Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ecology Impacts The Decrease of Spirochaetes and Prevotella In The Fecal Gut Microbiota of Urban Humans

Louise B. Thingholm
Christian Albrechts Universität zu Kiel: Christian-Albrechts-Universitat zu Kiel
Corinna Bang
Christian-Albrechts-Universitat zu Kiel https://orcid.org/0000-0001-6814-6151
Malte C. Rühlemann
Christian-Albrechts-Universität zu Kiel: Christian-Albrechts-Universitat zu Kiel
Annika Starke
Christian-Albrechts-Universitat zu Kiel Medizinische Fakultat
Florian Sicks
Tierpark Berlin-Friedrichsfelde GmbH
Verena Kaspari
Tierparkvereinigung Neumuenster e.V.
Anabell Jandowsky
Tierpark Arche Warder e.V.
Kai Fröhlich
Tierpark Arche Warder e.V.
Gabriele Ismer
Tierpark Gettorf GmbH & Co. KG
Andreas Bernhard
Zoo Leipzig GmbH
Claudia Bombis
Tierpark Hagenbeck Gemeinnütige Gesellschaft mbH
Barbara Struve
Leintalzoo Schwaigern
Philipp Rausch
Christian-Albrechts-Universität zu Kiel: Christian-Albrechts-Universitat zu Kiel
Andre Franke (a.franke@ikmb.uni-kiel.de)
Christian Albrechts Universität zu Kiel: Christian-Albrechts-Universitat zu Kiel https://orcid.org/0000-0003-1530-5811

Research

Keywords: mammals, phylogeny, ecology, gut microbiota, human health, physiology

Posted Date: January 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-147555/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: Compared to the huge microbial diversity in most mammals, human gut microbiomes have lost diversity while becoming specialized for animal-based diets – especially compared to chimps, their genetically closest ancestors. The lowered microbial diversity within the gut of westernized populations has also been associated with different kinds of chronic inflammatory diseases in humans. To further deepen our knowledge on phylogenetic and ecologic impacts on human health and fitness, we established the herein presented biobank as well as its comprehensive microbiota analysis. In total, 424 stool samples from 40 different animal species, including *Homo sapiens*, belonging to four diverse mammalian orders were collected at seven different locations and analyzed by 16S rRNA gene amplicon sequencing. Comprehensive data analysis was performed to (i) determine the overall impact of host phylogeny vs. diet, location, and ecology and to (ii) examine the general pattern of fecal bacterial diversity across captive mammals and humans.

Results: By using a controlled study design with captive mammals we could verify that host phylogeny is the most dominant driver of mammalian gut microbiota composition. However, the effect of ecology appears to be able to overcome host phylogeny and should therefore be studied in more detail in future studies. Most importantly, our study could observe a remarkable decrease of Spirochaetes and *Prevotella* in urban humans and New World monkeys, which is probably not only due to diet, but also to the social behavior of these communities.

Conclusion: Our study highlights the importance of phylogenetic relationship and ecology within the evolution of mammals' fecal microbiota composition. Particularly, the observed decrease of Spirochaetes and *Prevotella* in urban communities might be associated to lifestyle dependent rapid evolutionary changes, potentially involved in the establishment of dysbiotic microbiomes promoting the etiology of chronic diseases.

Background

The fecal gut microbiota is well-recognized as an integral part of the mammalian gastrointestinal tract that is crucially involved in host physiology (reviewed in [1-3]). Particularly, the overall diversity of microorganisms and their provided functions play a key role in host nutrition, *e.g.*, via energy production from otherwise indigestible foods [3, 4]. Further, by maintaining a close molecular crosstalk with the innate and adaptive immune system of its host, the mutualistic and commensal gut microbiota is crucially involved in the development and functionality of the mammalian immune response [5-7]. Consequently, changes in community composition, have been shown to contribute to a variety of systemic diseases in humans, including inflammatory bowel disease (IBD), allergies, diabetes and asthma [5, 8-10]. However, different environmental factors are known to substantially influence the gut microbiota, particularly host diet, but also the hosts' evolutionary history [11]. With respect to this, it is assumed that microbial communities might have influenced the evolution of multi-cellular organisms and to this end the evolution of humans [12]. Hence, one current research focus regarding microbial composition in health and disease is to understand the fundamental mechanisms shaping the gut microbiota of mammals. This knowledge can provide deep insights into the short- and long-term adaptation of bacterial communities to their respective hosts, as well as the potential role of maladaption in disease pathogenesis.

Both, diversity and function of the mammalian fecal microbiota, have been characterized in several studies [11, 13-16]. Despite distinctions, the major bacterial phyla (Firmicutes and Bacteroidetes) and sometimes even genera have been shown to be highly conserved during the mammalian evolution [11, 12, 16, 17]. Additionally, key signature taxa compositions known as enterotype-like clusters have been found in various mammalian species [18-20]. Thus, in general, commonalities among mammalian species and their evolutionary history appears to be reflected in the composition and common dynamics of their respective microbial communities. However, particularly for humans we currently lack understanding of the immense and recent changes in humans' gut microbiota due to cultural evolution and urbanization and this influences human health. Particularly in urban populations, a tremendous decrease of fecal bacterial diversity has been shown when compared to rural individuals; most probably arguing for a high and fast impact of lifestyle and diet [15, 21, 22]. With respect to this, it has been hypothesized that diet might mainly mediate the functional community assembly through environmental filtering, while host evolutionary history dictates the prevalence of specific heritable microbial taxa [11]. Nevertheless, the overall understanding of influencing factors resulting in differences between species and individuals, such as genetics and environmental factors like geography, nutrition, and ecology is far from complete [11, 13-15, 23]. This might not only be due to technical artifacts throughout published studies (captive vs. wild animals, mammals vs. non-mammals, different laboratory workflows, distinct analytical approaches), but also the inability to disentangle the effect of correlated patterns such as the dietary behavior of animals, which largely follows their phylogeny. This dependence makes partitioning the microbial variation between the influences of diet and host phylogeny even more difficult and calls for diverse data sets and comprehensive analytical strategies considering intrinsic and extrinsic confounders.

In this study we sampled more than 400 stool samples from 40 different mammal species including *Homo sapiens*, belonging to four diverse mammalian orders (**Table 1**). These samples included species from eight different locations across Germany sampled over four years to analyze the effect of location and habitat as well as host phylogeny and ecology. All samples underwent the same technical procedure (DNA extraction and amplicon sequencing) to ensure compatibility as this has been a notable confounder in earlier studies [24]. The obtained microbial profiles were used to decipher similarities as well as differences in the fecal microbial composition between the various host species with a specific focus on the Hominidae family. Here, we also included a dataset comprising fecal bacterial microbiota results of a large children cohort from Guinea-Bissau, Western Africa [25] to study rapid lifestyle dependent changes, which underlines the loss of bacterial diversity in westernized populations and provide interesting new findings for future research.

Methods

Cohort sampling

Stool samples from animals were collected in seven different zoos within Germany (Arche Warder, Gettorf and Neumuenster from Schleswig-Holstein, Berlin, Hagenbeck in Hamburg, Leipzig in Sachsen and Schwaigern in Baden-Württemberg) by the responsible zookeepers and kept frozen at -80 °C until shipment to the laboratory. Zookeepers ensured sample origin and provided dietary and relationship information on the animals included in this study.

To be able to find potential relationships between dietary factors, phylogenetics as well as dysbiosis associated to the western-lifestyle, 44 healthy human samples were included in the analysis (in addition to four zookeepers from Gettorf). These were recruited at the University Hospital Schleswig Holstein, Campus Kiel and comprised detailed phenotypic, disease related and dietary information. The study was approved by the local ethic committee in Kiel (D441). None of the participants had received any antibiotics or other medication two months prior to inclusion, and none reported any gastrointestinal complaints.

Stool sample processing and sequencing

DNA of samples was extracted using the QIAamp DNA fast stool mini kit automated on the QIAcube (Qiagen, Hilden, Germany). Therefore, material was transferred to 0.70 mm Garnet Bead tubes (Dianova, Hamburg, Germany) filled with 1.1 ml InhibitEx lysis buffer. Bead beating was performed using a SpeedMill PLUS (Analytik Jena, Jena, Germany) for 45 s at 50 Hz. Samples were then heated to 95 °C for 5 min with subsequent continuation of the manufacturer's protocol. Extracted DNA was stored at -20 °C prior to PCR amplification. Blank extraction controls were included during extraction of samples.

For sequencing, variable regions V1 and V2 of the 16S rRNA gene within the DNA samples were amplified using the primer pair 27F-338R in a dual-barcoding approach according to Caporaso et al. [44]. Stool DNA was diluted 1:10 prior PCR, and 3 µl of this dilution were finally used for amplification. PCR-products were verified using the electrophoresis in agarose gel. PCR products were normalized using the SequalPrep Normalization Plate Kit (Thermo Fischer Scientific, Waltham, MA, USA), pooled equimolarily and sequenced on the Illumina MiSeq v3 2x300bp (Illumina Inc., San Diego, CA, USA). Demultiplexing after sequencing was based on 0 mismatches in the barcode sequences.

Data processing

Data processing was performed using the DADA2 [45] workflow for big datasets (https://benjjneb.github.io/dada2/bigdata.html) resulting in abundance tables of amplicon sequence variants (ASVs). Briefly, all sequencing runs were handled separately (workflow adjusted for V1-V2 region can be found here: https://github.com/mruehlemann/ikmb_amplicon_processing/blob/master/dada2_16S_workflow.R) and finally collected in a single abundance table per dataset, which underwent chimera filtering. ASVs underwent taxonomic annotation using the Bayesian classifier provided in DADA2 and using the Ribosomal Database Project (RDP) version 16 release. Sequences that were not assignable to genus level were binned into the finest-possible taxonomic classification. Prior to further analysis, two samples were removed as they came from animals with only one representative species per family, and six samples were removed as they belonged to the outlying sparsely represented order Diprotodonts, which incorporates marsupial mammals only and whose occurrence is restricted to Australasia. ASV abundance tables and taxonomic annotation were passed on to the phyloseq package for random subsampling to 5,900 sequences per sample (rarefy_even_depth(), removing 6998 OTUs) or calculation of relative abundance (normalized by rowsum, minimum read depth 5900) and construction of phylum- to genus-level abundance tables (tax_glom(x, NArm=F)).

The final dataset comprised 417 samples, from 368 unique hosts, unknown host or "collective faecal samples", including 48 humans. A number of samples were collected from "collective faecal samples" in cages and the specific host species animal is therefore not known. For the remaining samples, the specific host animal was recorded by name or no specification was given. In analysis that include one sample per host, we removed all samples but one from named animals, and kept samples from unnamed or "collective faecal samples".

An additional dataset is included, taken from a large cohort of 1255 subjects with fecal bacterial microbiota data, collected in one of the poorest countries in the world, Guinea-Bissau, Western Africa, as described previously [24]. Data was processed with DADA2 and turned into a phyloseq object, after which a subset of 159 subjects were selected as the subset recruited at home (controls) and min. 10 years old. This dataset was then joined with the above described data to evaluate the associations of resent lifestyle vs phylogeny.

Data analysis

A phylogenetic "timetree" was built for mammal species using the website http://timetree.org/ and the resulting tree was imported to R as a phylo object and used throughout the analysis by sub-setting to species of interest for each specific analysis. A list of 42 mammal species was uploaded. As the database would not identify *Macropus rufogriseus* this species was substituted with the close relative *Macropus parryi* in order to keep the two species nodes of the Macropus genus.

Specificity of individual microbial taxa to subgroups of hosts was evaluated using the indicspecies::multipatt function in R with settings func = "IndVal.g" and control = how(nperm=10,000) for analysis within Hominidae and control = how(blocks = (factor(location)), nperm = 10,000) for analysis across host order clades. The function calculates if a taxon show specificity to any one or a combination of subgroups as defined by a cluster argument. For the identification of microbial species showing specificity to mammal order clade(s) or human subgroup (Kiel-area, Guinea-Bissau or zookeeper), species were first filtered to keep those present in at least 0.5% of samples, and only mammal subgroups with min. five subjects were included in the analysis. Tables of species abundance and prevalence across host groups was also generated using the multipatt function and used for visualization in **Figure 4**.

Alpha diversity was calculated from rarefied count data using diversity() and estimateR() from vegan R package. Differences in alpha diversity between groups was analyzed using ANOVA (stats::aov) while correcting for zoo location.

For dispersion analyses of mammals grouped by dietary behaviour (Carnivore, Herbivore, Omnivore) we performed dispersion analysis using vegan::betadisper with bias.adjust=T. To evaluate if the results were robust to variation in the number of different host species clades found in each diet group, we performed random selection of five host species for the two diet-groups herbivores and omnivores, and performed the dispersion analysis. We did so 100 times and calculated the mean dispersion and p-value. The Carnivora were not included due to very few different species in this group. The analysis was performed for microbial taxa at the ASV, species, genus and family level, filtered to keep taxa present in at least 1% of samples. For dataset including the Guinea-Bissau samples, the threshold was set to 0.5%. Permutational manova analysis was performed using vegan::adonis2, with species level data filtered as for betadisper. Animals clades were included if they contained min five samples, and analysis were run with 999 permutations and method="bray". Analysis of phylogeny was correcting for zoo, and vice versa.

The multiple regression on matrices (MRMs), as described previously [26], was used to calculate how much of the microbial variation could be assigned to host phylogeny and location (ecodist ::MRM with 1000 permutations). For each MRM analysis, one samples was selected per host species, and this was repeated 100 times. Then the median coefficient and p-value was calculated. The three-distance matrix was based on geographic coordinates for location (distGeo() function in geosphere), patristic distances for host phylogeny obtained from the website http://timetree.org/ (see above) and the species table for the microbiota filtered to keep species present in >1% of host animals (602 species). Bray Curtis dissimilarity and Jaccard distance was calculated using vegan::vegdist() function in R. Details on diet was available for 115 hosts comprising 15 species. The data was used to generate eight binary dietary variables that reflect the main dietary categories that could be extracted from the available information on the animal's diets. The eight binary variables included in the analysis (fruit, meat, vegetables, eggs, greenery, Herbs/tea, multivitamin/mineral and vitamin B) were used to generate a distance matrix based on the binary Jaccard distance. For the PGLS analysis the species microbiome profiles were used, with the host phylogenetic timetree expanded to include all samples as for MRM, and analysis performed using the caper:: comparative.data and caper::pgls functions with lambda='ML'. For the MPD analysis, data was prepared as for PGLS, but with the extra step of regressing out the effect of location using a linear regression with square root transformation of taxa abundance and selecting species present in >5 samples. PGLS analysis was performed with picante::ses.mpd with null.model = "richness", abundance.weighted = T and runs = 999.

Then, we looked for zoo-specific microbial species. The microbiome species were filtered to keep the 348 most abundant microbial species (min abundance 1e-5 in at least 3% of the 368 samples), and data converted to presence/absence. Using this data, we identified species only present in one location.

.

Illustrations

To generate stacked barplots, taxa with relative abundance not more than 0.05 in at least 10% of the samples were removed, remaining taxa renormalized to sum to 100 per sample, and plots made using function barplot in R package graphics. The plot.phylo() function in R package ape v5.3 was used to generate the phylogenetic tree in Figure 1. Boxplots were generated using ggplot. Figure panels were arranged using function ggpubr::ggarrange, or cowplot::plot_grid, and colors selected from RColorBrewer. Ordination plots in Figure 2 were made using vegan:: capscale with microbiome data aggregated at species level and filtered to species present in at least 1% of samples, dist=bray and metaMDS=F, and ggplot for illustration, while microbial species for Figure 5 were filtered to species present in at least 0.5% of samples. The heatmaps in Figure 4 were made using data extracted from the multipatt analysis, and pheatmap:: pheatmap.

Results

Study cohort

In total, 424 stool samples from 40 different animal species including *Homo sapiens*, belonging to four diverse mammalian orders were collected and microbiome profiled (**Table 1**). After filtering (see Methods) the dataset included microbiota profiles from 368 unique subjects across 38 mammalian species, including four zookeepers, 44 non-zoo-keeping humans and 324 zoo animals. These animals were sampled in seven different zoos across Germany (Berlin, Neumuenster, Gettorf, Warder, Hamburg, Leipzig, Schwaigern) to analyze the effect of location and habitat as well as host phylogeny and ecology. Mammalian orders in the cohort comprise Artiodactyla, Carnivora, Perissodactyla and Primates. Artiodactyla (also called even-toed

ungulates) encompass most of the world's species of large land mammals such as sheep, goats, camels, pigs, cows, and deer, from which ten species are included in this study. Six different species were included from the order Perissodactyla (also called odd-toed ungulates), which in general consists of about 17 species that are hoofed animals (*e.g.,* horses and rhinoceroses). The order Carnivora comprises over 280 species of placental mammals, from which five were sampled in this study. Additionally, 23 different species from the order Primates (old/new world) were analyzed, including the most closely human related primate species, *Pan paniscus* and *Pan troglodytes*. We would like to emphasize that all stool samples, as well as the corresponding extracted-DNAs are publicly available at the Institutes biobank for non-profit research purposes upon request.

Fecal microbiota composition highly reflects animals' phylogeny and thus, also their diet

In this study, the V1V2 region of the 16S rRNA gene from feces of 417 samples (post quality control) were sequenced and ASV tables were generated and annotated to species level. This resulted in 63,780 unique ASVs and 1381 species across the 417 samples. One sample per animal (or stool pool, see methods) was selected as some individual animals were sampled more than once, resulting in 368 samples from 38 species. The general prevalence of microbial species was restricted since 80.5% of microbial species were detected in less than 5% of the samples, likely reflecting the high diversity and distinctness of the various host species and their respective ecologies in the dataset [11].

As expected, the fecal microbiota profiles of the studied mammalian species followed their phylogenetic relationships regarding prevalence and abundance of microbial families **(Fig. 1)**. The stacked bar plot in Fig. 1 illustrates the average microbiota profile of each selected host species in the study and highlights how substantially the Carnivora microbiota differ from the remaining host clades. Another interesting finding from this evaluation was the restricted abundance of *Bacteroidaceae*, which only showed high abundances in *Homo sapiens*, *Callithrix jacchus*, *Varecia rubra* and *Suricata suricatta*. Relatively high proportions of unclassified bacterial families were found in the orders Artiodactyla, Diprotodontia and Perissodactyla, underlining the understudied fecal microbial diversity within these orders. Evaluating the relative frequency of species in each of the five mammalian order clades, we found support for Carnivora as the limiting group to define a mammalian core microbiota. Out of 603 microbial species, five were found with a relative frequency above 50% in all groups, one belonging to each of *Bacteroidaceae*, *Ruminococcaceae, Erysipelotrichaceae, Lachnospiraceae* and unclassified *Candidatus Saccharibacteria*. None of the five ASVs were annotated at species level. When disregarding the Carnivora clade, 27 species met the threshold, while when Artiodactyla, Diprotodontia, Perissodactyla or Primates are left out, only 5, 6, 7 or 6 species met the threshold, respectively.

To evaluate if microbial species showed specificity to any single or combination of host clades, we applied a multi-level pattern analysis (multipatt in R package indicspecies). We found 305 species with significant specificity (out of 602 tested) to any one combination of host order clades, with 82 assigned to Artiodactyla, 156 to Carnivora, 114 to Perissodactyla and 83 to Primates (multipatt, p.adj<0.05, **Table S1**). Interestingly, a large precentage of species associated with Carnivora were uniquely associated with this clade (139 of 156) and therefore show indicator tendency for this order and include two species of *Escherichia/Shigella* genus, and 21 species of the order Clostridiales. For primates, only 24 of the 83 species were specific to the primate order.

We then evaluated the phylogenetic relatedness of each species in the microbiome community and found a broad association of species abundance with host phylogeny. We calculated the mean phylogenetic distance (MPD) between all species and compared observed phylogenetic relatedness to the pattern expected under the null community randomized while holding species richness constant, as we observed some association of host phylogeny with microbiome alpha diversity diversity (picante::ses.mpd). We selected species present in >5 samples and regressed out the effect of location (see methods). A total of 231 species (out of 536) showed phylogenetic relatedness (p.adj<0.05, abundance weighted MPD model, 999 permutations, **Table S2**). Two *Prevotella* genera, namely *Prevotella copri* and one unclassified at species level, was tested in the model and both showed significant phylogenetic relatedness (p.adj<0.05). Five species of Spirochaetes were analyzed, with three Spirochaetales showing significant phylogenetic relatedness (all unclassified at species level, two annotated as *Sphaerochaeta* and *Treponema* at genus level). The remaining two Spirochaetes showed nominal association but did not passmultiple testing correction (p.adj of 0.026 and 0.047).

Dietary preferences are one factor that strongly varies between mammalian orders and diet is believed to be one of the most important factors influencing the gut microbiome, in addition to host genetics, ecology, or habitat. However, as (i) host-phylogenetic clusters have strongly correlated dietary behaviors which largely follow host phylogeny, (ii) our dietary data is restricted to main food categories as binary data *e.g.* overall intake of meat and/or plant-based diet, and (iii) as dietary data is only available for approximately one third of the cohort (incl. only two Carnivora), consistent and reliable segregation of the microbial variation between diet and host phylogeny is not fully feasible using this dataset (strongly nested). Still, with a large overlap in diets for different species and some within-species differences due to between-zoo differences, we found it of relevance to consider the dietary patterns as much as the data allowed.

First, we evaluated if dietary information could explain parts of the variation in the gut microbiota community composition of the mammals. To this end, we used the multiple regression on matrices (MRMs) model, as described in more detail below. Details on diet were available for 115 hosts belonging to 15 species. The data was used to generate eight dietary variables that reflect the main dietary categories from the animal diets *e.g.* fruit, meat, vegetables, greenery (see Methods), and a distance matrix was calculated based on shared dietary patterns (Jaccard distance). We calculated the variation explained by phylogeny, location and diet using the 115 samples with available dietary data. In this subset of samples, both with and

without the inclusion of dietary data, the effect of host phylogeny was significant (median p<0.05). Diet was not significant (median p>0.05) despite 12% variation explained, probably due to the high correlation between diet and host phylogeny in general. Above we observed an association of *Bacteroidaceae* with host phylogeny and identified two *Prevotella* species showing phylogenetic association with the host (namely *Prevotella copri* and one unclassified at species level). Abundance patterns of *Prevotella* has previously been found to be associated with diet with positive associations with fiber intake and negative associations with meat intake. Therefore, we zoomed in on these two species and evaluated the role of vegetables and meat on their abundance while considering location and host phylogeny via Phylogenetic Generalized Least Squares models (PGLS). The unclassified *Prevotella* significantly associated with meat intake (and host phylogeny and location) (PGLS, meat p<0.05, lambda=0.34 (95% CI 0.16-0.61)), while *Prevotella copri* only associated significantly with host phylogeny (lambda=0.96 (95% CI 0.90-0.98)). At the family level, we evaluated association of *Bacteroidaceae* with meat intake and identified both meat and host phylogeny as significantly factors influencing its abundance across hosts (PGLS p<0.05, lambda=0.49 (95% CI 0.26-0.74)).

To further understand the influence of dietary preferences on microbiome diversity, we compared the alpha diversity between hosts grouped into their five orders. Shannon diversity varied between all pairs of host orders that were not both predominantly herbivores (two-way anova correcting for location, q<0.001), while pairs comprising predominantly herbivores showed no significant difference (q>0.05) (no difference between Perissodactyla and Artiodactyls) (**Fig. 2**). Carnivora (carnivores and omnivores) and Primates (predominantly omnivores) had on average approximately half the community diversity of herbivores (Artiodactylas and Perissodactylas), probably highlighting the diversity increasing effect of hgher plant and fiber intake, which requires a rich enzyme repertoire. This observation of a dietary-driven alpha-diversity pattern was further supported by a comparison of Shannon diversity between mammals grouped by their dietary behaviour (as oppose to phylogenetic order) (**Fig 2B**). The observation of a pattern of alpha diversity that follows both the dietary behaviour and host phylogeny of the mammals, was further supported by a Phylogenetic Generalized Least Squares (PGLS) analysis of the association between alpha diversity and dietary behaviour, that becomes insignificant when considering host phylogeny (pgls in R package caper, lambda='ML', controlling for location, p>0.05 for both Shannon and Chao). Interestingly, within the Primate's order, the alpha diversity varied greatly. At the genus level, Hylobates, Macaca, Pan and Pongo, showed the highest diversity, while Varecia and Callithrix showed the lowest diversity (considering clades with min 5 individuals, **Figure S1**). A similar pattern was found for richness (Chao), with the lowest diversity found in Carnivora and Primates (**Figure S1 and S2)**. For richness, Perissodactyla showed a high diversity also compared to the other clade of predominantly herbivores (versus Artiodactylas $q=6.2*10^{-5}$). Host phylogeny directly dictates the animal's dietary preferences in part by shaping their digestive abilities such as the ruminant animals specialized stomach that give them the ability to acquire nutrients from plant-based food. The fermenting process is driven by microbial actions, and when comparing the microbial diversity between ruminants and none-ruminants in the dataset, we identified a significantly higher diversity in the ruminant mammals (**Figure 2D**). To evaluate if there was a detectable effect of individual food groups when controlling for phylogenetic relatedness and location, we applied a PGLS model to the 115 samples with available dietary data and evaluated the association of each of the eight food groups with Shannon diversity. The analysis detected a significant association for fruit, eggs and greenery (PGLS, p<0.05) and a trending association for multimineral/vitamin (p=0.05), all models retaining a significant lambda, indicating a role of both host phylogeny and intake of these food groups on microbial diversity.

To evaluate the variability of microbial communities within host groups with different dietary preferences (carnivores n=6, herbivores n=84, omnivores n=278), we performed dispersion analysis based on the Bray-Curtis diversity measure of dissimilarity between host's microbiome compositions (betadisper in R, bias.adjust=T). The analysis was performed for microbial taxa at the ASV, species, genus and family level, and at all levels the analysis detected a significant difference in variability. However, the pattern of variability between the diet groups changed when moving from the ASV to the higher taxonomic levels. At ASV level, the carnivores had the lowest dispersion, while there was no significant difference between the herbivores and omnivores (mean distance 0.59, 0.67, and 0.68, respectively). At species level, the herbivores diversity decreased drastically (mean distance 0.35), while the omnivores also decreased to 0.52 and carnivores remained largely unchanged (0.55). In addition, there was a significant difference between the herbivores and omnivores at species level (median p=0.0017). The carnivores changed from having the lowest dispersion to having the highest, just above the omnivores, and the pattern remained stable at higher microbial taxonomic levels. Even as we adjusted for sample bias, notable variation in the number of different species sampled within each host order clade remained. When calculating the difference between the herbivores and omnivores, we therefore performed 100 random samplings of five host species groups per diet-group, performed the analysis of variation on each subsampling and then calculated the mean dispersion and p-value across the 100 analyses. The observed pattern is likely caused by the lower species assignment rates in the plant-eating groups as compared to carnivores; carnivores had the lowest number of ASVs with 328 ASVs, and omnivores the highest with 6198 ASVs. At microbial species level the herbivores had 396 species (incl. unannotated) down from 3006 ASVs. The relative change in richness was very low for carnivores (1.77 times) as compared to herbivores and omnivores (7.59 and 10.58, respectively). The carnivores showed the highest percentage of annotated ASVs across microbial species, genera and family levels, followed by the omnivores (percentage annotated microbial families: 99.4% for Carnivora, 75.7% for Omnivore and 67.0% for Herbivore).

Host phylogeny remains an important factor in shaping gut microbiome also for captive and geographically separated mammals

Next, we evaluated whether the variation in the gut microbiota of the mammals is mainly explained by location (given by the Zoo's geographical locations and humans home-city) or by their phylogenetic relatedness. We included one sample per individual mammal (as some mammals had been sampled multiple times) and used multiple regression on matrices (MRMs), as described previously [11], to calculate how much of the microbial variation could be assigned to host phylogeny and location. The three-distance matrices were based on geographic coordinates for location, patristic

distances for host phylogeny and the species table for the microbiota (see **Methods**). To control for the effects of intra-species variation, we performed the analysis 100 times, each time with one randomly selected sample per host species. Thirty-eight host species had data points across all three matrices, and data was selected from a total of 386 samples. The analysis was performed considering both relative abundances (Bray-Curtis) and presence/absence (Jaccard) for microbiota composition. In both analyses, host phylogeny explained a significant amount of variation (median p-value<0.05, coefficient ~23% for BC and Jaccard), while the variation explained by location was insignificant (median p-value>0.05, coefficient -0.03% BC, -0.09% Jaccard, **Figure S3**). In contrast alpha diversity shows strong associations to the geographic location and the phylogenetic relationships between the animals (Shannon lambda= 0.77 (95% CI 0.63-0.87, lower and upper p<0.05), location p<0.05, R2=10%; Chao lambda= 0.80 (95% CI 0.68-0.89, lower and upper p<0.05), location p<0.05, R2=17%).

Despite the very limited variation in microbial community composition found to be explained by location when using the full host phylogenetic tree reflecting geological time and location reflecting geographic distance, further evaluation of the host phylogenetic subgroups using ANOVA and PerMANOVA (adonis) approaches did detect some influence of location. These analyses treat location and host phylogeny as categories unorganized by evolutionary distance or morphology (host taxonomy), or geographical distance (here zoo location). For host mammals grouped at genus level, the variation in microbial composition explained by location was 3.36% after adjusting for phylogeny, and variation explained by phylogeny was 34.7% (likewise after adjusting for location). These associations were highly significant (adonis2 p<0.001, species-level microbiome, 999 permutations, min. 5 animals per host group, **Table S3**). Visual evaluation of the community structure by host phylogeny and zoo location supported an effect of both factors (**Figure 3**), however only a small shift could be detected due to location. Whereas the order Carnivora again clusters decidedly different from all others, members of the Artiodactyla and Perissodactyla are more similar to each other even compared to the Primates, which displayed high variation (dispersion) among their microbial communities. Having a closer look into each order, Primates revealed a peculiar pattern in their microbial communities. Here, four human samples were included that did not belong to the human Kiel control group, but instead were sampled from two animal zookeepers from Gettorf, as well as from two workers not handling animals. The two samples of the animal zookeepers shifted, away from the human samples from the geographically close Kiel area, towards Gettorf zoo where they worked and for the zookeeper of lemurs, tamarins and squirrel monkeys, the shift was directed towards the *Saguinus oedipus* (tamarins) (see **Figure 3 "Primates"**), indicating their microbiomes are influenced by the animals they interact with. Otherwise, the clustering of primate species highly reflects their phylogeny, even though many of them live in different group sizes and together with many other species. A similar pattern could be observed for most other host orders, too, including *e.g.,* sheep and goats within in the Artiodactyla or the racoons within the Carnivora.

When considering the within-sample diversity (alpha diversity) using PGLS analysis, as opposed to the community composition (beta-diversity) evaluated above, support was detected for an effect of both host phylogeny and location, supporting the above observations for community structure (Shannon lambda= 0.77 (95% CI 0.63-0.87, lower and upper p<0.05), location p<0.05, R^2=10%; Chao lambda= 0.80 (95% CI 0.68-0.89, lower and upper p<0.05), location p<0.05, R2=17%).

One way by which the confinement of animals to specific zoos could influence a possible phylogeny-driven microbial composition could be through local community dynamics and restricted bacterial/host dispersal between locations, eventually leading to zoo specific microbial communities/signatures. Thus, we looked for zoo-specific microbial species within the 351 most abundant microbial species (min abundance 0.001% in at least 3% of samples). Only four species (*Alistipes finegoldii*, *Bacteroides stercoris*, *Bifidobacterium tissieri* and *Clostridium IV leptum*) were unique to one location, namely *A. finegoldii* and *C. IV leptum* to Kiel, and *B. stercoris* and *B. tissieri* to Neumuenster. The bacteria specific to Kiel originate completely from human hosts while bacteria specific to Neumuenster Zoo originate from Primates and Carnivora. In Neumuenster, marmosets (Primates) hosted by far the majority of the two species (found in 17 or the 19 marmosets' stool-pool samples), while samples from marmosets were also only available for this location. Both *B. stercoris* and *B. tissieri* were also found in one stool-pool from ring-tailed coati, and *B. tissieri* was found in one stool-pool of racoons. Both ring-tailed coatis and racoons were also sampled in other zoos (Gettorf and Berlin, respectively), indicating some cross-host species transfer within zoos or indicates similarities in the host's ecologies. However, the overall pattern does not indicate widespread zoo-specific microbial species. A PGLS analysis of *B. stercoris* and *B. tissieri* with location, confirmed the importance of host phylogeny over location (caper::PGLS, lambda='ML', *B. stercoris* lambda=0.75 (95% CI 0.61-0.85, p upper and lower <0.05), location p=0.85; *B. tissieri* lambda=0.67 (95% CI 0.50-0.80, p upper and lower <0.05), location p=0.74). PGLS cannot be used to evaluate the two taxa unique to Kiel due to Kiel location only containing samples from one host species which contributes all of these specific species (44 of 48).

Variation in the family Hominidae

A reduced diversity and an increase in dispersion have been observed for humans as compared to closely related taxa or other mammals. Across mammalian families, Hominidae showed a highly variable diversity **(Figure S1)**. The range overlapped with most other families but was clearly lower than most herbivorous families, and showed generally higher Shannon diversity than the Carnivora, and the primate families Callitrichidae (marmoset), Cebidae and Lemuridae. Our dataset includes three genera (comprising four species) within the family Hominidae. Since microbial diversity is also known to be decreased within westernized populations [15, 21], we here included another dataset comprising fecal bacterial microbiota results of a large children cohort from one of the poorest countries in the world, Guinea-Bissau (Western Africa) [25, 27]. We selected 159 individuals who were recruited at home (controls) at a minimum of 10 years of age and compared their alpha diversities with the human subjects of our study cohort. With regards to Shannon diversity, Guinea-Bissau human subjects had a significantly higher diversity compared to the German subjects ($q=3.1*10^{-12}$) but lower diversity when compared to the hominid primate genera *Pan* ($q=1.4*10^{-16}$) and *Pongo* (q=0.048) (**Figure S4**). As expected, human subjects from

Germany had an even lower alpha diversity when compared to *Pan* ($q=2.5*10^{-28}$) and *Pongo* ($q=7.4*10^{-6}$), whereas *Pan* and *Pongo* did not show differences to each other (q=0.19). PGLS analysis confirmed the importance of phylogeny in shaping diversity of these four hominid species clades (PGLS with location, lambda='ML', including only German humans, Shannon lambda=0.73 (95% CI 0.33-0.95, p<0.05), Chao lambda=0.83 (95% CI 0.51-0.97, p<0.05).

To further evaluate the microbial community (species) within the Hominidae, we evaluated the dispersion of samples in each subgroup of *Pan*, *Pongo* and humans grouped by location. The humans showed a higher dispersion as compared to Pan and Pongo, and the German subjects a higher dispersion compared to subjects from Guinea-Bissau ($p =6.2*10^{-5}$, German mean distance 0.42, Guinea-Bissau mean distance 0.36).

To further explore the variation in the Hominidae, we used the indicator species analysis (multipatt introduced above) to identify bacterial species that are specific to a certain host group (*Pan*, *Pongo*, and German and Guinea-Bissau human subjects). The analysis identified 141 species (out of 353 analyzed) with significant specificity to one or a combination of groups, with 2 assigned to specifically to humans from Guinea-Bissau (out of 38 assigned to a combination that include Guinea-Bissau), 35 specifically to humans from Germany (out of 71), 8 specifically to *Pan* (out of 59) and 20 specifically to *Pongo* (out of 76) (multipatt p.adj<0.05). **Figure 4** shows the relative abundance and frequency of the 53 most significant species (p.adj<0.001), the full list is presented in **Table S4**. Of the 35 species showing specificity to the German humans, out of the 71 species associated with Germans, eight belonged to the *Bacteroides* clade and five to *Alistipes*. Prior studies comparing mammal clades and westernized with non-westernized populations reported that Spirochaetes are increasingly absent from populations consuming a westernized diet [14, 21]. The current dataset includes six species assigned to the order Spirochaetes, of which four showed significant association to non-human subgroups, while two are highly abundant and prevalent in subjects from Guinea-Bissau (potentially *Brachyspira pilosicoli*, *B. aalborgi*). Also, *Prevotella* shows an association towards non-westernized subjects (*Prevotella copri* and one unclassified *Prevotella*, multipatt p.adj<0.05). Contrary to *Prevotella*, *Bacteroides* displayed a high abundance and prevalence only in human subjects from Germany (specificity to German humans, eight of 10 *Bacteroides* with multipatt p.adj<0.05). These findings outline previously reported and yet unknown phenomena along gradients of westernization in humans and the hominids in general, and argue for the diverse interaction between host background, lifestyle, and microbiome as shown in previous studies on humans as well [28].

There exist over 250 different species of monkeys, and these can be divided into two main groups; the Old World monkeys that are native to Africa and Asia, and the New World monkeys that are native to Central and South America. Our dataset also includes different species of New (33) and Old World monkeys (52), with Old World monkeys falling in the Cercopithecidae family and New World monkeys in the Callitrichidae and Cebidae families. Recently, Amato and colleagues found evidence that human stool microbiota composition is more similar to members of the Cercopithecidae than to those of *Pan* and *Pongo* [22]. Thus, we here analyzed compositional differences using data from zoo animals instead of wild ones. Surprisingly, data obtained confirmed the overall findings of Amato *et al.*, though all animals in zoos are in close contact with humans. Comparisons of the microbiota composition (species) of the Cercopithecidae mammals with German humans, Guinea-Bissau humans and *Pan*, using adonis2, showed a smaller difference between German humans and Cercopithecidae (R^2=0.36), as compared to German humans and *Pan* (R^2=0.40). When taking the Guinea-Bissau humans instead of the German humans, the pattern was the same, however both measures of difference was reduced (vs. Cercopithecidae R^2=0.25, vs. Pan R^2=0.30).

Globally over 250 different monkey species exist, which can be divided into two main groups; the Old World monkeys which are native to Africa and Asia, and the New World monkeys which are native to Central and South America. Our dataset also includes various species of New (33) and Old World monkeys (52), with among the families Cercopithecidae (Old World monkeys) and Callitrichidae and Cebidae in the group of New World monkeys. Recently, Amato and colleagues found evidence that human fecal microbiota composition is more similar to members of the Cercopithecidae than to those of hominid apes [23]. Thus, we analyzed compositional differences using data from zoo animals instead of wild ones to reduce the effects of parasitation, environment and diet. Surprisingly, we can confirm the overall findings of Amato *et al.*, even though zoo animals are in close contact with humans. Comparisons of the microbiota composition between members of the Cercopithecidae, *P. pansicus*, *P. troglodytes* and human subjects form Germany and Guinea-Bissau showed a smaller difference between German humans and Cercopithecidae (R2=0.36), as compared to German subjects and members of the genus *Pan* (R2=0.40). Human subjects from Guinea-Bissau displayed a similar trend, although differences were more subtle (vs. Cercopithecidae R2=0.25, vs. Pan R2=0.30).

Discussion

Compared to the vast microbial diversity in most mammals, human gut microbiomes have lost diversity while becoming specialized for animal-based diets – especially compared to their genetically closest ancestors and particularly in westernized societies [13, 18, 29, 30]. Meanwhile numerous studies have shown that low diversity of the gut microbiome associates with different types of chronic diseases in humans [9, 10]. To further expand our understanding of the role of evolutionary history for shaping the gut microbiome and how this might impact today's human health, we created this broad dataset and conducted comprehensive analyses.

Nearly 64,000 ASVs (bacterial) were detected in our dataset with highest taxonomic diversity in the orders Artiodactyla and Perissodactyla. Though a high diversity within Artiodactyla and Perissodactyla has been shown in various studies before [12], we found high proportions of not assignable bacterial families highlighting a yet hidden microbial diversity within these host groups. In contrast to these hosts, we observed a relatively less diverse but highly specific bacterial communities in samples from the order Carnivora, which are mainly composed of so far well characterized bacteria, as was also observed for the order Primates. In general, bacterial community structure largely followed the phylogenetic relationship of the studied

mammals with a common core microbiome of at least 30 bacterial species (relative frequency normalized by group size above 50%) in all groups, except for Carnivora. This finding might reflect the functional assembly of bacteria shared between the omnivorous and herbivorous hosts among the Artiodactyla, Perissodactyla, and Primates as described previously [11, 31]. This core microbiota contains bacterial strains that are essential for the breakdown of plant fibers and production of SCFAs and are thus crucial for host physiology [32]. Due to the limitations of 16S rRNA gene amplicon sequencing we were not able to define specific strains and genes that form part of this core, however ongoing studies in our laboratory using deep shotgun sequencing will uncover those in the future and extend upon previous low coverage analyses.

As summarized by Groussin et al. in 2017 [31], the consistent dominant drivers of animal gut microbiome diversity appear to be host evolutionary history and diet (including physiological adaptations), while also biogeography, sex, reproductive status, and social structure have been associated to microbial community differences. From an analytical point of view, dietary behavior of animals largely follows their phylogeny, thus making the partitioning of microbial variation between diet and host phylogeny difficult in such studies. While similar limitations apply to our dataset, captive animals' dietary sources are more controlled and therefore easier to record. Dietary data was gathered for a subset of the dataset and allowed for an evaluation of the role of both diet and host phylogeny. Interestingly, by using the subset with dietary information our analysis suggests that the effect of host phylogeny is still stronger than that of diet. Previous studies focusing on the strictly herbivorous Panda bears underline this finding by demonstrating that their fecal gut microbiota is more similar to their carnivorous and omnivorous relatives than their dietary behavior suggests [30, 33]. Having samples from seven different locations available for analysis we could also demonstrate that host location only has a minor influence on the bacterial community structure. More precisely, location/geography appears to influence only individual bacterial families in most animals, whereas the broad taxonomic levels are dictated by host phylogeny as has been shown in previous studies [34-36]. We could not observe any bacterial clades that were location-specific and prevalent, most probably due to the central role of host phylogeny for bacterial prevalence patterns as well as the well-defined and shared dietary information throughout German zoos.

During the last years, ecology of mammals and in particular of primates has come into focus of microbiome research which surprisingly shows many parallels to patterns observed in human associated microbial communities and has been shown to even overcome host phylogeny [37-39]. In line with the study of Amato *et al.* 2019, we found that the fecal bacterial community between humans and studied individuals belonging to the New World monkeys is more similar than between humans and *Pan* or *Pongo*, which are phylogenetically closer [13, 23, 40]. As hypothesized by Amato *et al.*, this effect might originate from the human dietary niche and associated physiological adaptations, which are more similar to those of cercopithecines than to other apes [23]. Thus, the findings of Amato *et al.* could be confirmed even in captive animals living in close contact with humans – strongly supporting the influence of phylogeny, ecology, and associated physiological adaptations in shaping the microbial community of primates.

After having assessed the importance of host phylogeny, location, and diet in shaping the gut microbiome in the full dataset, we turned our focus to the family Hominidae to further understand the observed decrease of fecal bacterial diversity in westernized humans during the last 100 years. To further explore this trend, we included not only a human cohort from North Germany, but also a dataset of fecal microbiotas determined from a rural community of children living in Guinea-Bissau, Western Africa. When comparing the fecal bacterial diversity between all host groups of the Hominidae family, we justified previous observations about the westernized (local to Kiel, Germany) humans: this cohort had the lowest diversity as well as the highest specificity and dispersion (reviewed in [30]). In addition, we could confirm earlier findings of reduced relative abundance of *Prevotella* – probably replaced by increased relative abundances of *Bacteroides* – and an increase of several clades that are associated with carnivorous diet in other mammals, *e.g., Enterobacteriaceae* [30, 39]. This trend was also reported in studies comparing urban and rural human cohorts only and hypothesized to be due to consumption of sugar, animal fat, and calorie-dense foods in industrialized countries [14, 15]. Additionally, we observed a decrease of members of the Spirochaetes specifically for the German (urban) human cohort, highlighting a trend of urbanization-associated lose within the human species that was described in other studies as well [14, 21]. This finding, however, could be expanded on with this dataset as we observe high proportions of Spirochaetes in fecal samples from nearly all other mammal orders with the exception of Carnivora and - more noteworthy - primates belonging to the New World monkeys. It has been speculated that the presence of Spirochaetes and *Prevotella* relates to the high fiber intake from ingested plant polysaccharides to produce high levels of short-chain fatty acids and thereby maximizing metabolic energy extraction [14, 21]. However, a study in baboons demonstrated that social interactions are an important determinant of gut microbiota composition as well [37]. Here, Tung and colleagues studied baboons belonging to two different social groups and found that direct physical contact during social interactions plays a role in transmitting gut bacteria between members of the same social group. Particularly, anaerobic and non-spore-forming bacteria such as Spirochaetes and *Prevotella* were, found amongst those "socially structured" microbes [37]. This would indicate that the observed changes in abundances of Spirochaetes and *Prevotella* might also depend on host's ecology. Unlike most Old World monkeys, many New World monkeys form monogamous pair bonds and live only in small groups [41]. Urban human communities also tend to live separated in small groups or families, whereas rural communities are living closely together, sharing smaller rooms and often live in close connection to their livestock. Thus, consistent changes in fecal microbiota composition among phylogenetically very close related mammals highly correlate with their social relationships. The samples from two zookeepers from Gettorf that cluster with the animals they are taking care of highly underlines this assumption though with very limited sample size in our study. However, the same observation has been made before for human family members that share the fecal microbiota with their dogs despite having different diets [42]. However, it shall be noted that Spirochaetes are mainly living in anaerobic sediments [43] to which both – urban human communities (mainly living in cities) as well as New World monkeys (mainly arboreal) – have less contact. These ecological characteristics could of course play an essential role for the loss of these bacteria and thus warrants future studies.

One of the main technical artifacts in studies elucidating mammal fecal microbiota has been noted to be sampling in zoos instead of wild animals [3, 12, 22, 44]. However, this could not be confirmed in our study. Instead, we think that the studied mammals from different locations in Germany were

better comparable due to standardized food and rhythms in their diet. This is underlined by the limited microbial variation found to be explained by location or diet when using the full host phylogenetic tree. In addition, the clustering of primate species highly reflects their phylogeny, though many of them live in different group sizes and with diverse other species within the zoo itself. A similar pattern could be observed for most other host orders, too, including *e.g.,* sheep and goats within the Artiodactyla or the racoons within the Carnivora. Recently, Nishida *et al.* also detected no significant microbial variation between wild and captive animals [24], though of course antibiotic influences have been reported [22] and other studies highlighted large transitions after the transfer of wild animals into laboratory environments [20]. Moreover, we could confirm various earlier observations that rely on samples of wild animals. Thus, we find that it is of much higher importance to employ the same standardized and well-controlled sample processing pipeline throughout a study. In addition, while effects of location and housing conditions are present and an important factor in microbiome studies, the role of host phylogeny remains strong and appears resilient to other confounders.

In conclusion, our study highlights the importance of phylogenetic relationship and ecology within the evolution of mammals' fecal microbiota composition. Particularly, the tremendous decrease of Spirochaetes and *Prevotella* in urban communities might be associated to lifestyle dependent evolutionary fast-track changes, potentially involved in the establishment of dysbiotic microbiomes promoting the etiology of chronic diseases [45]. Consequently, the observed findings urgently need deeper analysis based on shotgun metagenomics and metatranscriptomic studies to gain insights into the functional loss as well as the immunogenic consequences that might be associated.

Declarations

Ethics and consent to participate

The study was conducted in accordance with the Declaration of Helsinki, and ethical approval was granted by the ethics committees at Kiel University (D441, German controls) and at Guinea-Bissau (Comité Nacional de Ética na Saúde) (ref. no. 062/CNES/INASA/2017) as well as of Region of Southern Denmark (ref. no. S-20160138). All volunteers or parents of volunteers provided oral and written informed consent.

No ethical approval for animal faecal samples was obtained because this study did not involve a prospective evaluation, did not involve laboratory animals and only involved non-invasive procedures.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analysed during the current study are available at the NCBI SRA repository (Accession no. PRJNA685981 and PRJNA642721). Codes used for data analyses and supportive data files are available at https://github.com/LouiseBThingholm/Zoo-Microbiome-Project-2020.

Competing interests

The authors declare no competing interests.

Funding

This study was supported by the Deutsche Forschungsgemeinschaft (DFG) Research Training Group 1743 and received infrastructure support from the Collaborative Research Center 1182 ‚Origin and Function of Metaorganisms' (www.metaorganism-research.com, no: SFB1182).

Authors Contribution

This study was designed by LBT, CB, AS and AF. CB, AS, FS, VK, AJ, KF, GI, AB, CBo and BS were responsible for sample collection and handling. Funding was acquired by AF. CB and AS assisted with sample analysis (laboratory work and data handling). Bioinformatics were performed by LBT, MCR and PR. The study was supervised by CB and AF. LBT, CB, MCR, PR and AF wrote the original draft. All authors read and approved the final manuscript prior to submission.

Acknowledgements

We would especially like to thank all animal keepers, who were responsible for stool sample collection and who are taking care on the animals every day, namely these were René Viete und Andrea Fleischer from Berlin, Dr. Silke Plagmann from Gettorf, Frank Maxis from Leipzig, Carolin Reimertz from Warder, and all others involved from Neumeunster, Hamburg and Schwaigern. We would also like to thank Ms. Ilona Urbach, Ms. Ines Wulf and Mr.

Tonio Hauptmann of the IKMB microbiome laboratory and the staff of the IKMB sequencing facilities for excellent technical support. We would also like to thank our collaboration partners Sebastian von Huth and Uffe Holmskov from the Cancer and Inflammation Research Group University of Southern Denmark.

References

1. Ley RE, et al. Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol. 2008;6(10):776–88.
2. McFall-Ngai M, et al. Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci USA. 2013;110(9):3229–36.
3. Clayton JB, et al. The gut microbiome of nonhuman primates: Lessons in ecology and evolution. Am J Primatol. 2018;80(6):e22867.
4. den Besten G, et al. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. J Lipid Res. 2013;54(9):2325–40.
5. Hill DA, Artis D. Intestinal bacteria and the regulation of immune cell homeostasis. Annu Rev Immunol. 2010;28:623–67.
6. Macpherson AJ, Harris NL. Interactions between commensal intestinal bacteria and the immune system. Nature Reviews Immunology. 2004;4(6):478–85.
7. Artis D. Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. Nature Reviews Immunology. 2008;8(6):411–20.
8. Moore WE, Moore LH. Intestinal floras of populations that have a high risk of colon cancer. Appl Environ Microbiol. 1995;61(9):3202–7.
9. Proceedings of the National Academy of Sciences of the United States of America. Amaral FA, et al., Commensal microbiota is fundamental for the development of inflammatory pain. Proc Natl Acad Sci USA, 2008. 105(6): 2193–7.
10. Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. Curr Opin Gastroenterol. 2015;31(1):69–75.
11. Youngblut ND, et al. Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. Nat Commun. 2019;10(1):2200.
12. Colston TJ, Jackson CR. Microbiome evolution along divergent branches of the vertebrate tree of life: what is known and unknown. Mol Ecol. 2016;25(16):3776–800.
13. Amato KR, et al. Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes. ISME J. 2019;13(3):576–87.
14. De Filippo C, et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A. 2010;107(33):14691–6.
15. Yatsunenko T, et al. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222–7.
16. Ley RE, et al. Evolution of mammals and their gut microbes. Science. 2008;320(5883):1647–51.
17. Muegge BD, et al. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science. 2011;332(6032):970–4.
18. Moeller AH, et al. Chimpanzees and humans harbour compositionally similar gut enterotypes. Nat Commun. 2012;3:1179.
19. Moeller AH, Ochman H. Microbiomes are true to type. Proc Natl Acad Sci U S A. 2014;111(26):9372–3.
20. Wang J, et al. Dietary history contributes to enterotype-like clustering and functional metagenomic content in the intestinal microbiome of wild mice. Proc Natl Acad Sci U S A. 2014;111(26):E2703-10.
21. Angelakis E, et al. species enrich the gut microbiota of traditional rural populations but are absent from urban individuals. New Microbes New Infect. 2019;27:14–21.
22. Campbell TP, et al. The microbiome and resistome of chimpanzees, gorillas, and humans across host lifestyle and geography. ISME J. 2020;14(6):1584–99.
23. Amato KR, et al. Convergence of human and Old World monkey gut microbiomes demonstrates the importance of human ecology over phylogeny. Genome Biol. 2019;20(1):201.
24. Nishida AH, Ochman H. Rates of gut microbiome divergence in mammals. Mol Ecol. 2018;27(8):1884–97.
25. von Huth S, Kofoed PE, Holmskov U, *Prevalence and potential risk factors for gastrointestinal parasitic infections in children in urban Bissau, Guinea-Bissau.* Trans R Soc Trop Med Hyg, 2019.
26. Lichstein JW. Multiple regression on distance matrices: a multivariate spatial analysis tool. Plant Ecol. 2007;188(2):117–31.
27. von Huth S, et al. Minor compositional alterations in faecal microbiota after five weeks and five months storage at room temperature on filter papers. Sci Rep. 2019;9(1):19008.
28. Clemente JC, et al., *The microbiome of uncontacted Amerindians.* Sci Adv, 2015. 1(3).
29. Walter J, Ley R. The human gut microbiome: ecology and recent evolutionary changes. Annu Rev Microbiol. 2011;65:411–29.
30. Davenport ER, et al. The human microbiome in evolution. BMC Biol. 2017;15(1):127.
31. Groussin M, et al. Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. Nat Commun. 2017;8:14319.

32. Rowland I, et al. Gut microbiota functions: metabolism of nutrients and other food components. Eur J Nutr. 2018;57(1):1–24.
33. Xue Z, et al. The bamboo-eating giant panda harbors a carnivore-like gut microbiota, with excessive seasonal variations. mBio. 2015;6(3):e00022-15.
34. Couch CE, et al. Bighorn sheep gut microbiomes associate with genetic and spatial structure across a metapopulation. Sci Rep. 2020;10(1):6582.
35. Phillips CD, et al. Microbiome analysis among bats describes influences of host phylogeny, life history, physiology and geography. Mol Ecol. 2012;21(11):2617–27.
36. Sylvain F, et al., *Fish Skin and Gut Microbiomes Show Contrasting Signatures of Host Species and Habitat*. Appl Environ Microbiol, 2020. 86(16).
37. Tung J, et al., *Social networks predict gut microbiome composition in wild baboons*. Elife, 2015. 4.
38. Gilbert SF. A holobiont birth narrative: the epigenetic transmission of the human microbiome. Front Genet. 2014;5:282.
39. Moeller AH, et al. Social behavior shapes the chimpanzee pan-microbiome. Sci Adv. 2016;2(1):e1500997.
40. Amato KR, et al. Phylogenetic and ecological factors impact the gut microbiota of two Neotropical primate species. Oecologia. 2016;180(3):717–33.
41. Rylands AB, Mittermeier RA, *The Diversity of the New World Primates (Platyrrhini): An Annotated Taxonomy*, in *South American Primates: Comparative Perspectives in the Study of Behavior, Ecology, and Conservation*, P.A. Garber, et al., Editors. 2009, Springer New York: New York, NY. p. 23–54.
42. Song SJ, et al. Cohabiting family members share microbiota with one another and with their dogs. Elife. 2013;2:e00458.
43. Harwood CS, Canale-Parola E. Ecology of spirochetes. Annu Rev Microbiol. 1984;38:161–92.
44. Frankel JS, et al. The effect of captivity on the primate gut microbiome varies with host dietary niche. Am J Primatol. 2019;81(12):e23061.
45. Finlay BB, Humans C, Microbiome. Are noncommunicable diseases communicable? Science. 2020;367(6475):250–1.

Table

Table 1. Summary of animals and samples included in the present study. A total of 38 different mammalian species were sampled across seven different locations in Germany. The table summarizes the number of animals per species, the locations where each species was sampled, the phylogeny of the species, their characteristic and dietary behavior.

No. of Individuals	No. of Locations	Species	Species (scientific)	Characteristics	Genus	Family	Diet	Order
6	2 [a,b]	*zebu*	*Bos primigenius f. taurus*	Ruminants Social	Bos	Bovidae	Herbivores, feeding on grass, foliage, and plant products	Artiodactyla
7	3 [a,b,c]	*goat*	*Capra aegagrus f. hircus*	Ruminants Social	Capra			
10	2 [a,b]	*sheep*	*Ovis orientalis f. aries*	Ruminants Social	Ovis			
3	1 [b]	*eland*	*Taurotragus oryx*	NonRuminant Solitary	Taurotragus			
2	1 [b]	*camel*	*Camelus ferus f. bactrianus*	Ruminants Social	Camelus	Camelidae		
8	1 [b]	*vicuna*	*Vicugna vicugna f. pacos*	NonRuminants Social	Vicugna			
2	1 [c]	*elk*	*Alces alces alces*	NonRuminants Solitary	Alces	Cervidae		
5	1 [c]	*reindeer*	*Rangifer tarandus fennicus*	Ruminants Social	Rangifer			
14	1 [a]	*wild boar*	*Sus scrofa*	NonRuminants Social	Sus	Suidae	Omnivores, eating grass, leaves, roots, insects, worms	
23	2 [b,d]	*ring-tailed coati*	*Nasua nasua*	NonRuminants Social	Nasua	Procyonidae	Omnivore	Carnivora
21	2 [c,d]	*racoon*	*Procyon lotor*	NonRuminants Solitary	Procyon			
2	1 [b]	*black bear*	*Ursus americanus*	NonRuminants Solitary	Ursus	Ursidae		
4	2 [c,d]	*polar bear*	*Ursus maritimus*	NonRuminants Solitary			Carnivore	
2	1 [b]	*meerkat*	*Suricata suricatta*	NonRuminants Social	Suricata	Herpestidae		
13	1 [a]	*donkey*	*Equus asinus*	Ruminants Social	Equus	Equidae	Herbivores, feed on grasses, leaves, and other plant parts (hindgut fermenters)	Perissodactyla
11	3 [a,b,c]	*horse*	*Equus ferus caballus*	Ruminants Social				
4	1 [b]	*zebra*	*Equus quagga boehmi*	NonRuminants Social				
2	1 [b]	*tapir*	*Tapirus terrestris*	NonRuminants Solitary	Tapirus	Tapiridae		
19	1 [d]	*marmoset*	*Callithrix jacchu*	NonRuminants Social	Callithrix	Callitrichidae	Omnivores, eating insects, fruit, and the sap or	Primates
1	1 [b]	*emperor*	*Saguinus*	NonRuminants	Saguinus			

		tamarin	*imperator subgrisescens*	Social			gum from trees
3	1 [b]	*white-lipped tamarin*	*Saguinus labiatus*	NonRuminants Social			
2	1 [c]	*red-handed tamarin*	*Saguinus midas*	NonRuminants Social			
3	1 [b]	*cotton-top tamarin*	*Saguinus oedipus*	NonRuminants Social			
5	1 [b]	*squirrel monkey*	*Saimiri sciureus*	NonRuminants Social	Saimiri	Cebidae	
2	1 [b]	*Diana monkey*	*Cercopithecus diana*	NonRuminants Social	Cercopithecus	Cercopithecidae	Omnivores, eating mainly fruits, but also flowers, leaves, bulbs and rhizomes, insects, snails, small mammals
1	1 [b]	*black crested mangabey*	*Lophocebus aterrimus*	NonRuminants Social	Lophocebus		
6	1 [b]	*celebes crested macaque*	*Macaca nigra*	NonRuminants Social	Macaca		
37	3 [b,c,d]	*barbary macaque*	*Macaca sylvanus*	NonRuminants Social			
2	1 [e]	*drill*	*Mandrillus leucophaeus*	NonRuminants Social	Mandrillus		
1	1 [f]	*mandrill*	*Mandrillus sphinx*	NonRuminants Social			
3	1 [e]	*hamadryas baboon*	*Papio hamadryas*	NonRuminants Social	Papio		
48	1 [g, b]	*human*	*Homo sapiens*	NonRuminants Social	Homo	Hominidae	Omnivores, with fruit as the preferred food among all but some human groups
1	1 [h]	*bonobo*	*Pan paniscus*	NonRuminants Social	Pan		
56	2 [b, f]	*chimp*	*Pan troglodytes*	NonRuminants Social			
6	1 [e]	*Sumatran orangutan*	*Pongo abelii*	NonRuminants Solitary	Pongo		
12	2 [b, f]	*white-handed gibbon*	*Hylobates lar*	NonRuminants Social	Hylobates	Hylobatidae	Omnivores, eating mainly fruits, but also flowers, leaves and insects
10	2 [b, c]	*ring-tailed lemur*	*Lemur catta*	NonRuminants Social	Lemur	Lemuridae	Omnivores
11	2 [b, c]	*red ruffed lemur*	*Varecia rubra*	NonRuminants Social	Varecia		Herbivorous, eating mainly fruits and leaves

[a]Arche Warder, [b]Gettorf, [c]Berlin, [d]Neumuenster, [e]Hagenbeck, [f]Schwaigern, [g]Kiel, [h]Leipzig

Figures

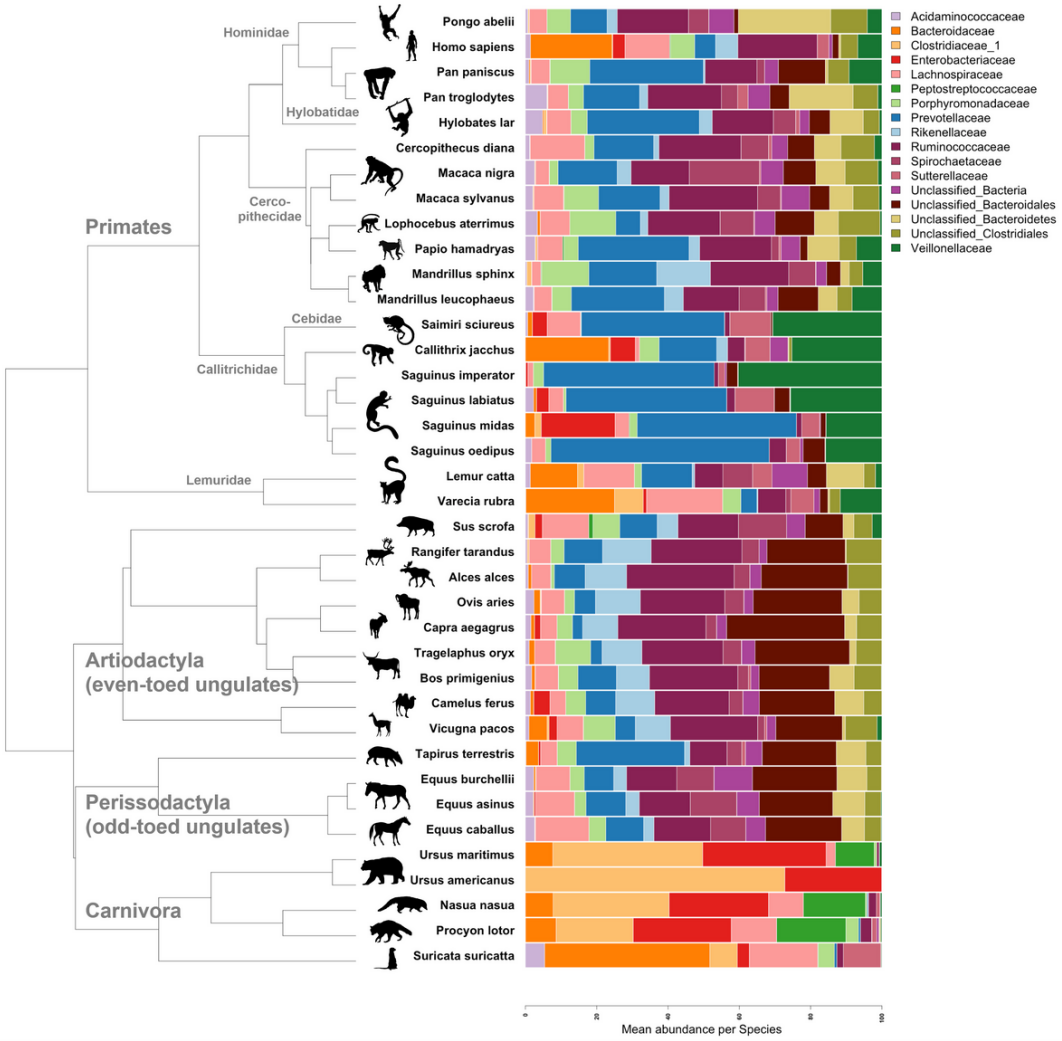

Figure 1

Mammals gut bacterial profile by host phylogeny. Mapping of family-level microbiome relative abundances (17 most abundant) onto host phylogenetic tree (built using http://timetree.org/) revealed clear clustering of microbiome profile by host clade. A total of 38 different host species with microbiome data are included, and microbiome data aggregated at family level keeping unannotated clades (seen as unclassified in legend). Icons taken from http://phylopic.org/. Credits to Rebecca Groom, Roberto Díaz Sibaja, Sarah Werning.

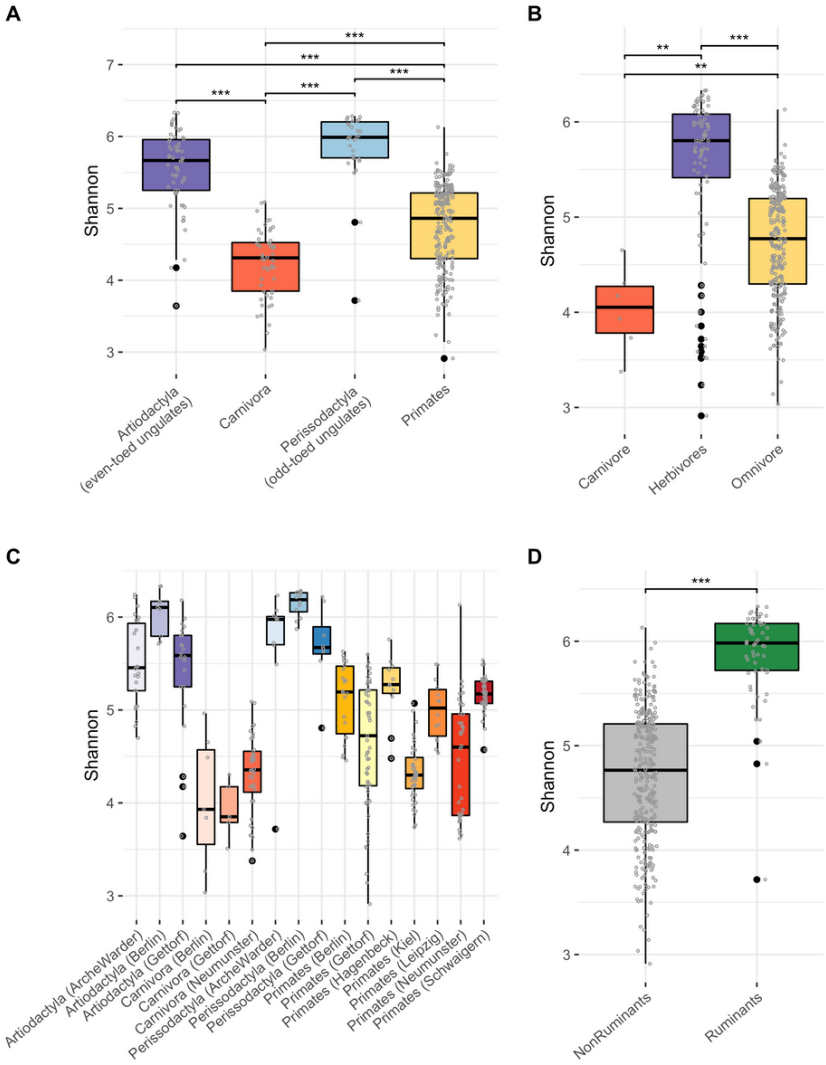

Figure 2

Comparison of alpha diversity between host clades (A) at order-level, (B) by dietary behavior, (C) order level sub-stratified by sampling location and (D) ruminant phenotype. Alpha diversity measured as Shannon diversity, differed between host order clades in a manner that largely reflected dietary preferences but with little association to location. Analysis of pairwise differences was performed using ANOVA, correcting for location. ***: p.adj < 0.001, **: p.adj < 0.01.

Figure 3

Graphical summary of the community structure by host phylogeny and zoo location. Panel (A) show host animals colored by phylogenetic order and shaped by location (Zoo or hometown) indicating only small effects of location. Panels (B-E) show animals belonging to each of the four different phylogenetic orders samples in the cohort. Panel (B) include enhanced circles and labels for the two animal zookeepers included in the dataset. The clustering of species in those orders highly reflects their phylogeny, though many of them live in different group sizes and with diverse other species within the zoos itself. Each panel shows host animals coloured by phylogenetic species and shaped by location (Zoo or hometown). Plots are unconstrained principal coordinates analysis made with vegan::capscale, with dist="bray", metaMDS = F. Percentages given at each axis present the proportion of variance explained on the axis.

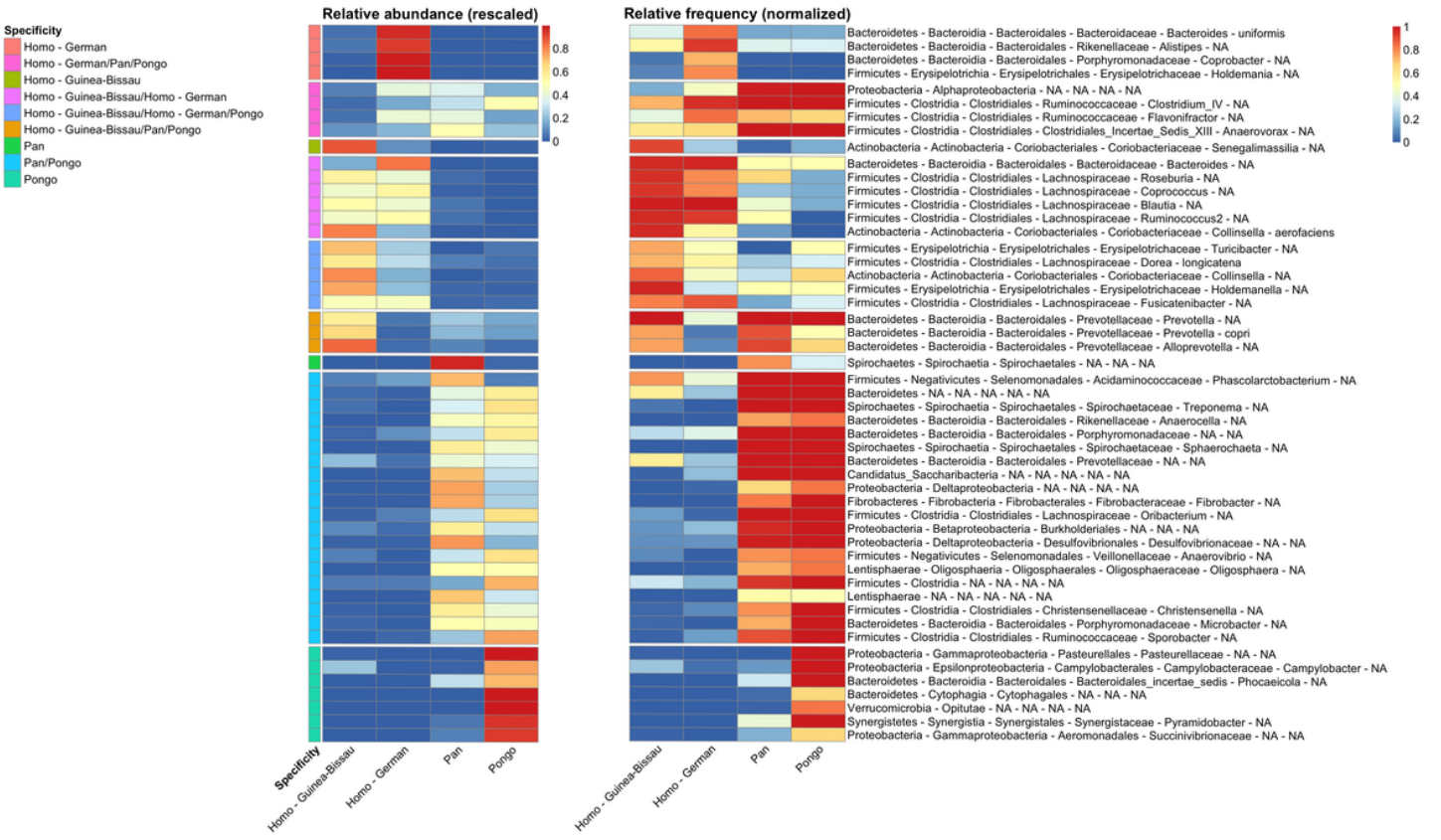

Figure 4

Microbial variation within the Hominidae family. Heatmap of Hominidae displaying the relative abundance (rescaled by rowsum for each species) and frequency of the 53 most significant species found by multi-level pattern analysis (multipatt in R package indicspecies), which was used to identify species that showed specificity in terms of abundance and prevalence to one of the subgroups Pan, Pongo, German or Guinea-Bissau human subjects, or a combination of those. Figure generated using pheatmap::pheatmap and arranged using Inkscape [26]. Relative abundance is rescaled.

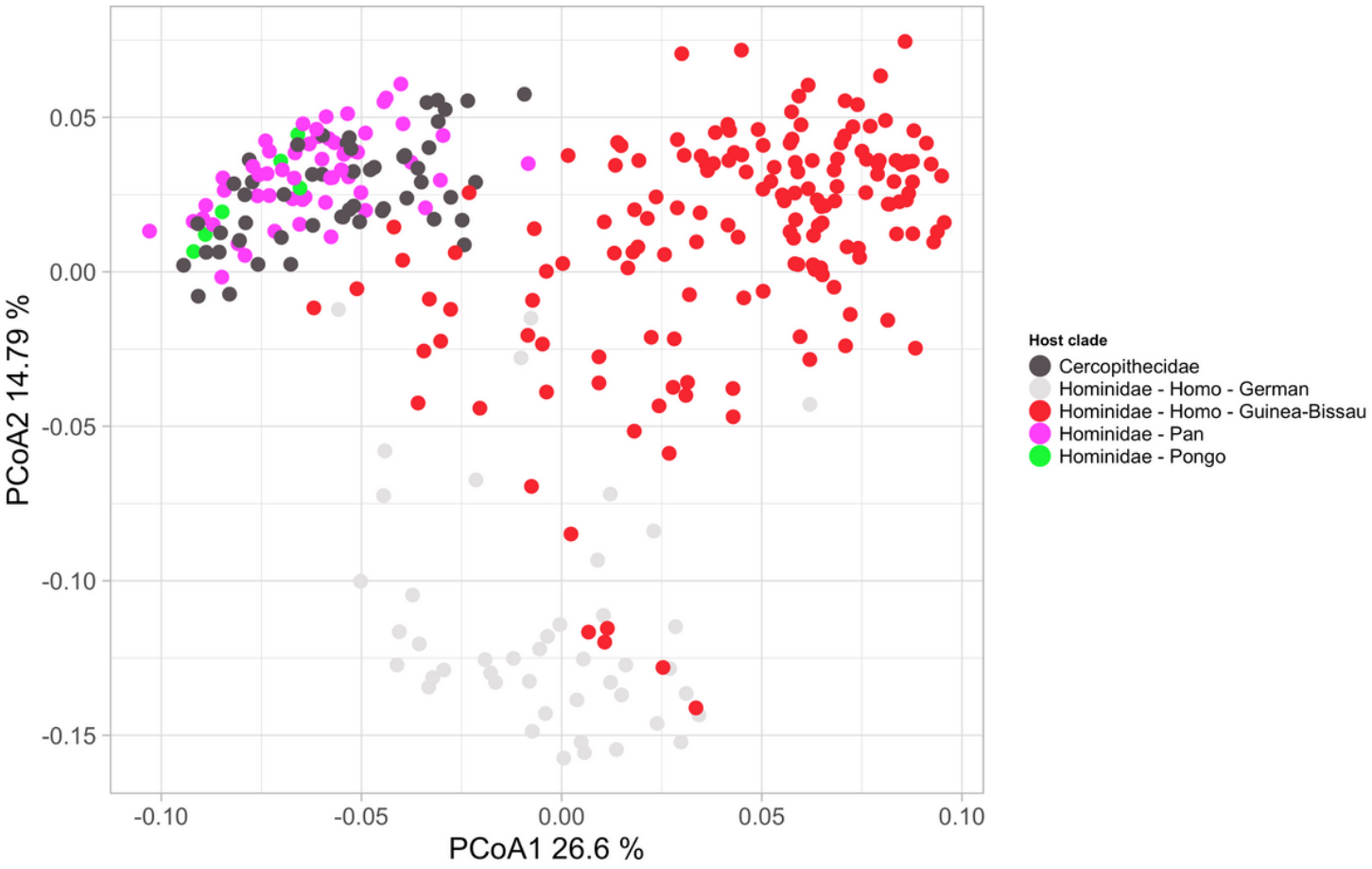

Figure 5

Community structure by host phylogeny for humans and New and Old World monkeys, with humans separated by sampling location. Samples ordinated by microbial Bray-Curtis dissimilarity showed a pattern of microbial composition dictated more by ecology than by host phylogeny, when comparing the New and Old World monkeys to humans. The plot show host animals colored by phylogeny, and location for the humans. The first ordination axis separates the humans from the non-human species, while the second axis separates the westernized and rural human samples. The ordination is unconstrained principal coordinates analysis made with vegan::capscale, with dist="bray", metaMDS = F. Percentages given at each axis present the proportion of variance explained on the axis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BMCMicrobiomeSuppl..docx
- TableS1TAXAmultipattHostOrder.xls
- TableS2TAXAsesmpdspecies.n536.xls
- TableS3SummarystatsadonisPhylogenymin1samp.xls
- TableS4TAXAmultipattHominidea.xls