used in work published previously[31] in a broader meta-analysis framework, where a putative convergence was observed across various primate species (including humans) with increasing levels of generalized lifestyle disruption. While the broadness of such an overview was valuable in demonstrating overall trends across species, it may also mask intraspecific effects underlying a lifestyle gradient, and also limits the resolution and interpretability of correlations that can be drawn to specific lifestyle components, including diet and health, which themselves may play very different biological roles in the various species under investigation. By focusing in depth on the dietary and microbial facets associated with a single species across increasingly unnatural lifestyle conditions, we are powered to make specific conclusions relating these covariates in a common context.

Methods

All work in this study was carried out following the International Primatological Society Ethical Guidelines for the Use of Non-Human Primates in Research, and the American Society for Primatologist's Principles for the Ethical Treatment of Non-Human Primates. All work in this study was reviewed and approved by the University of Minnesota Institutional Animal Care and Use Committee, and it was determined that no formal approval was needed because the work did not involve the handling of live animals. All laboratory protocols were reviewed and approved by the University of Minnesota Institutional Biosafety Committee under protocol number 1303–30480 H. Approvals were obtained from appropriate governmental and organizational authorities for plant and animal feces collection at all sites. These included approvals from Philadelphia Zoo, Singapore Zoo, Endangered Primate Rescue Center, Department of Forest Protection (Da Nang City, Vietnam), Da Nang University (Da Nang City, Vietnam), and Son Tra Nature Reserve Ranger Force (Da Nang City, Vietnam).

Study site, subjects, and sample material. Fecal samples (n = 111 samples, at least 35 subjects) were collected opportunistically immediately after defecation from captive (n = 12 samples, 2 subjects), semi-captive (n = 15 samples, 7 subjects), semi-wild (n = 18 samples, 18 subjects), and wild (n = 66 samples, 8 or more subjects) doucs (*Pygathrix nemaeus*) in 2012–2013. For the wild individuals, fecal samples (n = 26 samples) were collected from eight uniquely identifiable individuals. The remaining fecal samples (n = 39) collected from the wild population originated from individuals that could not be uniquely identified, and as such these samples could

have been from any of the 8 identified wild individuals or others never identified. Samples were freshly frozen after collection at −20 °C, and remained frozen during air transport to the USA. Once in the USA, samples were frozen at −80 °C until processing. The microbiota of the fecal samples were analyzed previously in a meta-analysis examining broader relationships between captivity and the microbiome[31]. In the wild, it can be difficult to distinguish which fecal samples came from which subjects, thus we have clarified differences between sample number and subject number above. For this analysis, doucs housed at the Endangered Primate Rescue Center (EPRC) in Cuc Phuong National Park, Ninh Binh, Vietnam served as the semi-wild population, as the doucs there live a lifestyle (including environment and diet) representing an intermediate state between wild and semi-captive. Specifically, these doucs are fed a wide variety of plants foraged daily by EPRC staff and allowed to choose which to eat. They are also fed 3x per day and allowed to eat throughout the day, which is more representative of the natural condition. Lastly, for two hours each day (approximately 12–2 pm), all staff and visitors leave the compound to allow all the primates to sleep, which is similar to how the doucs behave in the wild (Clayton, personal communication). They are not offered any supplemental dietary items, such as ripe fruits, vegetables, or vitamin supplements, all of which are fed to doucs housed at traditional zoological institutions. Additionally, while the doucs are housed in enclosures, they are kept outside year-round. However, Cuc Phuong National Park is outside of the wild douc home range. Doucs housed at the Singapore Zoo served as the semi-captive population, as they live a lifestyle (including environment and diet) representing an intermediate state between semi-wild and captive. Specifically, they are fed a diet richer in plant species compared to the captive population (15 species vs. 1 species), however, as in the captive population, their diet includes ripe fruits, vegetables, and vitamin supplements. The plants offered to the doucs are picked by staff daily and delivered to them. Additionally, the climate of Singapore is closer to that of their native habitat in Vietnam, as both countries are located in Southeast Asia. Doucs housed at the Philadelphia Zoo served as the captive population, as they live in artificial environments compared to their semi-wild and wild counterparts. The doucs are fed only a single plant species, and consume a traditional zoological diet. They also remain indoors year-round. Doucs inhabiting Son Tra Nature Reserve, Da Nang, Vietnam (16°06′–16°09′N, 108°13′–108°21′E) served as the wild population in this comparative study[32,33] (Supplemental Fig. 1). Son Tra is located only 10 km from the heart of Da Nang City, which is the third largest city in Vietnam. The nature reserve is comprised of 4,439 total ha and, of those, 4,190 ha is covered by both primary and secondary forests[32]. Da Nang has two seasons every year, wet (September until March) and dry (April until August). We collected samples from March 2013–June 2013. The samples collected in March were collected at the end of the month. Thus, the wild douc samples were collected in the dry season. The wild doucs are the only population who forage for their preferred native plant species on their own.

Defining the "lifestyle" concept. Central to our re-analysis of these douc samples is the concept of lifestyle. Here we define lifestyle as a synthetic variable, an amalgamation of environmental factors that comprise the living conditions of a population. As such, this umbrella variable is a proxy for climate, diet, and geographically-localized microbial exposure, among potentially numerous other variables. This is broadly similar to the meaning of lifestyle as applied to human populations, where the descriptive comparison of different global cultures is similarly subject to innumerable covariates, known and unknown, under a similar umbrella[34,35]. We independently assess dietary information, which has been shown to be a critical component of lifestyle in context of the microbiome, but it is not our intention to attribute changes in microbiota, function, or health indicators to diet alone. Hence, our focus is foremost on performing correlations and gradient analysis with the backbone gradient comprised of the latent ordinal correlate "lifestyle." We accordingly pepper this central analysis with auxiliary analyses and conjecture where appropriate to explain the trends observed.

Genomic DNA extraction. Total DNA from each fecal sample was extracted as described with some modifications[36]. Briefly, two rounds of bead-beating were carried out in the presence of NaCl and sodium dodecyl sulfate, followed by sequential ammonium acetate and isopropanol precipitations; precipitated nucleic acids were treated with DNase-free RNase (Roche); and DNA was purified with the QIAmp® DNA Stool Mini Kit (QIAGEN, Valencia, CA), according to manufacturer's recommendations. DNA quantity was assessed using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific Inc, Massachusetts, USA).

Bacterial 16S rRNA PCR amplification and Illumina MiSeq sequencing. The bacterial 16S rRNA gene was analyzed using primers 515F and 806R, which flanked the V4 hypervariable region of bacterial 16S rRNAs[37]. The oligonucleotide primers included Illumina sequencing adapters at the 5′ ends and template specific sequences at the 3′ ends. The primer sequences were: 515F (forward) 5′ GTGCCAGCMGCCGCGGTAA 3′ and 806R (reverse) 5′ GGACTACHVGGGTWTCTAAT 3′[37]. The 16S rRNA PCR amplification protocol from the earth microbiome project was used[38]. Each sample was amplified in two replicate 25-μL PCR reactions and pooled into a single volume of 50 μL for each sample. The amplification mix contained 13 μL of PCR grade water (MoBio, Carlsbad, CA), 10 μL of 5 PRIME HotMasterMix (5 PRIME, Gaithersburg, MD), 0.5 μL of each fusion primer, and 1.0 μL of template DNA in a reaction volume of 25 μL. PCR conditions were an initial denaturation at 94 °C for 3 m; 35 cycles of 94 °C 45 s, 50 °C for 60 s, and 72 °C for 90 s; and a final 10 m extension at 72 °C. Following PCR, concentration of PCR products was determined by a PicoGreen assay. Equal amounts of samples were pooled, and size selection was performed using the Caliper XT (cut at 386 bp +/− 15%). Final quantification was performed via a PicoGreen assay and assessment on a Bioanalyzer 2100 (Agilent, Palo Alto, California) using an Agilent High Sensitivity chip. The PCR amplicons were sequenced at the University of Minnesota Genomics Center (UMGC) using Illumina MiSeq and 2 × 300 base paired-end reads (Illumina, San Diego, California).

16S Data analysis. Raw sequences were analyzed with QIIME 1.8.0 pipeline[39]. The demultiplexed sequences from the UMGC were subjected to the following quality filter: 150 bp < length < 1,000 bp; average quality score

>25. Preprocessed sequences were then clustered at 97% nucleotide sequence similarity level. For the diversity and taxonomic analyses, the open-reference-based OTU picking protocol in QIIME was used with GreenGenes 13_8 as the reference database[40] using the USEARCH algorithm[41]. Unmatched reads against the reference database which also did not cluster later in the open reference pipeline were excluded from the downstream analysis. Read depth was relatively uniform across lifestyles (Supplemental Fig. 2). Taxonomy information was then assigned to each sequence cluster using RDP classifier 2.2[42]. Closed-reference OTUs of chloroplast origin were filtered out with QIIME, and samples were rarefied to 52,918 reads for the downstream analysis.

For the closed-reference-only analyses, including PICRUSt and chloroplast analyses, the raw FASTQ files were processed with SHI7[43], a wrapper script that detected and removed TruSeq v3 adaptors with trimmomatic[44], stitched the R1 and R2 reads together with FLASh[45], performed quality trimming from both ends of the stitched reads until a minimum quality score ≥32 was reached, and filtered out reads with average quality score <36. 88.5% of all original sequences were retained after QC, resulting in an average read length of 254 bases and average quality score of 37.6. Closed-reference picking was performed at 95% similarity level with the taxonomy-aware exhaustive optimal alignment program BURST[46] against a database of all RefSeq chloroplast sequences in phyla Chlorophyta (green algae) and Streptophyta (land plants) as of 06/27/2017, a total of 1,506 chloroplast reference sequences. The same closed-reference procedure was also used to re-pick OTUs against GreenGenes 13_8 for use with PICRUSt, as the latter is reliant on closed-reference GreenGenes IDs for functional prediction.

Alpha diversity (including chao1, shannon, and simpson diversity metrics) and beta diversity analysis (including Bray-Curtis, weighted and unweighted UniFrac metrics)[47], were performed and plotted through a combination of wrapper scripts in QIIME and custom R scripts using the vegan, ape, ggplot2, and phyloseq packages[48–51]. ANOVA was used to assess the statistical significance of alpha diversity variation among populations. We used R's default t-test, the Welch's t-test, for pairwise comparisons of alpha diversity, as it is robust to unequal variances and unequal sample sizes between populations. Adonis was used to assess whether populations significantly differed by beta diversity[49].

The functional profiles of the microbial sample were investigated using PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States)[52], which predicts Kyoto Encyclopedia of Genes and Genomes (KEGG) module abundances within a microbial community based on 16S rRNA surveys. Within this pipeline, relative abundances of OTUs were normalized by 16S rRNA copy number, after which centered log ratio transformation was applied with detection-limit zero replacement. Metagenomic contents were predicted from the KEGG catalogue[53]. The mean Nearest Sequenced Taxon Index (NSTI) for all lifestyles was below 0.18 (Supplemental Fig. 3). To assess degree of correlation between each functional module and population group ordered by degree of wildness, polyserial correlation was used with ordered lifestyle groups (captive < semi-captive < semi-wild < wild) via the polycor package[54]. Statistical significance of this correlation was ascertained by computing a p-value from the polyserial chi-square rho and degrees of freedom, followed by Holm family-wise error rate correction (alpha <0.05). Additionally, the opposite ends of the lifestyle spectrum were compared using pairwise statistical tests in order to filter and corroborate the polyserial correlations. To retain comparability with the groupings used previously[31] as well as accumulate sufficient samples to power pairwise comparison, the "captive" end of the spectrum was represented by combining both the captive and semi-captive groups into a single "(Semi-)captive" group. These two groups were compared pairwise using the non-parametric Wilcoxon Rank-Sum test followed by Holm adjustment. All associations for which absolute rho < 0.3, adj. p > 0.05, or rho confidence p > 0.05 were considered insignificant.

Differential taxon abundance testing was also performed. OTUs were binned additively according to taxonomy at the genus level (or at lowest characterized level if genus was uncharacterized for a given OTU). The resulting taxa were then filtered such that only taxa present in at least 3 samples and with 0.01% average abundance throughout the dataset were retained, leaving 75 distinct taxa at or below genus level. The OTU table was normalized by centered log-ratio with least-squares zero interpolation[55] to allow for valid compositional covariate testing[56]. Similarly to the PICRUSt KEGG module differential abundance testing described above, statistical significance of association between each taxon and the sample populations was assessed with polyserial correlation across groups in order of wildness, as well as Wilcoxon rank-sum tests of the two extrema (combined captive & semi-captive group vs wild group), followed by Holm adjustment. All associations for which absolute rho < 0.3, adj. p > 0.05, or rho confidence p > 0.05 were considered insignificant. For heatmaps, only features with absolute rho > 0.6, adj. p < 0.05, and rho confidence p < 0.05 are displayed for clarity.

To ensure robustness of statistical analyses in light of the varying number of fecal samples per individual in the lifestyle groups, we collapsed all replicates that could have possibly originated from a single douc into one averaged sample for that douc. We also collapsed all stool samples originating from potentially ambiguous wild doucs into a single "unknown" wild sample. This reduced the number of samples considered to 36 (a single sample per uniquely identifiable individual, plus one additional sample representing the average of all unknown wild individuals in the wild lifestyle). Specifically, after collapse, there were 2 captive samples, 7 semi-captive samples, 18 semi-wild samples, and 9 wild samples. Data from this secondary validation round are presented in supplementary material for each relevant result. P-values reported by these tests were FDR-corrected using the Benjamini-Hochberg method.

Data deposition. All sequencing data are deposited at the European Bioinformatics Institute under project number PRJEB11414. Additionally, all R code and raw non-sequence data used for these analyses is freely available on the project GitHub site located at https://github.com/jbclayton83/douc-microbes-paper.

Analysis of diet components. One population of wild doucs was observed between January and August 2013 in Son Tra Nature Reserve, Da Nang, Vietnam. All occurrences of observed feeding behaviors were recorded.
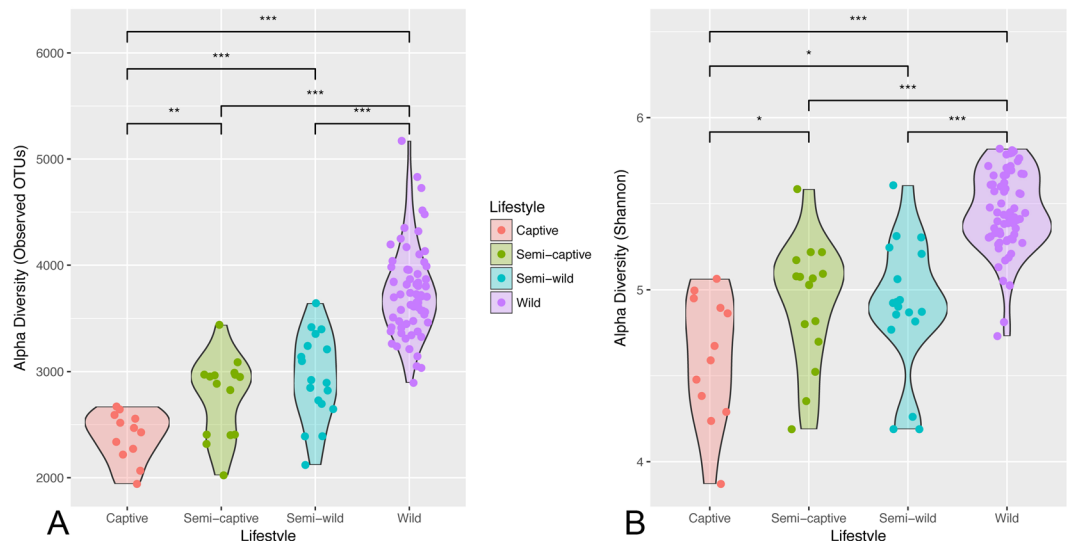
Figure 1. Diminished alpha diversity in red-shanked douc microbiomes across lifestyles. Violin plots of gut microbial alpha diversity across the 4 lifestyles according to (**A**) the number of species-like operational taxonomic units (OTUs) generated by open-reference OTU picking in the gut microbiome, and (**B**) the Shannon diversity index. The width of the shape corresponds to the distribution of samples (strips overlaid as strip chart), and asterisks denote significant differences at Welch's t-test *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$. Under both metrics, the wild population exhibits the highest biodiversity, which appears to diminish as a gradient with level of captivity to the captive population, which has the lowest.

Identified plant parts ingested were recorded and reachable feeding trees were marked. The plant parts of specific trees which were prevalent in their diet and were available in sufficient quantities were sampled and dried to 95% dry matter as per a previously established method[57]. Samples were sent to the Biochemical Lab at The Agriculture and Forestry University in Ho Chi Minh City, Vietnam, for chemical analysis. Concentrations of crude protein, simple sugars, crude fiber, calcium, sodium, manganese, potassium and iron were determined on a dry matter basis, all of which follow AOAC methods 920.152, 973.18C, and 974.06[58]. Additionally, all plants fed to semi-wild doucs during a two-week period in October 2012 were also sent for chemical analysis for comparison. Chemical compositions of semi-captive and captive diets were constructed using the precise dietary components administered by the facilities. Nutrient contents were compiled from laboratory nutrient analysis on a concentration per dry matter basis. Wild and semi-wild nutrient contents were constructed by observed frequency.

Results

Microbiome diversity declines according to lifestyle and habitat disruption.   Fecal microbiome diversity showed a steady decline from wild towards captive environments. The number of OTUs in the doucs decreased in a gradient-like fashion with the highest number in wild doucs (4231.68 ± 584.37 OTUs), and the lowest number in captive doucs (2332.08 ± 180.30 OTUs). Consistent with the gradient hypothesis, the semi-wild doucs (2845.50 ± 494.98 OTUs) and semi-captive doucs (2696.93 ± 417.00 OTUs) were intermediate. Pairwise comparison of all populations by OTU abundance showed statistically significant differences in OTU count biodiversity between all groups ($p < 0.01$) (Fig. 1A). In addition to investigating overall OTU diversity (i.e., number of OTUs) amongst the four unique douc populations, other alpha diversity (i.e., within-sample diversity) metrics were calculated. The mean Shannon diversity index, which measures species evenness, was highly significant across the four douc populations (wild: 7.86 ± 0.34; semi-wild: 7.07 ± 0.55; semi-captive: 7.11 ± 0.53; captive: 6.65 ± 0.53; ANOVA, $p = 4.3 \times 10^{-18}$). Based on the calculated Shannon diversity indices, the wild douc microbiome was the most even of the four douc populations (Fig. 1B). These trends are robust to collapsing samples by individual (Supplemental Fig. 4).

Beta diversity calculations were performed to assess whether significant differences between populations were present, using unweighted UniFrac distance (Fig. 2), as well as the weighted UniFrac and non-phylogenetic Bray-Curtis metrics (Supplemental Fig. 5). Analysis of unweighted UniFrac distance measurements is most effective at detecting differences in community membership when considering abundance differences among rare taxa[59]. An Adonis test on unweighted UniFrac distances revealed that fecal microbiome grouped statistically by douc population (Adonis $p = 0.001$), suggesting that each douc population had a unique microbiome. It also suggests that lifestyle has a major influence on gut microbial community structure, as doucs living under the most unnatural conditions had gut microbiomes most disparate from free-living wild doucs. Overall, the results of our beta diversity analyses indicated that microbiome composition was distinct for each of the four douc populations examined in this study at the 97% OTU and genus levels. Further, Fig. 2 reveals a clear gradient by naturalness of lifestyle along PC1, the primary axis of differentiation.
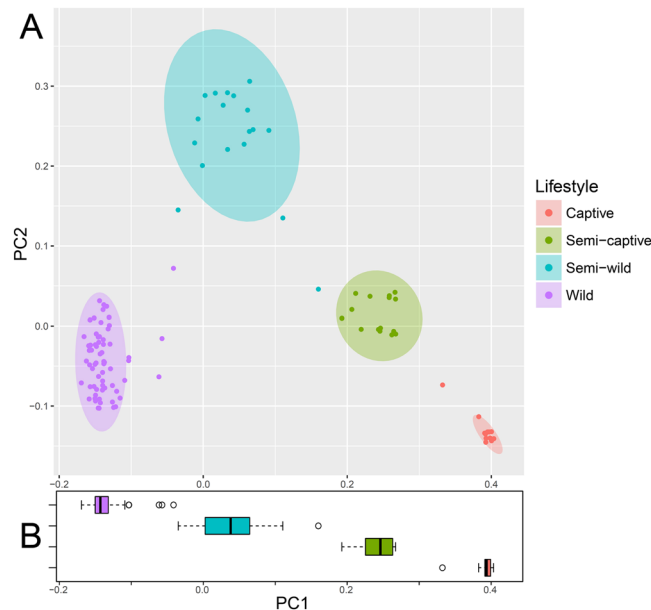
Figure 2. Principal coordinates plot showing (**A**) unweighted UniFrac ordination and (**B**) box plot of PC1 by population showing ecological distance between gut microbial communities in wild, semi-wild, semi-captive, and captive red-shanked doucs. All samples were obtained with the same protocol for V4 16S rRNA sequencing, and open-reference OTU picking was used. Douc microbiomes clearly clustered by population suggesting that each douc population had a unique microbiome, and thus were highly distinctive. Lifestyle has a major influence on gut microbial community structure, as doucs living under the most unnatural conditions (captive) had gut microbiomes most disparate from wild doucs (i.e., natural).

Differential taxonomic abundance analysis by lifestyle. Broad phylum-level taxonomic summarization revealed trends among the fecal microbiomes of the four douc populations included in this study. The fecal microbiomes of wild, semi-wild, semi-captive, and captive doucs were dominated by the phylum Firmicutes. Bacteroidetes was found in very low abundance in both the wild and semi-wild populations. In contrast, Bacteroidetes was the second most abundant phylum found in both the semi-captive and captive populations. Additionally, Verrucomicrobia was much more abundant in the semi-wild fecal microbiome than the other lifestyles examined (Supplemental Figs 6,7). *Bacteroides* and *Prevotella*, as well as *Methanosphaera*, *CF231*, *Treponema*, and *YRC22* were highly positively correlated with captivity level (all polyserial rho ≥ 0.71, $p < 1 \times 10^{-6}$) (Figs 3 and 4; Supplemental Fig. 7; Supplemental Table 1a). Conversely, *Adlercreutzia*, *Anaerostipes*, *Blautia*, *Campylobacter*, *Dehalobacterium*, *Dorea*, and *Oscillospira* were much less abundant with increasing captivity (all polyserial rho ≤ -0.65, $p < 1 \times 10^{-6}$) (Figs 3 and 4; Supplemental Fig. 7; Supplemental Table 1a). Although the genus *Akkermansia* shows a similar trend (polyserial rho $= -0.57$, $p = 1.29 \times 10^{-09}$), its abundance peaks slightly in the semi-wild population (Figs 3 and 4; Supplemental Fig. 7). These taxonomic trends are robust to collapsing samples by individual (Supplemental Table 1b; Supplemental Fig. 8).

A full-rank, species-level, plant-inclusive heatmap was also generated under similar feature selection criteria (absolute polyserial rho > 0.3, rho $p < 0.05$, pairwise $p < 0.05$) but with sample (column) clustering also performed to ascertain whether hierarchical clustering of these taxa abundance profiles alone would be able to recover a similar separation between lifestyles as observed in the Unweighted UniFrac ordination. The unsupervised clustering of these features correctly recovered the group membership of all samples, as observed by the preservation of the sample group labels without gaps or shuffling between lifestyles (Supplemental Fig. 9).

In addition to examining relative abundances of bacterial taxa between douc groups, we calculated and compared the log *Firmicutes* to *Bacteroidetes* ratio for each douc group (Supplemental Fig. 10A,B). The log of the *Firmicutes* to *Bacteroidetes* (F:B) ratio, which has been suggested as a measure of energy harvest capacity by microbial communities[5,19,60], was higher in the wild population than in the semi-wild, semi-captive, and captive populations (4.64 ± 0.94; 3.78 ± 1.14; 1.94 ± 0.81; 1.43 ± 0.50, respectively). A Kruskal-Wallis test indicated there is a significant difference in F:B ratio between at least one lifestyle group and the others. Pairwise significance between specific groups, with the exception of captive versus semi-captive (Wilcoxon rank-sum $p = 0.05$), were significantly different from one another for all lifestyle pairs with $p < 0.01$. In fact, there appears to be a relationship between lifestyle and the F:B ratio, as we see the highest ratio in wild doucs, the second highest ratio in the semi-wild doucs, the third highest ratio in semi-captive doucs, and finally the lowest ratio in captive doucs (Supplemental Fig. 10a). These trends are robust to collapsing samples by individual (Supplemental Fig. 10b).

Red-shanked douc metagenome: Functional analysis using PICRUSt. The functional profiles of the microbial sample in this study were investigated employing PICRUSt. In captivity, we observed a general trend toward increased protein metabolism at the expense of fatty acid metabolism. Specifically, the KEGG
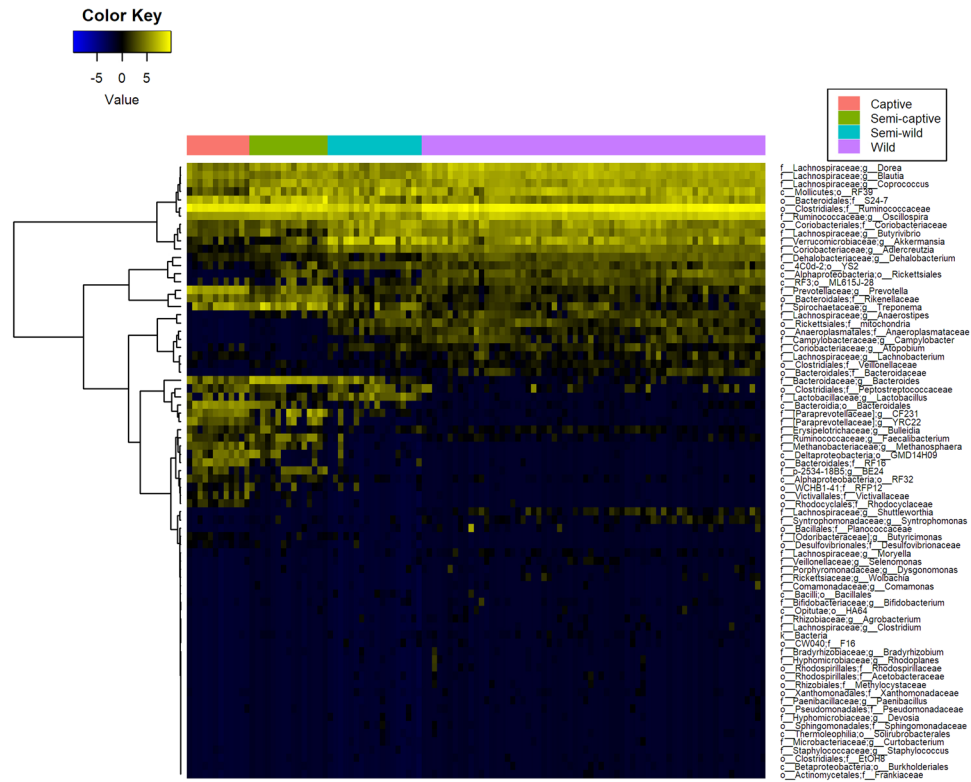
Figure 3. Heatmaps of differentially abundant microbial taxa at the genus level in red-shanked doucs living four distinct lifestyles. Taxa are displayed with polyserial correlations (rho) above 0.3, rho estimate adjusted p < 0.05, and (pairwise) Wilcoxon rank-sum adjusted p-value for combined captive & semi-captive versus wild lifestyles <0.05. Color represents intensity of centered log ratio abundances along gradient of color scale shown.

Ortholog (KO) super-heading "Amino acid metabolism" was highly correlated with captivity status (polyserial rho = 0.85, p = 2.2×10^{-11}) and the super-heading "Lipid metabolism" was highly anticorrelated (polyserial rho = −0.89, p = 7.3×10^{-12}). Perhaps due to the presence of chloroplasts in the closed-reference data used for PICRUSt analysis, the photosynthesis and antenna proteins pathway was downregulated in captivity, but the porphyrin and chlorophyll metabolism pathway was upregulated. With the exception of tetracycline biosynthesis, antibiotics-related pathways, including vancomycin biosynthesis, beta-lactam resistance, and penicillin & cephalosporin biosynthesis, were positively associated with captivity. Certain xenobiotic (mainly industrial pollutants) degradation pathways were positively associated with captivity, including ethylbenzene, styrene, and toluene. Other xenobiotic pathways (such as plant toxins and wartime chemicals) were negatively associated with captivity, including xylene, dioxins, atrazine, and chloroalkanes & chloroalkenes. Lastly, chemotaxis, invasion, flagellar assembly, and cytoskeleton genes were enriched in wild doucs. All p-values for results in this paragraph were less than <1×10^{-2} (Fig. 5; Supplemental Table 2a). These trends are robust to collapsing samples by individual (Supplemental Table 2b).

Composition of douc diets. The diets of the douc populations were compared to determine what factors, if any, could have contributed to the differences in microbiome composition observed (Table 1). Wild doucs fed on 57 different plant species. Sixty-one percent of all identified plant parts observed being ingested were collected and chemically analyzed. The semi-wild douc population were offered 60 plant species over the course of one year, 16 of which were never consumed[61]. In contrast to the high diet diversity (i.e., number of plant species) consumed by the wild and semi-wild doucs, the semi-captive and captive doucs consumed a much less diverse diet. Specifically, the semi-captive doucs were presented with approximately 15 plant species and the captive douc diet only contained one plant species[31] (Fig. 6A). The semi-wild population was observed feeding on 35 different plant genera over one year, and the wild population was observed feeding on 41 different plant genera over approximately seven months (Fig. 6B; Table 2).

We estimated total raw plant dietary content for the four douc populations, using chloroplast sequences observed in the 16S amplicon sequencing data. We found that chloroplast content was substantial in wild and semi-wild populations, but that chloroplast content decreased dramatically in semi-captive and captive doucs (Fig. 7A). Alignment of the 16S sequences to known plant reference genomes at 95% identity yielded a heatmap similar to Supplemental Fig. 5b. Overall, there is a clear trend toward increased plant abundances in the semi-wild and wild lifestyles, although certain orders display different abundances between lifestyle groups. Some orders can be seen to overlap between lifestyles, while others do not (Fig. 7B; Supplemental Table 3).
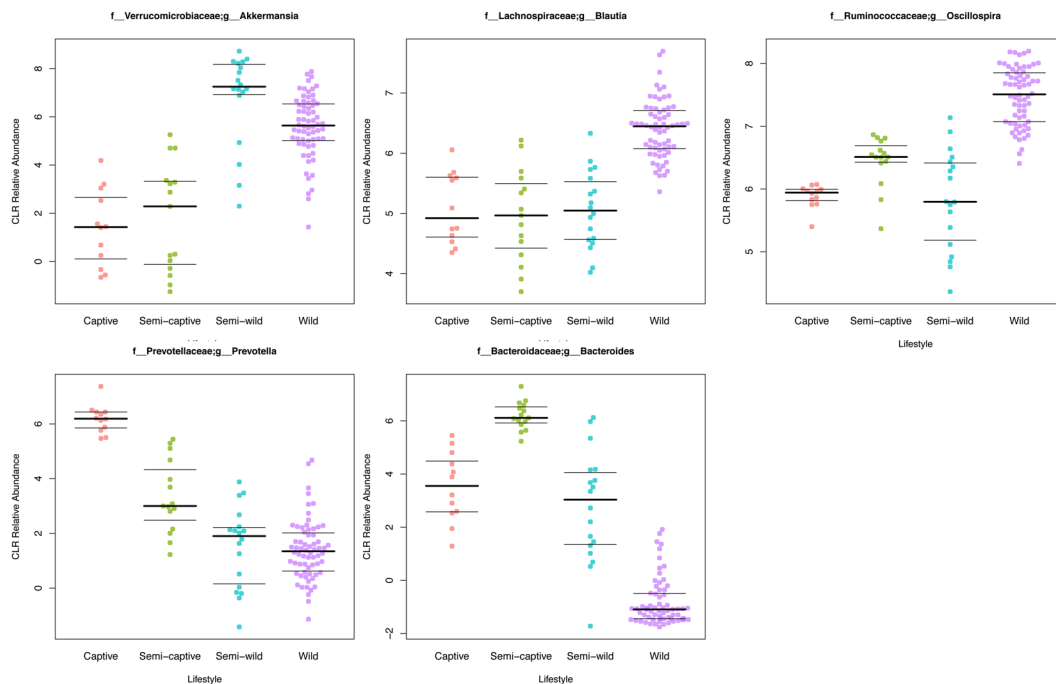
Figure 4. Beeswarm plots displaying gradient-like patterns of selected microbial taxa. A beeswarm plot of the arcsine square root relative abundance of bacterial genera *Bacteroides*, *Prevotella*, *Oscillospira*, *Blautia*, and *Akkermansia* shown in wild, semi-wild, semi-captive, and captive douc populations. All samples were obtained with the same protocol for V4 16S rRNA sequencing, and open-reference OTU picking was used. Red-shanked doucs acquire *Bacteroides* and *Prevotella*, and lose *Oscillospira*, *Blautia*, and *Akkermansia* in captivity. The presence of *Akkermansia* was most associated with a semi-wild lifestyle.
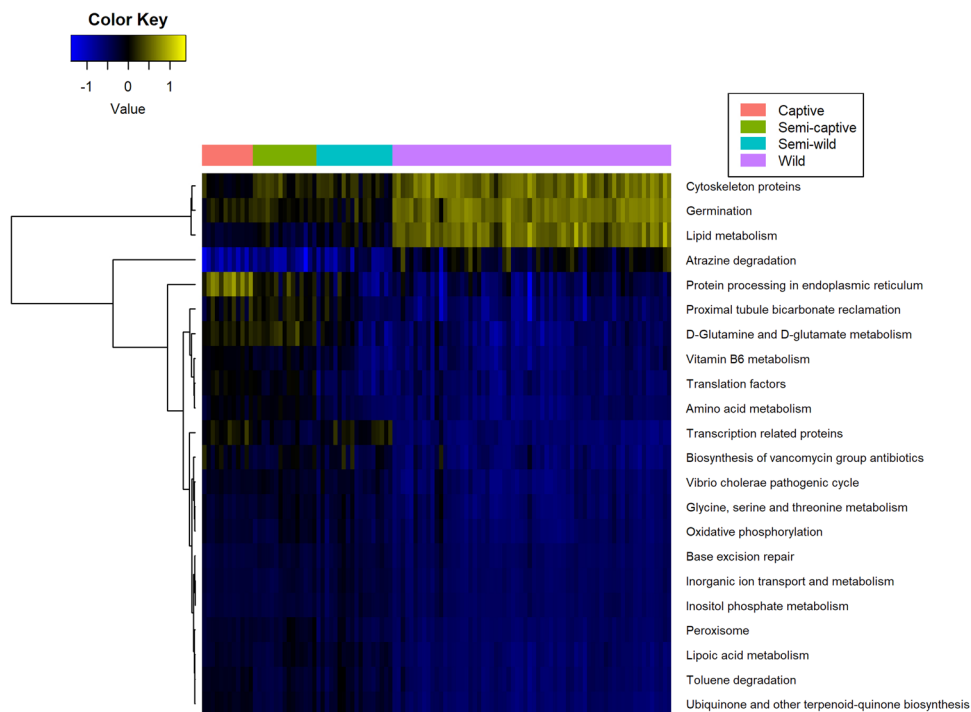
Figure 5. Heatmap of KEGG level 3 metabolic pathways in red-shanked douc groups living four distinct lifestyles. Pathways are displayed with polyserial correlations (rho) above 0.75, rho estimate adjusted p < 0.05, and (pairwise) Wilcoxon rank-sum adjusted p-value for combined captive & semi-captive versus wild lifestyles <0.05. Color represents intensity of unit-normalized centered log ratio abundances along gradient of color scale shown.

Dietary Groups	Proportion of Diet (%)			
	Wild	Semi-wild	Semi-captive	Captive
Leaves*	65.50	100.00	66.10	1.40
Flowers	5.30	0.00	0.00	0.00
Seeds	2.00	0.00	0.00	0.00
Unripe Fruit	17.80	0.00	0.00	0.00
Other plant parts**	9.40	0.00	0.00	0.00
Pellets	0.00	0.00	0.90	7.30
Fruits and Vegetables	0.00	0.00	33.00	90.90
Cereals	0.00	0.00	0.00	0.40

Table 1. Dietary components of red-shanked doucs living four distinct lifestyles, including wild, semi-wild, semi-captive, and captive. *Values for leaves may also include petioles or stems. **Other plant parts include pith, bark and leaf buds.
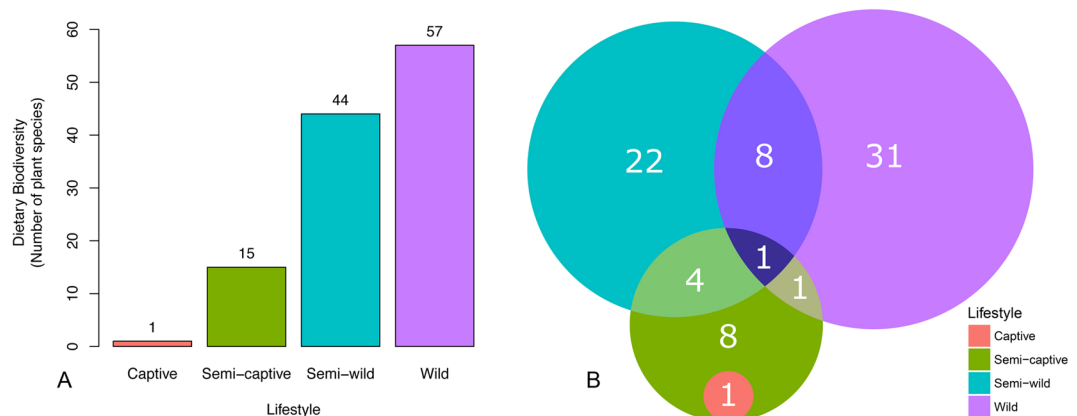
Figure 6. Plant diversity in red-shanked douc diet reflects dietary diversity across populations. (**A**) Bar plots of dietary biodiversity, as measured by the number of plant species consumed by wild, semi-wild, semi-captive, and captive populations of red-shanked doucs. Wild doucs feed on 57 different plant species, whereas the semi-wild doucs feed on 44 different plant species annually[61,113]. In contrast to the high dietary diversity consumed by the wild and semi-wild doucs, semi-captive and captive doucs are fed far fewer plant species. Specifically, semi-captive doucs are feed on approximately 15 plant species and the captive doucs are fed a single plant species[113]. Fruits and vegetables consumed by semi-captive and captive doucs were not a component of the "plant species" dietary category referred to in the Fig. 6. (**B**) Venn diagram depicting the number of plant genera consumed by the wild, semi-wild, semi-captive, and captive douc populations, while the numbers in overlaps representing the genera eaten by the constituent populations. Number of genera for the wild population was obtained from Clayton (unpublished). Number of genera for the semi-wild population was obtained from a combination of Clayton (unpublished) and Otto[61].

Nutritional composition of food items included in wild, semi-wild, semi-captive, and captive douc diets differed. Specifically, the crude protein concentration of the semi-wild, semi-captive, and captive douc diets was higher than that of the wild douc diet. The wild and semi-wild douc diets contained much more Acid Detergent Fiber (ADF) and Neutral Detergent Fiber (NDF) than did the semi-captive and captive diets. We examined the four douc diets for differences in amount of three macrominerals, including calcium, potassium, and sodium. Of the diets examined, the semi-wild douc diet contained more calcium than did the wild, semi-captive, and captive douc diets. Additionally, the diet consumed by wild doucs contained more potassium than did the diets consumed by semi-wild, semi-captive, or captive doucs. The diets of semi-wild and wild doucs contained considerably less sodium than the semi-captive and captive doucs. The semi-wild douc diet contained more iron and zinc than the wild, semi-captive, or captive diets. The concentration of sugar was not available for all lifestyle groups. The captive douc diet had the highest amount of soluble sugars compared to wild and semi-wild diets (Table 3).

Discussion

In this study, the douc was used to study the relationships between lifestyle and microbial composition and function within the gastrointestinal tract. Doucs are folivorous Old World monkeys, that are anatomically, physiologically, and ecologically unique amongst the living primates[62]. They possess specialized GI systems similar to ruminants, allowing for the digestion and utilization of extremely high fiber diets[10,11]. For doucs, mutualistic microbial populations are indispensable to digestive processes such as the fermentation of polysaccharides and subsequent production of short-chain fatty acids[63–65]. Although the digestive specializations possessed by

Diets by Lifestyle	Captive	Semi-captive	Semi-wild[a]	Wild
Plant genus/genera	*Morus	Acalypha **Adenanthera Asystasia **Cinnamomum **Hibiscus Khaya **Leucaena *****Litsea Moringa *Morus Polygonum Pterocarpus ***Syzygium Tamarindus Terminalia	**Adenanthera Alangium Albizzia Alchornea Averrhoa Bidens Cassia ****Celtis **Cinnamonum ****Claoxylon Clerodendrum Crateva ****Dalbergia Delonix Derris Dimocarpus Elaeocarpus Euodia ****Ficus **Hibiscus ****Ilex **Leucoena *****Litsea Maesa Micromelum Mussaenda ****Oroxylon Phyllanthus Rhus Saurauia Sterculia Triadica ****Vitex Wrightia ****Zanthoxylum	Acacia Amesiodendron Ancistrocladus Antidesma Aporusa Barringtonia Beilschmiedia Bischofia Brownlowia Callerya ****Celtis ****Claoxylon Cleidion ****Dalbergia Diospyros Dipterocarpus ****Ficus Glochidion Gluta Gmelina ****Ilex Litchi Lithocarpus *****Litsea Mallotus Millettia Mischocarpus Nauclea Ochrocarpus Ormosia ****Oroxylon Parashorea Quercus Rothmannia Sandoricum Schefflera Scolopia ***Syzygium Uvaria ****Vitex ****Zanthoxylum

Table 2. Plant genera observed in douc populations by lifestyle. [a]Combination of data collected by JBC for this study and data lifted from Otto[61]. *Shared between captive and semi-captive. **Shared between semi-captive and semi-wild. ***Shared between semi-captive and wild. ****Shared between semi-wild and wild. *****Shared between semi-captive, semi-wild, and wild.
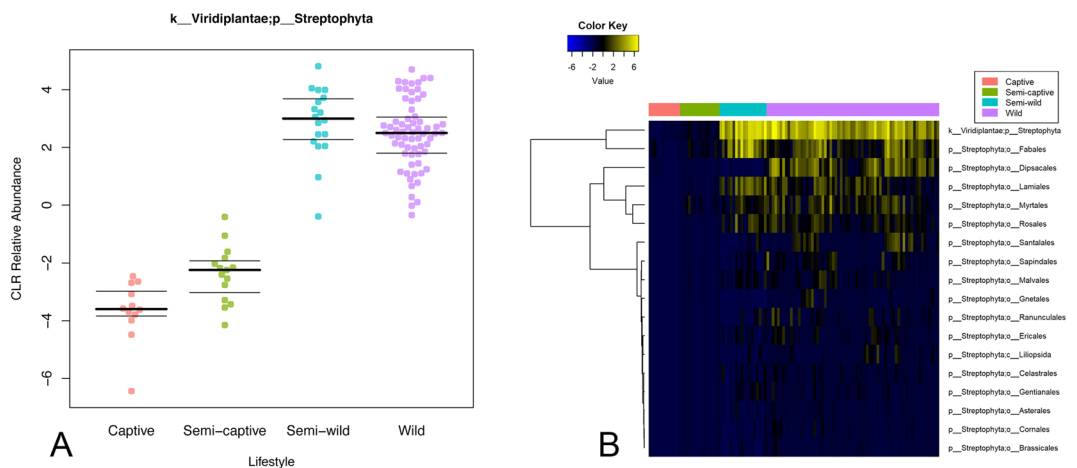
Figure 7. Plant chloroplast material in douc feces increases with wildness of lifestyle. (**A**) Centered-log-ratio-corrected relative abundance beeswarm plot. (**B**) Heatmap of class/order-level representation of plant taxa by lifestyle. All land plants (phylum Streptophyta) for which refseq chloroplast sequences were available were searched for matches with the 16S data. There is trend toward increased plant matter with wildness (polyserial rho = −0.81, pairwise combined captive & semi-captive versus wild Wilcoxon rank-sum adjusted p = 8.65×10^{-13}). Certain plant taxa display different patterns of overlapping abundance between lifestyles.

doucs have allowed them to thrive in their native habitat, the same specializations appear to challenge their survival in captivity, as they are highly susceptible to gastrointestinal disorders when maintained on commercially prepared diets in captive situations[12–14]. In order to better understand the link between lifestyle, gut microbial

Diet	Wild[1]	Semi-wild[a]	Semi-captive[a]	Captive[a]
Crude Protein (%)	9.46	16.52	13.37	16.70
Crude Fat (%)	—	3.23	3.12	3.71
Soluble Sugars (%)	2.70	2.28	—	7.90
ADF (%)	46.76[b]	23.20[c]	23.07	8.65
NDF (%)	53.67[b]	35.60[c]	31.97	12.64
Calcium (%)	0.49	1.05	0.22	0.72
Potassium (%)	0.96	0.76	0.21	0.29
Sodium (%)	0.01	0.01	0.24	0.27
Zinc (mg/kg)	19.39	101.59	8.25	26.30
Iron (mg/kg)	26.50	337.32	20.74	64.33

Table 3. Nutrient content on a dry matter basis from red-shanked doucs living four distinct lifestyles, wild, semi-wild, semi-captive, and captive. Semi-captive diet also included a vitamin and mineral supplement which was not included in the analysis. [a]Values from Clayton et al.[31]. NDF (Neutral Detergent Fiber) and ADF (Acid Detergent Fiber) were not available from the laboratory analyses, therefore values for [b]were taken from Ulibarri[33] and [c]were lifted from Otto[61] in order for a comparison to be possible.

communities, and health, we examined the fecal microbiomes of four douc populations living four distinct lifestyles (wild, semi-wild, semi-captive, and captive).

Microbial diversity. Studies have shown that present-day (i.e., modern) humans have lost a considerable portion of their natural (i.e., historical) microbial diversity[35,66,67]. Reduced bacterial diversity is often viewed as a negative indicator of health[68]. 16S rRNA sequencing results revealed that captive doucs had a marked reduction in gut bacterial alpha diversity when compared to wild and semi-wild doucs, as we reported recently[31]. Considering that doucs often fail to thrive in captivity, this was a salient finding. Not only was a reduction in diversity detected in captive doucs, but a gradient-like decrease in diversity related to lifestyle was observed, as the level of diversity observed in the semi-wild doucs was intermediate between wild and semi-captive doucs, and the level of diversity seen in semi-captive doucs was intermediate between semi-captive and captive doucs. This trend is consistent between metrics of richness and evenness, and suggest that lifestyle factors, especially dietary composition, and gut bacterial diversity are interrelated for doucs. This is similar to what has been shown in humans and other organisms[35,69–73].

Beta-diversity metrics revealed each douc population had a unique microbiome. Weighted UniFrac ordination, although maintaining clear group separation, did not recover a clear gradient, whereas Bray-Curtis appears similar to Unweighted UniFrac in visibly resolving the lifestyle gradient with PC1. This implies that the taxonomic membership of the gut microbiome, more than the abundance of each member, plays a dominant role in uncovering the gradient between lifestyles. Moreover, tree-independent unsupervised hierarchical clustering, utilizing only the taxonomic features significantly correlated to lifestyle gradient, preserved lifestyle group membership without erroneously assigning any samples to the wrong lifestyle groups, and found further structure within each group. This highlights the potential for using these taxa, or a subset thereof, as biomarkers capable of accurately differentiating between lifestyles.

Lifestyle factors drive gut microbial community structure. Establishing a link between lifestyle and the microbiome, with specific emphasis on diet as a major component of lifestyle, was a major focal point of this study. GI microbiome composition is shaped by host genetics and environment, among many factors[25,74]. Examining four populations of the same NHP species living four very different lifestyles enabled assessment of the contribution of environmental factors independent of interspecific host variation towards shaping the microbiome. The various interacting environmental differences to which each douc population is exposed, such as climate and diet, suggest that lifestyle plays a fundamental role in shaping gut microbiome composition in wild and captive NHPs. Of these lifestyle factors that contribute to the establishment and maintenance of the gut microbiota, diet is likely the most influential, as studies have shown that changes in diet are directly associated with shifts in gut microbial community structure[74–78]. Examples exist of species adapting to specific dietary niches in both wild[79] and captive[73] settings via changes in their gut microbiota.

Many of our results suggest the existence of a relationship between microbiome composition and dietary patterns specifically. The relative abundance of select bacterial genera, *Bacteroides*, *Prevotella*, *Oscillospira*, and *Blautia*, and differences in dietary composition between lifestyles, warrant further discussion in this regard. *Prevotella*, which is involved in the digestion of simple sugars and carbohydrates[80], was notably higher in the captive doucs than in the other three douc populations examined. One very different component in the diets of wild versus captive doucs is the inclusion of produce in the captive diets. Fruits consumed by captive primates have much different nutrient profiles than fruits consumed by wild primates[81]. They have been cultivated for human consumption to be lower in fiber and protein and higher in sugar as opposed to wild fruits which are, in general, the exact opposite[82]. Given the diet of captive doucs contained more than a threefold increase in the percentage of sugars compared to wild and semi-wild douc diets, there seems to be a clear relationship between sugar consumption and *Prevotella* abundance. Unlike captive doucs, wild doucs consumed unripe fruit, which is drastically different than ripe fruit fed to captive individuals, and thus its impact on the douc microbiome is different than

cultivated fruits would have. The low relative abundance of *Prevotella* in the wild douc microbiome provides further evidence that diet, perhaps in tandem with other lifestyle components, is a major driver of microbiome composition.

Semi-captive and captive doucs also harbored more *Bacteroides* than semi-wild or wild doucs. In humans, *Bacteroides* are found in higher abundance in individuals who consume diets high in fat and protein[34,77]. Interestingly, the captive douc diet contained more protein than the wild douc diet, which may explain why *Bacteroides* was a dominant member of the captive douc microbiome, yet virtually absent from the wild douc microbiome.

Another notable example of the specific diet-microbiome relationship seen in our analysis was with the genus *Oscillospira*, which has a known association with the consumption of plant material, including leaves and grass cuticles[83–86]. *Oscillospira* was markedly increased in the wild doucs compared to the other douc populations, and was more abundant in the semi-wild and semi-captive doucs than in the captive population. The observed differences in abundance of *Oscillospira* between douc populations was likely a function of the stark differences in dietary consumption between populations, and most importantly, the difference in diversity and proportion of plants and plant parts consumed by the different douc populations examined. Wild, semi-wild, and, to a lesser degree, semi-captive doucs all consume diets that contain a higher proportion and diversity of plants compared to the captive population. Aside from *Oscillospira*, overall douc microbiome composition seemed to be, at the very least, partially driven by plant abundance and diversity in the diet. While a high variety of plant species is bound to impact the gut microflora, the sheer differences in proportion of the diet which is plants is likely to be equally as much of a causative factor[87]. In this study, we utilized both known dietary makeup and measured chloroplast content (via 16S rRNA sequencing) to detail plant consumption by each douc population. We observed a striking downward trend in the number of plant genera and species consumed by increasing captivity of lifestyle (Fig. 6A,B). The measured chloroplast content of the stool mirrored this downward trend (Fig. 7A). Further, we were able to observe trends in plant taxon composition, although the resolution of the chloroplast analysis was limited approximately to the plant class level. Some of the patterns of overlap observed mirrored the overlap in plant genera fed to the doucs, but we were unable to confirm whether the presence of the plant classes/orders reported in the chloroplast analysis indeed coincided with the specific plants fed to the doucs in the different lifestyles. Further targeted plant genomic screens may expand upon this proof of concept in the future.

An unexpected result found in this study was the high abundance of the genus *Akkermansia* found in the semi-wild doucs. While *Akkermansia* was most abundant in the semi-wild doucs by far, this genus was also more abundant in the wild doucs than doucs living semi-captive and captive lifestyles. Members of the genus *Akkermansia*, such as *Akkermansia muciniphila*, are known for their roles in mucin-degradation, and have been suggested to play protective roles in the gut[88,89]. Everard *et al.*[89] showed that obese and type 2 diabetic mice had decreased abundance of *A. muciniphila*, and treatment with this microbe reversed high-fat diet-induced metabolic disorders. Interestingly, a recent study examining the link between gut microbiota and primate GI health found GI-unhealthy doucs had reduced relative abundances of *Akkermansia*[90]. Another study examining gut microbiome composition of a colobine primate, *Rhinopithecus brelichi*, showed that *Akkermansia* was more abundant in captive individuals when compared to their wild counterparts, which is different than what was seen in doucs[91]. The extremely high abundance of *Akkermansia* in the semi-wild doucs combined with the higher level seen in wild doucs compared to semi-captive and captive doucs suggests that the microbe is indirectly linked to diet, as the diets of wild and semi-wild doucs contain much more fiber compared to those of the captive doucs. Considering that high fiber diets increase mucin thickness, and *Akkermansia* is a strict mucin degrader, the higher abundance of *Akkermansia* seen in doucs consuming higher fiber diets logical[92].

We examined the F:B ratio, as this ratio is important in humans in terms of dietary energy extraction[5,93]. We saw a higher F:B ratio in wild and semi-wild doucs compared to semi-captive and captive doucs. Ley *et al.*[19] found an increased presence of Firmicutes with a corresponding decrease of Bacteroidetes correlating with an overall greater energy harvest[19]. Based on our results, it appears a decrease in the F:B ratio was clearly associated with lifestyle, notably diet, as the wild doucs had the highest ratio, followed by the semi-wild doucs, semi-captive doucs, and captive doucs. As previously mentioned, wild and semi-wild diets contained substantially more plant matter than captive diets. Naturally this equates to diets much higher in fiber fractions (ADF, NDF). Due to the scarcity of high-quality food items in the wild, we witnessed doucs ingesting very fibrous plant parts such as bark, mature leaves, flowers, seeds and unripe fruit. In the semi-wild facility, the doucs are habituated and know that they will receive leaf meals which provides them with a balance of fiber and soluble nutrients, making the ingestion of very fibrous items such as bark unnecessary. This can partially explain the higher reported NDF values in wild doucs when compared to semi-wild doucs. This relationship between F:B ratio and diet was expected, as our results show captive populations have diets lower in fiber fractions and higher in soluble carbohydrates, notably sugars, when compared to wild or semi-wild populations. It is plausible that wild doucs, which consume lower quality food items (bark, etc.), rely on a more efficient energy harvest strategy to obtain nutrients from their diet, and thus survival. Overall, the differences in the F:B ratio observed between populations living in natural versus unnatural settings, in addition to the presence of GI symptoms observed in the semi-captive and captive doucs in this study, suggests the ratio is a potential indicator of GI perturbation. A higher ratio is associated with a higher fermentation efficiency and increased VFA production[5,79], but also obesity in humans[5]. Furthermore, doucs living under artificial (i.e., captive) conditions, which had a lower ratio in this study, often suffer from a wasting syndrome[94,95].

Putative functional associations with lifestyle.
PICRUSt-predicted functional pathways show a few interesting trends. Interestingly, captivity appears to be correlated with pathways spanning metabolism of antibiotics, nutrients, and xenobiotics, as well as other potentially relevant trends. In terms of antibiotics, resistance to beta-lactam antibiotics (penicillin and cephalosporin) was predicted to be enriched in captivity alongside biosynthesis of the same. Tetracycline biosynthesis genes, in contrast, were predicted to be elevated in the wild lifestyle

and were not accompanied by a significant elevation in tetracycline resistance pathways. This pattern raises the possibility that the putative antibiotics pathways upregulated in captivity may be adapted for competition and virulence factor regulation[96], whereas the pathways upregulated in the wild may play more of a quorum sensing role[97].

Another interesting trend is the apparent tradeoff between amino acid metabolism in the wild and lipid metabolism in captivity, two important umbrella pathways (hierarchically from top level of KEGG, Metabolism ->Lipid Metabolism and Metabolism ->Amino Acid Metabolism). This pronounced trend is likely indicative of the highly different diets received by the populations and may reflect differences in plant species consumed, the differences in microbiome composition in response to different nutrient profiles, or other lifestyle factors including antibiotics exposure. Additionally, the differentially increased motility of the members of the wild microbiome (evidenced by putative upregulated cytoskeletal regulation, flagellar and motility proteins, and chemotaxis) may imply increased vigor and facilitate nutrient scavenging. The putative increasing differential abundance of sporulation and germination pathways with wildness may also highlight the resilience of the wild microbiota, and may be due in part to various members of order Clostridiales (many of which can form spores and later germinate from them) also being differentially more abundant.

There are some particularly intriguing trends concerning xenobiotic (pollutant) degradation. Xylene, dioxins, atrazine, and chloroalkane/ene degradation levels were predicted to be significantly associated with increasing wildness of lifestyle. Setting aside the expected presence of toxic chemicals in the douc's natural diet, all of these compounds have been associated with wartime chemicals. Specifically, agent orange and other war chemicals deployed during the Vietnam Conflict are atrazines and dioxins. Dioxins are also byproducts of forest fires and several types of manufacturing processes. They are persistent in the environment[98], and studies have shown that manufacturing workers who are in direct contact with dioxins have an increased risk for the development of cancer[99]. Son Tra Nature Reserve, the research site where wild douc fecal samples were collected, is located approximately 8 km from Da Nang International Airport. This airport is considered one of the world's most dioxin-contaminated sites. There are over 187,000 square meters of contaminated soil located in several sites near the airport[100]. Similarly, jet fuel and industrial solvents often contain xylene, another xenobiotic upregulated in wilder lifestyles. The semi-wild population, which shows a peak in xylene degradation, is located in Singapore, which contains the largest xylene plant[101]. When refined into xylyl, xylene was also used as a wartime riot control agent[102]. Interestingly, the reverse (i.e. showing an increase with captivity) is observed for the degradation pathways of other pollutants often associated with industrialized economies such as ethylbenzenes[103], styrenes (often industrially synthesized from ethylbenzenes)[104], and toluene[105], which may indicate forms of air contamination or other "first-world" pollutants. Given this, coupled with the fact that wild doucs consume plants rich in toxic compounds, we hypothesize that the microbiome of wild doucs serves a detoxification role for the animal and therefore is enriched for taxa with the ability to degrade local contaminants. Another interesting implication is that the microbiome may function as a geochemical sensor of environment, where passage through a host may modulate detection sensitivity for certain compounds through dietary biomagnification.

Viewing the microbiome as compositional data.   From a statistical standpoint, the treatment of microbiome data in a modern compositional framework[106] frees microbial composition data from the simplex and allows many standard univariate statistical analysis techniques to apply[107]. With the observation that most bacterial abundances post-CLR transformation appeared to be roughly gaussian, the polyserial correlation test became applicable to assess the degree by which each microbial or functional pathway abundance was correlated with the latent continuous variable captured by the four ordered "lifestyle" categories. For additional conservatism, the nonparametric Wilcoxon rank-sum test was also used on the two extrema of the assumed gradient (the "Wild" and combined "(Semi-)captive" lifestyles), and the significance criterion for an association was amended to require both strong polyserial correlation (absolute rho above 0.3) as well as corrected Wilcoxon p-values less than 0.05. Importantly, this compositional framework allows for the highly-powered and gradient-centric interpretation and visualization of microbial and functional data. As such, the pairwise testing of extrema is best regarded as a means to contextualize and filter the results of the gradient analyses, which account for all of the data and are better suited to characterize gradients.

By modeling the lifestyle groups themselves as a latent gradient variable, we are assessing microbial and functional relationships with a composite metric of "wildness" or "lifestyle perturbation," and hence avoid the use of other potentially confounded covariates such as climatic factors, diversity of plant species consumed, frequency of human contact, or health markers as individual proxies. In so doing, these and other covariates can be interpreted in relation to one another. Our focus in this work was to ascertain which microbes and microbially-deduced functional pathways were correlated with lifestyle and interpret these correlations in context of collected health and dietary data.

In terms of available health data, we did have access to the health records for the semi-captive and captive doucs. After reviewing the health records, and speaking personally with our collaborators at the institutions where the doucs were housed, we determined that four of seven semi-captive doucs and one of two captive doucs we sampled died of gastrointestinal-related diseases, including wasting and gastroenteritis, within one year after sample collection took place. Given that the douc lifestyles had very different gut microbiomes, including semi-captive to wild and captive to wild, is suggestive of an association between douc lifestyle, gut microbiome, and disease. However, we were unable to test this association properly due to a limited sample size. To properly test this association, and produce conclusive evidence, would require a larger sample size.

This analysis revealed a selection of strongly correlated, statistically significant microbial biomarkers indicative of douc lifestyle, which may be related to health and wellbeing. These trends may have implications to human and livestock health, as the douc may serve as a genetically similar model for the former, and a digestively similar model for the latter. For example, in humans, increases in *Bacteroides* and *Preveotella*, which were detected in

the semi-captive and captive doucs, have also been seen in individuals with colorectal cancer. This is one major example of a potential human health link highlighted by this study, which warrants further investigation in future work[108–112]. Since the microbiome itself is associated with host genetics as well as digestive functions and diseases[23], this study provides the framework for and invites further investigation of this potential model organism for the applicability of these findings within other species.

References

1. Cho, I. & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nature Reviews Genetics* **13**, 260–270 (2012).
2. Clemente, J. C., Ursell, L. K., Parfrey, L. W. & Knight, R. The impact of the gut microbiota on human health: an integrative view. *Cell* **148**, 1258–1270 (2012).
3. Consortium, H. M. P. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
4. Bercik, P., Collins, S. & Verdu, E. Microbes and the gut-brain axis. *Neurogastroenterology & Motility* **24**, 405–413 (2012).
5. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031, https://doi.org/10.1038/nature05414 (2006).
6. Smith, M. I. *et al.* Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* **339**, 548–554 (2013).
7. Turnbaugh, P. J., Bäckhed, F., Fulton, L. & Gordon, J. I. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell host & microbe* **3**, 213–223 (2008).
8. Turnbaugh, P. J. *et al.* The Effect of Diet on the Human GutMicrobiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice. *Science Translational Medicine* **1**, 6ra14–16ra14, https://doi.org/10.1126/scitranslmed.3000322 (2009).
9. Cho, I. *et al.* Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature* **488**, 621–626 (2012).
10. Chivers, D. J. in *Colobine monkeys: their ecology, behaviour and evolution*. 205–227 (Cambridge University Press, 1994).
11. Lambert, J. E. Primate digestion: Interactions among anatomy, physiology, and feeding ecology. *Evolutionary Anthropology* **7**, 8–20, https://doi.org/10.1002/(sici)1520-6505 (1998).
12. Agoramoorthy, G., Alagappasamy, C. & Hsu, M. J. Can proboscis monkeys be successfully maintained in captivity? A case of swings and roundabouts. *Zoo Biology* **23**, 533–544 (2004).
13. Nijboer, J. *Fibre intake and faeces quality in leaf-eating primates*. (University of, 2006).
14. Power, M. L., Toddes, B. & Koutsos, L. In *Nonhuman Primates in Biomedical Research (Second Edition)* (eds Mansfield, K., Tardif, S. & Morris, T.) 269–286 (Academic Press, 2012).
15. Bauchop, T. & Martucci, R. W. Ruminant-like digestion of the langur monkey. *Science* **161**, 698–700 (1968).
16. Edwards, M. S., Crissey, S. D. & Oftedal, O. T. In *Fact Sheet 007* (ed. Nutrition Advisory Group) 7 (1997).
17. Ensley, P. K. *et al.* Intestinal obstruction and perforation caused by undigested Acacia sp leaves in langur monkeys. *J Am Vet Med Assoc* **181**, 1351–1354 (1982).
18. Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci USA* **104**, 13780–13785, https://doi.org/10.1073/pnas.0706625104 (2007).
19. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–1023, https://doi.org/10.1038/4441022a (2006).
20. Overskei, T. L. *et al.* Entamoeba histolytica infection in hanuman (Semnopithecus entellus) and purple-faced (Trachypithecus vetulus) langurs. *Journal of Zoo and Wildlife Medicine* **25**, 240–247 (1994).
21. Zhang, C. H. *et al.* Interactions between gut microbiota, host genetics and diet relevant to development of metabolic syndromes in mice. *Isme J* **4**, 232–241, https://doi.org/10.1038/ismej.2009.112 (2010).
22. Amato, K. R. *et al.* Habitat degradation impacts black howler monkey (Alouatta pigra) gastrointestinal microbiomes. *The ISME journal* **7**, 1344–1353 (2013).
23. Knights, D. *et al.* Complex host genetics influence the microbiome in inflammatory bowel disease. *Genome medicine* **6**, 107 (2014).
24. David, L. A. *et al.* Host lifestyle affects human microbiota on daily timescales. *Genome biology* **15**, R89 (2014).
25. Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. *Cell* **159**, 789–799, https://doi.org/10.1016/j.cell.2014.09.053 (2014).
26. Ngoc Thanh, V., Lippold, L. K., Timmins, R. J. & Manh Ha, N. In *IUCN 2011* (2008).
27. Anderson, K. E. *et al.* Highly similar microbial communities are shared among related and trophically similar ant species. *Mol Ecol* **21**, 2282–2296, https://doi.org/10.1111/j.1365-294X.2011.05464.x (2012).
28. Roeselers, G. *et al.* Evidence for a core gut microbiota in the zebrafish. *Isme J* **5**, 1595–1608, https://doi.org/10.1038/ismej.2011.38 (2011).
29. Wang, M. *et al.* Mode of delivery and early nutrition modulate microbial colonization and fermentation products in neonatal piglets. *The Journal of nutrition* **143**, 795–803, https://doi.org/10.3945/jn.112.173096 (2013).
30. Zhao, L. *et al.* Quantitative genetic background of the host influences gut microbiomes in chickens. *Scientific reports* **3**, 1163, https://doi.org/10.1038/srep01163 (2013).
31. Clayton, J. B. *et al.* Captivity humanizes the primate microbiome. *Proc Natl Acad Sci USA* **113**, 10376–10381, https://doi.org/10.1073/pnas.1521835113 (2016).
32. Lippold, L. K. & Thanh, V. N. The time is now: survival of the douc langurs of Son Tra, Vietnam. *Primate Conservation* **23**, 75–79 (2008).
33. Ulibarri, L. R. *The socioecology of red-shanked doucs (Pygathrix nemaeus) in Son Tra Nature Reserve, Vietnam* PhD thesis (University of Colorado, USA, 2013).
34. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227, https://doi.org/10.1038/nature11053 (2012).
35. Clemente, J. C. *et al.* The microbiome of uncontacted Amerindians. *Sci Adv* **1**, https://doi.org/10.1126/sciadv.1500183 (2015).
36. Yu, Z. & Morrison, M. Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* **36**, 808–813 (2004).
37. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Isme J* **6**, 1621–1624, https://doi.org/10.1038/ismej.2012.8 (2012).
38. Gilbert, J. A. *et al.* Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project. *Standards in genomic sciences* **3**, 243–248, https://doi.org/10.4056/sigs.1433550 (2010).
39. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335–336, https://doi.org/10.1038/nmeth.f.303 (2010).
40. DeSantis, T. *et al.* NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic acids research* **34**, W394 (2006).
41. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461, https://doi.org/10.1093/bioinformatics/btq461 (2010).
42. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**, 5261–5267, https://doi.org/10.1128/AEM.00062-07 (2007).
43. Al-Ghalith, G. A., Ang, K., Hillmann, B. & Knights, D. *SHI7: A streamlined short-read iterative trimming pipeline.* https://github.com/knights-lab/shi7 (2017).

44. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
45. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
46. Al-Ghalith, G. A. & Knights, D. *BURST enables optimal exhaustive DNA alignment for big data*, https://github.com/knights-lab/burst (2017).
47. Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**, 8228–8235, https://doi.org/10.1128/AEM.71.12.8228-8235.2005 (2005).
48. McMurdie, P. J. & Holmes, S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one* **8**, e61217 (2013).
49. Oksanen, J. *et al.* The vegan package. *Community ecology package* **10**, 631–637 (2007).
50. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
51. Wickham, H. *ggplot2: elegant graphics for data analysis*. (Springer, 2016).
52. Langille, M. G. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature biotechnology* **31**, 814–821, https://doi.org/10.1038/nbt.2676 (2013).
53. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27–30 (2000).
54. Fox, J. Polycor: polychoric and polyserial correlations. *R package version 0.7-5*, http://cran/.R-project.org/package=polycor (2007).
55. Templ, M., Kowarik, A. & Filzmoser, P. Iterative stepwise regression imputation using standard and robust methods. *Comput Stat Data An* **55**, 2793–2806, https://doi.org/10.1016/j.csda.2011.04.012 (2011).
56. Tsilimigras, M. C. & Fodor, A. A. Compositional data analysis of the microbiome: fundamentals, tools, and challenges. *Annals of epidemiology* **26**, 330–335, https://doi.org/10.1016/j.annepidem.2016.03.002 (2016).
57. Barnett, A. *Expedition field techniques primates*. (Royal Geographical Society, 1995).
58. AOAC. *Official methods of analysis of AOAC International*. (AOAC International, 2012).
59. Chen, J. *et al.* Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* **28**, 2106–2113 (2012).
60. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484, https://doi.org/10.1038/nature07540 (2009).
61. Otto, C. *Food intake, nutrient intake, and food selection in captive and semi-free douc langur*. (Schuling Verlag, 2005).
62. Davies, A. G. & Oates, J. F. *Colobine Monkeys*. 415 (Cambridge Univ Press, 1994).
63. Jablonski, N. G. *The Natural History of the Doucs and Snub-nosed Monkeys*. 382 (World Scientific Publishing, 1998).
64. Nijboer, J., Clauss, M., Everts, H. & Beynen, A. *Effect of dietary fibre on the faeces score in colobine monkeys in Dutch Zoos* (2006).
65. Nijboer, J., Clauss, M. & Olsthoorn, M. Effect of diet on the feces quality in Javan langur (Trachypithecus auratus auratus). *J Zoo Wildl Med* (2006).
66. Martinez, I. *et al.* The gut microbiota of rural papua new guineans: composition, diversity patterns, and ecological processes. *Cell reports* **11**, 527–538, https://doi.org/10.1016/j.celrep.2015.03.049 (2015).
67. Moeller, A. H. *et al.* Rapid changes in the gut microbiome during human evolution. *Proc Natl Acad Sci USA* **111**, 16431–16435, https://doi.org/10.1073/pnas.1419136111 (2014).
68. Fujimura, K. E., Slusher, N. A., Cabana, M. D. & Lynch, S. V. Role of the gut microbiota in defining human health. *Expert Rev Anti-Infe* **8**, 435–454, https://doi.org/10.1586/ERI.10.14 (2010).
69. Kueneman, J. G. *et al.* In *Proc. R. Soc. B*. 20161553 (The Royal Society).
70. Kohl, K. D. & Dearing, M. D. Wild-caught rodents retain a majority of their natural gut microbiota upon entrance into captivity. *Environmental microbiology reports* **6**, 191–195 (2014).
71. McKenzie, V. J. *et al.* The Effects of Captivity on the Mammalian Gut Microbiome. *Integrative and Comparative Biology* **57**, 690–704 (2017).
72. Menke, S. *et al.* Gut microbiomes of free-ranging and captive Namibian cheetahs: diversity, putative functions, and occurrence of potential pathogens. *Molecular Ecology* (2017).
73. Kohl, K. D., Skopec, M. M. & Dearing, M. D. Captivity results in disparate loss of gut microbial diversity in closely related hosts. *Conservation physiology* **2**, cou009 (2014).
74. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563, https://doi.org/10.1038/nature12820 (2014).
75. Gophna, U. Microbiology. The guts of dietary habits. *Science* **334**, 45–46, https://doi.org/10.1126/science.1213799 (2011).
76. Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332**, 970–974, https://doi.org/10.1126/science.1198719 (2011).
77. Wu, G. D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108, https://doi.org/10.1126/science.1208344 (2011).
78. Xu, Z. & Knight, R. Dietary effects on human gut microbiome diversity. *The British journal of nutrition* **113**, S1–5, https://doi.org/10.1017/S0007114514004127 (2015).
79. Amato, K. R. *et al.* The role of gut microbes in satisfying the nutritional demands of adult and juvenile wild, black howler monkeys (Alouatta pigra). *Am J Phys Anthropol* **155**, 652–664, https://doi.org/10.1002/ajpa.22621 (2014).
80. Purushe, J. *et al.* Comparative genome analysis of Prevotella ruminicola and Prevotella bryantii: insights into their environmental niche. *Microbial ecology* **60**, 721–729, https://doi.org/10.1007/s00248-010-9692-8 (2010).
81. Oftedal, O. T. & Allen, M. E. In *Wild mammals in captivity: principles and techniques*. (eds Devra G. Kleiman *et al.*) 148–157 (University of Chicago Press, 1996).
82. Schwitzer, C. & Kaumanns, W. In *Zoo Animal Nutrition Vol. II* (eds Fidgett, A. *et al.*) 247–265 (Filander Verlag, 2003).
83. Clarke, R. T. Niche in Pasture-Fed Ruminants for the Large Rumen Bacteria Oscillospira, Lampropedia, and Quin's and Eadie's Ovals. *Appl Environ Microbiol* **37**, 654–657 (1979).
84. Mackie, R. I. *et al.* Ecology of uncultivated Oscillospira species in the rumen of cattle, sheep, and reindeer as assessed by microscopy and molecular approaches. *Applied and Environmental Microbiology* **69**, 6808–6815, https://doi.org/10.1128/AEM.69.11.6808-6815.2003 (2003).
85. Yanagita, K. *et al.* Flow cytometric sorting, phylogenetic analysis and *in situ* detection of Oscillospira guillermondii, a large, morphologically conspicuous but uncultured ruminal bacterium. *Int J Syst Evol Microbiol* **53**, 1609–1614, https://doi.org/10.1099/ijs.0.02541-0 (2003).
86. Zoetendal, E. G. *et al.* Distinct Microbiotas Are Present in Urban and Rural Native South Africans, and in African Americans. *Gastroenterology* **144**, S347–S347 (2013).
87. de Menezes, A. B. *et al.* Microbiome analysis of dairy cows fed pasture or total mixed ration diets. *FEMS microbiology ecology* **78**, 256–265 (2011).
88. Belzer, C. & de Vos, W. M. Microbes inside–from diversity to function: the case of Akkermansia. *Isme J* **6**, 1449–1458, https://doi.org/10.1038/ismej.2012.6 (2012).
89. Everard, A. *et al.* Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 9066–9071, https://doi.org/10.1073/pnas.1219451110 (2013).
90. Amato, K. R. *et al.* Using the gut microbiota as a novel tool for examining colobine primate GI health. *Global Ecology and Conservation* **7**, 225–237, https://doi.org/10.1016/j.gecco.2016.06.004 (2016).

91. Hale, V. L. R. *Co-evolution of gut microbes in colobine monkeys*. (Purdue University, 2014).
92. Desai, M. S. *et al*. A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell* **167**, 1339–1353. e1321 (2016).
93. Ley, R. E. *et al*. Evolution of Mammals and Their Gut Microbes. *Science* **320**, 1647–1651, https://doi.org/10.1126/science.1155725 (2008).
94. Crissey, S. D. & Pribyl, L. S. Utilizing wild foraging ecology information to provide captive primates with an appropriate diet. *The Proceedings of the Nutrition Society* **56**, 1083–1094 (1997).
95. Lacasse, C. *et al*. Taxus sp. intoxication in three Francois' langurs (Trachypithecus francoisi). *Journal of Veterinary Diagnostic Investigation* **19**, 221–224 (2007).
96. Balasubramanian, D. *et al*. Co-regulation of β-lactam resistance, alginate production and quorum sensing in Pseudomonas aeruginosa. *J Med Microbiol* **60**, 147–156 (2011).
97. Lu, L. *Autoinducer 2-based quorum sensing response of Escherichia coli to sub-therapeutic tetracycline exposure*. (Texas A&M University, 2006).
98. Schecter, A., Birnbaum, L., Ryan, J. J. & Constable, J. D. Dioxins: an overview. *Environmental research* **101**, 419–428 (2006).
99. Kogevinas, M. *et al*. Soft tissue sarcoma and non-Hodgkin's lymphoma in workers exposed to phenoxy herbicides, chlorophenols, and dioxins: two nested case-control studies. *Epidemiology*, 396–402 (1995).
100. Minh, N. H., Boivin, T., Canh, P. N. & Son, L. K. Comprehensive assessment of dioxin contamination in Da Nang airbase and its vicinities: Environmental levels, human exposure and options for mitigating impacts. *Interdisciplinary Studies on Environmental Chemistry-Environmental Research in Asia*, 21–29 (2009).
101. Tremblay, J.-F. Making Aromatics in Singapore. *Chemical & Engineering News Archive* **89**, 18–19, https://doi.org/10.1021/cen-v089n038.p018 (2011).
102. Olajos, E. J. & Stopford, W. *Riot Control Agents: Issues in Toxicology, Safety & Health*. (CRC Press, 2004).
103. Welch, V. A. & Fallon, K. J. Ethylbenzene. *Ullmann's Encyclopedia of Industrial Chemistry* (2005).
104. James, D. H. & Castor, W. M. Styrene. *Ullmann's Encyclopedia of Industrial Chemistry* (1994).
105. Fishbein, L. Toluene: uses, occurrence and exposure. *IARC scientific publications*, 97–108 (1987).
106. Gloor, G. B., Wu, J. R., Pawlowsky-Glahn, V. & Egozcue, J. J. It's all relative: analyzing microbiome data as compositions. *Annals of epidemiology* **26**, 322–329, https://doi.org/10.1016/j.annepidem.2016.03.003 (2016).
107. Van den Boogaart, K. G. & Tolosana-Delgado, R. *Analyzing compositional data with R*. Vol. 122 (Springer, 2013).
108. Chen, W., Liu, F., Ling, Z., Tong, X. & Xiang, C. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PloS one* **7**, e39743 (2012).
109. Sobhani, I. *et al*. Microbial dysbiosis in colorectal cancer (CRC) patients. *PloS one* **6**, (e16393 (2011).
110. Gorvitovskaia, A., Huse, S. M. & Holmes, S. P. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. *Microbiome* **4**, 15 (2016).
111. Brennan, C. A. & Garrett, W. S. Gut microbiota, inflammation, and colorectal cancer. *Annual review of microbiology* **70**, 395–411 (2016).
112. Wang, T. *et al*. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *The ISME journal* **6**, 320–329 (2012).
113. Clayton, J. B. *Associations Between Nutrition, Gut Microbial Communities, and Health in NonhumanPrimates* PhD thesis. (University of Minnesota, USA, 2015).

Acknowledgements

Author Contributions

J.C., T.J., H.L., V.M., D.T., M.M., H.C., K.G., conceived the study. J.C., B.T., N.T., N.D., T.N., B.T. and J.S. collected and processed samples. J.C., G.G., B.T., F.C., H.H., P.V., T.W., R.S., D.K. and T.J. analyzed the data. J.C., G.G., F.C., H.H., M.M., D.K. and T.J. wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-29277-x.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.