Eukaryotic composition across seasons and social groups in the gut microbiota of wild baboons

Mary Chege[1,2,3*†], Pamela Ferretti[4], Shasta Webb[3], Rosaline W. Macharia[2], George Obiero[2], Joseph Kamau[1], Susan C. Alberts[5,6], Jenny Tung[5,6,7,8], Mercy Y. Akinyi[1], Elizabeth A Archie[3†]

1. One Health Centre, Kenya Institute of Primate Research, Nairobi, Kenya
2. Department of Biochemistry, University of Nairobi, Nairobi, Kenya
3. Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, USA
4. Department of Medicine, Genetic Medicine Section, University of Chicago, Chicago, USA
5. Departments of Biology and Evolutionary Anthropology, Duke University, Durham, NC, USA
6. Duke University Population Research Institute, Duke University, Durham, NC, USA
7. Department of Primate Behavior and Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany
8. Canadian Institute for Advanced Research, Toronto, Ontario, Canada

†corresponding authors:
Mary Chege: mchege89@gmail.com
Elizabeth Archie: earchie@nd.edu

Abstract

Background: Animals coexist with complex microbiota, including bacteria, viruses, and eukaryotes (e.g., fungi, protists, and helminths). While the composition of bacterial and viral components of animal microbiota are increasingly well understood, eukaryotic composition remains neglected. Here we characterized eukaryotic diversity in the microbiomes in wild baboons and tested the degree to which eukaryotic community composition was predicted by host social group membership, sex, age, and season of sample collection.

Results: We analyzed a total of 75 fecal samples collected between 2012 and 2014 from 73 wild baboons in the Amboseli ecosystem in Kenya. DNA from these samples was subjected to shotgun metagenomic sequencing, revealing members of the kingdoms Protista, Chromista, and Fungi in 90.7%, 46.7%, and 20.3% of samples, respectively. Social group membership explained 11.2% of the global diversity in gut eukaryotic species composition, but we did not detect statistically

significant effect of season, host age, and host sex. Across samples, the most prevalent protists were *Entamoeba coli* (74.66% of samples), *Enteromonas hominis* (53.33% of samples), and *Blastocystis subtype* 3 (38.66% of samples), while the most prevalent fungi included *Pichia manshurica* (14.66% of samples), and *Ogataea naganishii* (6.66% of samples).

Conclusions: Protista, Chromista, and Fungi are common members of the gut microbiome of wild baboons. More work on eukaryotic members of primate gut microbiota is essential for primate health monitoring and management strategies.

keywords: eukaryotes, gut microbiome, wild baboons, fungi, protists, social groups

Introduction

Vertebrate gut microbiota are complex and dynamic communities of bacteria, viruses, archaea, and eukaryotes [1,2]. To date, most research on vertebrate gut microbiota has focused on bacteria, in part because bacteria are both easy to characterize via 16S rRNA gene sequencing [3,4] and because bacteria are important to host health, with well-known effects on host metabolism, vitamin biosynthesis, and immune modulation [5–8]. However, vertebrate gut microbiota also contain eukaryotes, such as protists, metazoans, and fungi. Research on these eukaryotic communities remains neglected, in part because genetic and bioinformatic methods to characterize these communities are less developed than those for bacteria [9–12].

To date, the best characterized eukaryotic gut communities are those found in humans and laboratory mice. Humans living in a wide range of geographic locations harbor gut eukaryotes, including protists such as *Blastocystis*, *Entamoeba*, and *Enteromonas*, and fungi such as *Saccharomyces*, *Candida*, *Penicillium*, *Aspergillus*, and *Malassezia* [13–15]. The relevance of

these taxa to host health often remains unclear [16]. Some of these protists, such as *Entamoeba histolytica*, *Blastocystis hominis, and Aspergillus fumigatus,* may have pathogenic effects on hosts [14,17–21], while many others likely have commensal relationships with their hosts, and their presence may indicate a normal, healthy gut microbiota [16,22,23]. For instance, in mice, *Tritrichomonas musculis* is associated with host immune modulation and protection against bacterial mucosal infections [24].

In contrast, little is known about the eukaryotes living in the intestines of wild animals, including non-human primates [25–27]. Here we characterize eukaryotic members of the gut microbiota of wild baboons (*Papio sp.*) living in the Amboseli ecosystem in Kenya [28]. Prior work in this population has revealed several variables that influence individual exposure and susceptibility to gut bacteria, and we predicted that these same variables would also be important in predicting gut eukaryotic composition, including host social group membership [29–32], the season of sample collection [33–39], and host age [34,40–42]. For instance, social group membership influences baboon ranging patterns, resource use, and social relationships, which might influence microorganism exposure and transmission. In support, baboons from different social groups show distinct gut bacterial communities, and social group membership explains more variance in the gut bacterial microbiome than host age or sex [29]. Seasonality also shapes baboon diet, water source use, and other aspects of the environment, leading to systematic fluctuations in the abundance of several bacterial taxa as a function of the season of sample collection [43].

Our objectives in this paper were to characterize common gut eukaryotes in wild baboons and test the effects of seasonality, host social group membership, host sex, and host age on baboon gut eukaryotic community composition and diversity. We accomplished this objective by leveraging two shotgun metagenomic data sets: (i) samples from 48 individual baboons living in

two social groups in July and August 2012 [29], and (ii) samples from 27 individual baboons collected during the wet (April and May) and dry seasons (September) of 2014 (two individuals were included in both data sets so the total sample size was 75 samples from 73 individual baboons). We expected that gut eukaryotic communities would be influenced by similar factors to gut bacterial communities in this population. Our results provide new insights into the gut eukaryotic composition of wild baboons, contributing useful information for understanding the biology and health of wild primates.

Materials and methods

Study subjects and sample material

The baboons sampled in this study were studied by the Amboseli Baboon Research Project (ABRP) in the Amboseli ecosystem in Kenya. Founded in 1971, the ABRP conducts longitudinal research on known individual baboons living in several social groups [28]. Members of this population are hybrids between yellow and anubis baboons (*Papio cynocephalus* and *P. anubis*) [44–46]. The Amboseli ecosystem experiences a 5-month dry season from June through October, followed by a 7-month wet season with highly variable rainfall [47].

Data on the Amboseli baboons is collected by experienced observers year-round during 5-hour monitoring sessions, six days per week. During these sessions, observers collect fecal samples opportunistically from individuals, all of which are known to the observers from distinctive physical features. For each sample, the baboon's social group membership is known from group censuses collected during each monitoring session. Age is known to within a few days' error for all individuals born into ABRP study groups (n=64 baboons in this study), and estimated for other individuals based on observable morphological characteristics and body condition (n=9 baboons

who were all immigrant males, as males but not females disperse from their natal groups in this population). Sex is known based on morphological differences. All sampled individuals appeared healthy upon visual examination. All fecal samples are collected within a few minutes of defecation, thoroughly mixed, and preserved in 95% ethanol (2:5 feces to ethanol).

We analyzed eukaryotic composition in two sets of fecal samples. The first set was used to test for social group effects on eukaryotic composition. This set included samples from 48 adult baboons living in two social groups: 'Mica's group', (11 females and 8 males), and 'Viola's group', (20 females and 9 males). These samples were collected within a one-month window during the dry season of 2012 (**Figure 1**; **Supplementary Table 1;** an analysis of bacterial microbiome composition from these samples was published in Tung et al. [29]). The second dataset was used to test for seasonal differences in eukaryotic composition. It included samples from 27 baboons collected during the dry season in September 2014 (n=15) and the wet season between April and May 2014 (n=12) (**Figure 1; Supplementary Table 1**). Host sex and age were known for individuals in both data sets and two individuals occurred in both data sets.

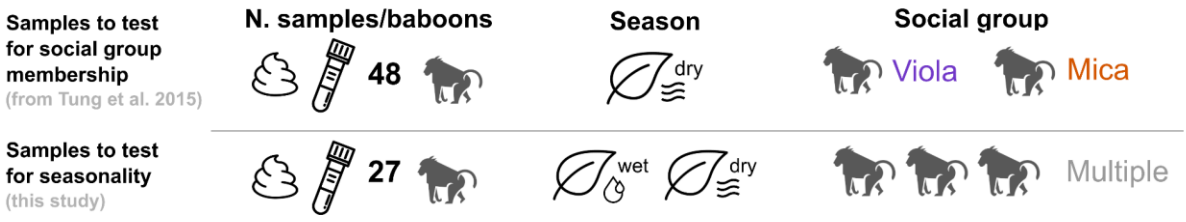

Figure 1. Schematic of the two sets of fecal samples investigated in this study (**Supplementary Table 1**). Shotgun metagenomic data for the first set of samples on social group membership was published in 2015 by Tung et al. [29]; these data are publicly available in the NCBI's Short Read Archive (Bioproject PRJNA271618). Shotgun metagenomic data for the second set of samples on seasonality were generated for the present study; these data are publicly available on NCBI's Short Read Archive (Bioproject PRJEB81717). Data on host sex and age were available for samples in both data sets.

Genomic DNA extraction and sequencing

Total DNA was extracted from the first (social group) set of fecal samples using the MO BIO Laboratories, Inc., Carlsbad, CA, PowerSoil DNA Isolation kit. Samples for the second (seasonal) set were extracted using Qiagen's DNeasy PowerSoil kit (Venlo, Netherlands). Both protocols were performed according to the manufacturer's instructions.

For samples in the first (social group) dataset [29], 200 ng of DNA were prepared for metagenomic sequencing on an Illumina HiSeq 2500, using Kapa Biosystems Library Preparation Kits (Kapa Biosystems, Wilmington, MA). The DNA samples were sheared to an average size of 400 base pairs, followed by ligation to barcoded adapters. The libraries were subjected to 100 base pair paired-end sequencing at the UCLA Neuroscience Genomics Core. In total, 1.4 billion raw, paired-end Illumina sequences were generated, with a mean ± SD of 14.4 ± 13.7 million read pairs per sample. The raw metagenomic sequencing data are deposited in the NCBI's Short Read Archive (BioProject PRJNA271618).

For samples in the second (seasonal) dataset, 200 ng of DNA were prepared for metagenomic sequencing on the Illumina NovaSeq X, utilizing the SeqWell purePlex DNA Library Preparation Kit (SeqWell, Beverly, MA). Samples were subjected to 150 base pair paired-end sequencing at the University of Minnesota Genomics Core. In total, 2.54 billion raw, paired-end Illumina sequences were generated, with a mean ± SD of 94 ± 31.2 million read pairs per sample. All the raw reads for the second dataset (n = 27) are deposited in NCBI's Short Read Archive (BioProject PRJEB81717).

145

146 **Identifying dominant eukaryotic members of the baboon gut microbiome**

147 The raw data were processed using FastQC [48] to assess the quality of the reads. Duplicate reads
148 were removed using FastUniq [49] and trimmed to remove adapter sequences and low-quality
149 bases (PHRED score <20) using Trimmomatic (v0.39) in paired-end mode [50]. We also removed
150 read pairs where one read was shorter than 75 base pairs after trimming.

151 We then used two previously-developed pipelines to further filter our sequences and
152 identify the presence/absence of eukaryotic taxa in each sample: EukDetect [51]
153 (https://github.com/allind/EukDetect) and its descendent pipeline gutprotist-search
154 https://github.com/allind/gutprotist-search). In brief, EukDetect aligns reads to a database
155 consisting of conserved eukaryotic marker genes from curated whole genome assemblies [51].
156 Gutprotist-search, developed alongside EukDetect, complements the approach in EukDetect by
157 using a database of NCBI sequences for particular taxonomic identifiers for eukaryotic taxa that
158 lack genome assemblies. Because genome sequences are unavailable for many gut eukaryotes,
159 gutprotist-search helps identify taxa that might otherwise go undetected.

160 Both pipelines were run using the recommended Snakemake workflow engine [51,52]
161 with default parameters. The metagenomic reads were aligned to the EukDetect marker database
162 and the gutprotist-search database using Bowtie2 [53], followed by stringent quality filtering
163 based on mapping quality to retain reads with a PHRED score ≥30, ensuring high base-call
164 accuracy. Sequence complexity was assessed using a complexity score threshold of ≥0.5 to retain
165 only high-complexity reads, reducing spurious alignments due to low-complexity regions.
166 Additional filtering as described in EukDetect protocol extended these steps to refine taxonomic
167 assignments by addressing off-target alignments arising from false positives. Specifically, reads

mapping to multiple species within a genus were compared for sequence identity and coverage, with lower confidence species excluded, and the ETE3 toolkit [54] was used to validate alignment accuracy. Following recommended practice, only taxa with more than four reads aligning to at least two distinct marker genes were retained, but unfiltered results were also explored for low-abundance species as recommended by the authors [51]. The EukDetect and gutprotist-search pipelines output a list of the eukaryotic taxa identified in the samples and associated relevant statistics such as the number of observed marker genes per sample per taxon, number of reads mapping to the marker genes, total marker gene coverage and identity percentage. We combined the results of these pipelines for each sample into a single table for our analyses (**Supplementary Table 2**). In this table, taxa were marked present (1) in a given sample if they were detected in one or both pipelines and absent (0) from a sample if they were not detected by either pipeline. Finally, the insect metazoan *Callosobruchus maculatus* was detected in 1 sample, but excluded from downstream analyses as it was likely acquired from food, and therefore not part of the gut eukaryotic community.

Statistical analyses

All statistical analyses were conducted in the R statistical environment (R version 4.3.3) [55]. We began our analyses by reporting the eukaryotic taxa that we identified across both sets of samples (N=75 samples; 48 from the first (social) set and 27 from the second (seasonal) set).

Next, to test whether eukaryotic species richness differed between baboon social groups (first sample set) or season (second sample set), we calculated the Simpson's alpha diversity index using the package *vegan* [56]. We then tested for differences in alpha diversity as a function of social group, season, host age, host sex, and the number of paired-end sequences generated for each sample using a linear model implemented in the *stats* package [55].

To test for compositional differences between eukaryotes as a function of baboon social group or season, we calculated Jaccard distance matrices within each data set using the vegdist function in *vegan* [56] (Jaccard distances measure dissimilarity between samples based on the presence/absence of taxa in each sample). To visualize how eukaryotic species composition varied between social groups or seasons, we used non-metric multidimensional scaling (NMDS). To determine the proportion of variance in community composition attributable to social group or season and host sex or age, we conducted a permutational multivariate analysis of variance (PERMANOVA) implemented in the adonis function in *vegan*, with 10,000 permutations [56] To test the effect of social group and season on the presence of each eukaryotic species within each data set using a binomial generalized linear model (GLM) implemented in the *stats* package. The presence or absence of each eukaryotic species was modeled as a response variable while the predictor variables were social group membership (Mica's or Viola's), season (wet or dry), host sex, host age, and read count.

Ethics statement

All protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Notre Dame to cover behavior observations and fecal sample collections in baboons at Amboseli, under protocol number 22-05-7259.

Results

Metagenomic analysis reveals diverse gut eukaryotic communities in wild baboons

Across both the social (n=48) and seasonal (n=27) data sets, we found genetic evidence for 21 eukaryotic species (**Supplementary Figure 1; Supplementary Table 3;** range=0 to 7 taxa per

sample). The mean number of eukaryotic taxa per sample across both data sets was 3.10 (SD= 1.60). Within each dataset, the average number of eukaryotic taxa present per sample was 3.09 (SD=1.47) in the social group dataset and 3.12 (SD= 1.84) in the seasonal dataset. Information on the potential transmission mode and health relevance of each taxon is in **Supplementary Table 4**. Additionally, we detected genetic evidence for the metazoan arthropod, *Callosobruchus maculatus* in 1.33% (n=1 of samples). Because baboons frequently ingest insects, these sequences may come from a closely-related insect in the baboon diet, and are not living members of the microbiome community. Sequences attributed to *C. maculatus* are removed from our analyses.

Overall, Protista was the most well-represented kingdom (n=68 samples; 90.7%), followed by Chromista (n=35 samples; 46.7%), and Fungi (n=22 samples; 29.3%; **Supplementary Figure 1**). Over half of the samples in the data set (n=56 samples; 74.7%) contained at least two detectable eukaryotic taxa, while 2.7% (n=2 samples) contained 7 eukaryotic taxa, the maximum number we observed. Only 2 samples (2.67%) had no detectable eukaryotes (one in the social group data set and one in the seasonal data set). Across all samples, the five most prevalent species were *Entamoeba coli* (n=56 samples; 74.7%), *Enteromonas hominis* (n=40 samples; 53.3%), *Blastocystis subtype 3* (n=29 samples; 38.7%), *Iodamoeba sp* (n=27 samples; 36%), and *Chilomastix mensnili* (n=26 samples; 34.7%; **Supplementary Table 3**). We also detected a set of species found in one or only a handful of samples, including (*Candida blattae, Malassezia restricta, Ogatea naganashii, Preussia sp.*, and *Aspergillus sydowii* (see **Supplementary Table 3**).

Social group membership predicts gut eukaryotic diversity and composition

In the first (social group membership) data set, we found significant differences in the gut eukaryotic alpha diversity for baboons in different social groups (linear model test, p=0.015;

Figure 2A; Supplementary Table 5). Samples from baboons living in Viola's group exhibited higher Simpson's diversity compared to those living in Mica's group (**Figure 2A**).

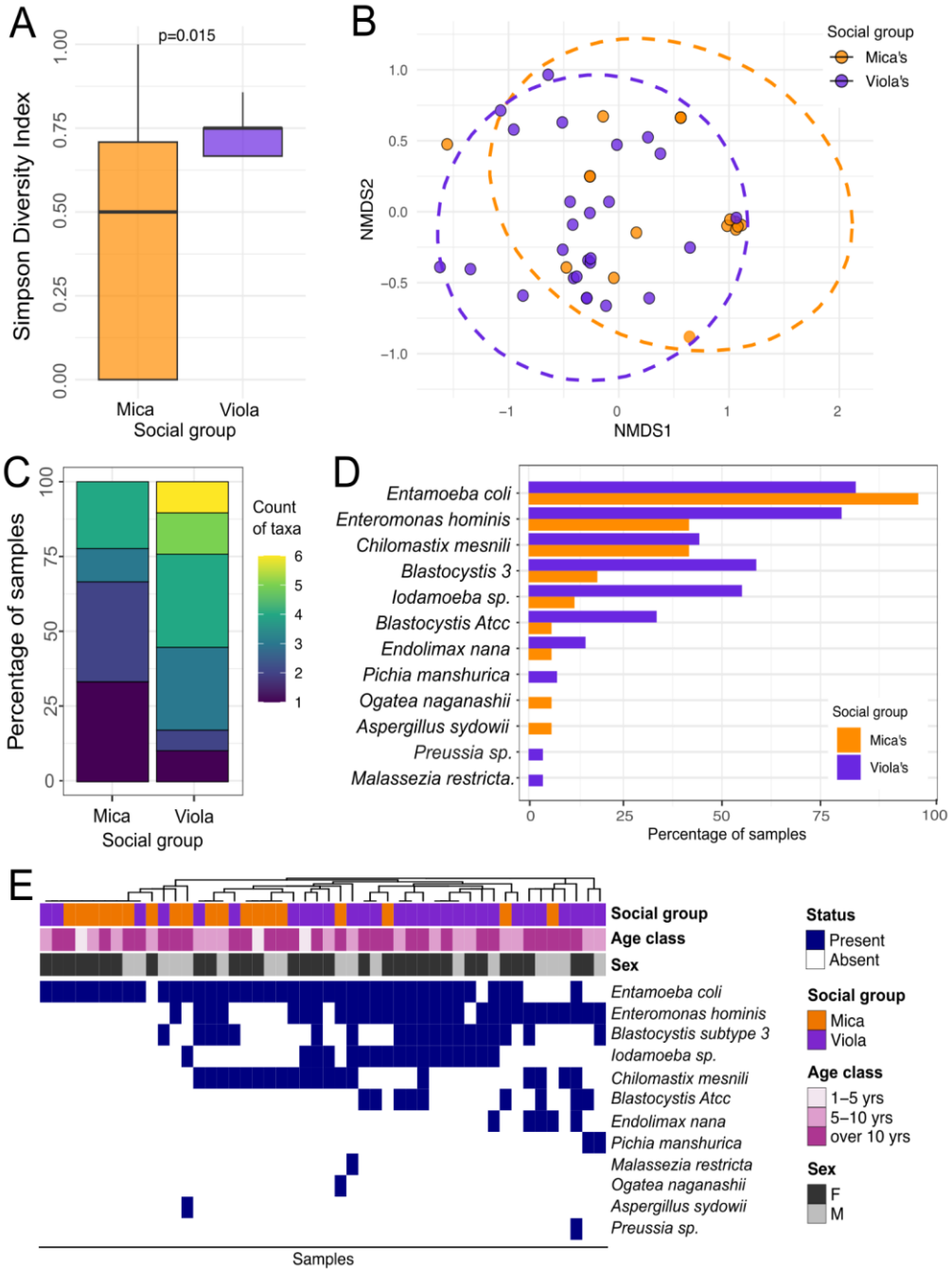

Figure 2. Social group membership is correlated with baboon gut eukaryotic composition. (A) Eukaryotic species diversity between Viola's and Mica's social groups using the Simpson's diversity index. A linear model was used to calculate statistical significance. (B) Non-metric multidimensional scaling (NMDS) plot of gut eukaryotic composition as measured using Jaccard index dissimilarity matrices. (C)

Number of eukaryotic taxa per sample in Mica and Viola's social groups and (D) prevalence of eukaryotic taxa in the two social groups. (E) Heatmap of the eukaryotic taxonomic composition of samples used to test for social group membership, with each column representing one fecal sample and metadata on social group, season, age class, and sex of the host. Samples are clustered using Euclidean distance of eukaryotic community composition.

Eukaryotic composition was also significantly different between Viola's and Mica's groups, explaining 11.2% of the variation in eukaryotic composition (PERMANOVA: $r^2 = 0.111$ $p = 1x10^{-3}$; **Figure 2B; Supplementary Table 6**). We also found a trend such that the number of paired end sequences in each sample had a small effect on eukaryotic composition ($r^2 = 0.045$; $p = 0.05$). By contrast, host age and sex did not make statistically significant contributions to gut eukaryotic composition (age, $p = 0.45$; sex, $p = 0.504$; see **Supplementary Table 6)**.

On average, 50% of Viola's and 12.5% of Mica's samples had at least 3 eukaryotic taxa (**Figure 2C**). In support of the patterns of alpha diversity we observed (**Figure 3A**), most taxa were more prevalent in Viola's group as opposed to Mica's group, including *Blastocystis subtype* 3 (coefficient: -2.03, p=0.01; **Supplementary Table 7**), *Blastocystis. ATCC* (coefficient: -1.99, p=0.08), *Enteromonas hominis* (coefficient: -1.77, p=0.01), *Iodamoeba sp.* (coefficient: -2.21, p=0.01) and *Endolimax nana* (coefficient: -1.82, p=0.20). Only one taxon, *Entamoeba coli,* was more prevalent in the Mica's group compared to the Viola's group (**Figure 2D-E**). *Ogataea naganishii* and *Aspergillus sydowii* were found at low prevalence (2%) only in samples from Mica's group, while *Preussia sp. BSL10* and *Malassezia restricta* were exclusively found in samples from Viola's group (2%, **Figure 2D-E**).

Season did not predict gut eukaryotic diversity and composition

We found no significant differences in the alpha diversity of the eukaryotes between samples collected in the dry season compared to the wet season (Simpson's diversity: effect size = 3.6%, Linear model test $p = 0.23$; **Figure 3A**). Across all 27 samples in the seasonal dataset, no

significant variation in taxonomic composition was explained by season, age, sex, or number of paired end sequences (PERMANOVA, $r^2 = 0.036$; p = 0.48; age, p = 0.20; sex, p = 0. 64, number of paired end sequences, p = 0.13; **Figure 3B**).

On average, 25.9% of samples collected in the dry season and 22.2% of samples collected in the wet season contained at least 3 detectable eukaryotic taxa (**Figure 3C**). The eukaryotic communities across all the samples in the seasonal data set were clustered according to their diversity and composition, which indicated small taxon-level distinctions between the wet and the dry seasons, despite no strong evidence for global seasonal shifts. No species differed in prevalence between seasons (**Supplementary Table 7**). In line with our results on the social group dataset, *Entamoeba coli* emerged as the most prevalent and dominant species (n= 16 samples; 59.3%), followed by *Enteromonas hominis* (n= 11 samples; 40.7%), *Blastocystis* subtype 3 and *Iodamoeba* sp. (n= 10 samples each; 37% each), *Pichia manshurica* (n= 9 samples; 33.3%), *Chilomastix mesnili* (n= 7 samples; 25.9%), *Ogatea naganishii* (n= 4 samples; 14.8%), *Blastocystis* ATCC (n= 3 samples; 11.1%) and *Byssochlamys* sp. AF001 (n= 2 samples; 7.4%; **Figure 3D-E**) .

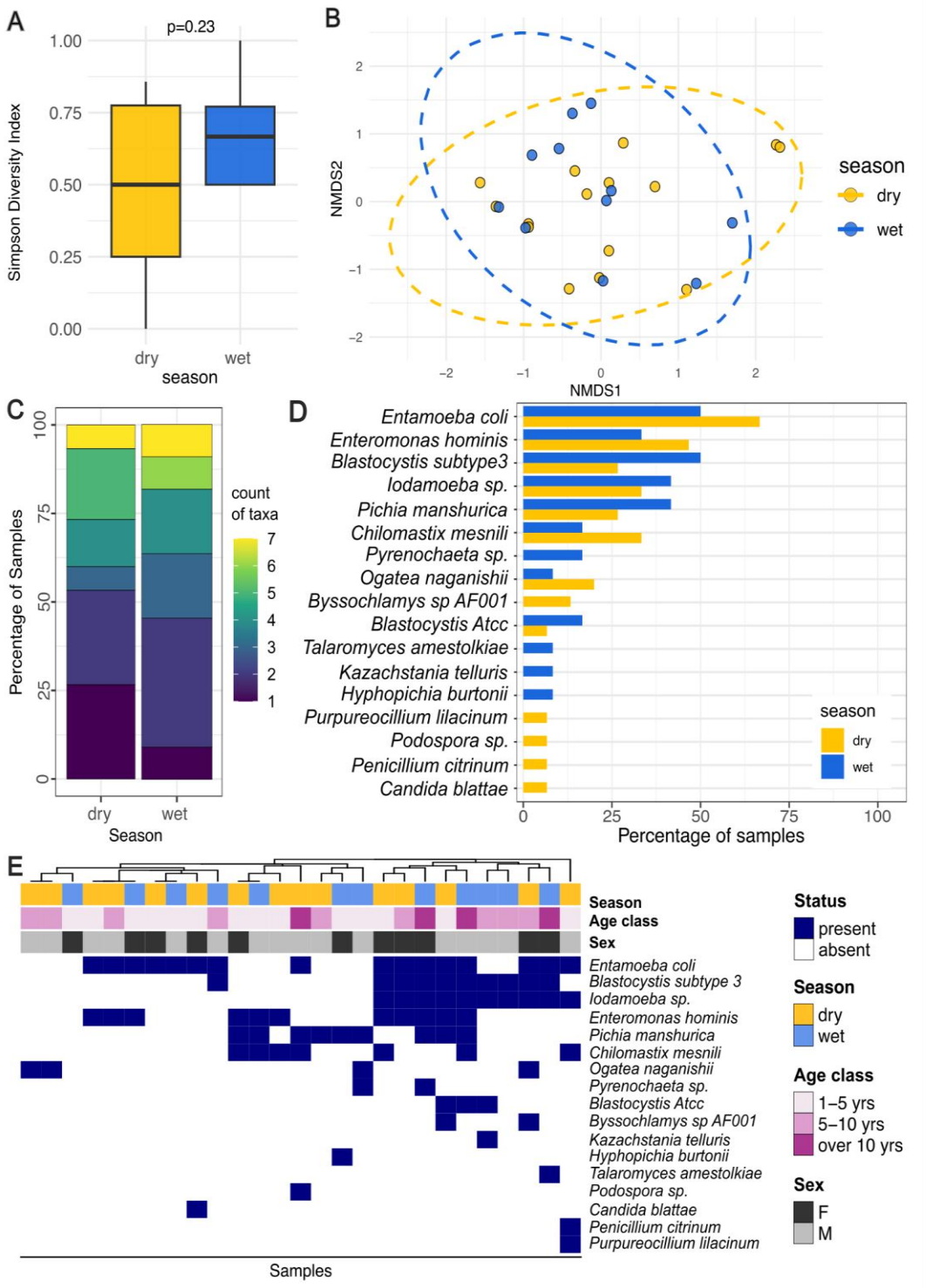

Figure 3. Season did not predict baboon gut eukaryotic diversity and composition. (A) Simpson's diversity index of baboon gut microbiome composition between the wet and dry seasons. A linear model was used to calculate statistical significance. (B) Non-metric multidimensional scaling (NMDS) plot of gut eukaryotic microbiome composition as measured using Jaccard index dissimilarity matrices. (C) Number of eukaryotic taxa per sample in the dry and wet season. (D) prevalence of eukaryotic taxa in the wet and

dry seasons. (E) Heatmap of the eukaryotic taxonomic composition of samples used to test for seasonality, with metadata on social group, season, age class, and sex of the host. Each column represents one fecal sample; samples are clustered using Euclidean distance of eukaryotic community composition.

Discussion

Although eukaryotes are commonly found in mammalian gut microbiomes, little is known about their importance for host health or the factors that drive their prevalence in wild non-human primates. Here, we investigated the composition and diversity of the eukaryotic communities inhabiting the gastrointestinal tract of wild baboons, focusing on the explanatory power of host social group membership, season, host sex, and host age. We detected genetic evidence for several eukaryotic taxa, spanning 3 kingdoms (Fungi, Protista, and Chromista in order of prevalence) and 21 species. Most of the taxa we found are considered non-pathogenic commensals and have been previously found in human populations across the world [57]. Among primates, *Entamoeba coli* has been identified in mountain gorillas (*Gorilla beringei beringei*) [58], western lowland gorillas (*Gorilla gorilla gorilla*) [59], red colobus (*Procolobus badius tephrosceles*) [60,61], red-tailed monkeys (*Cercopithecus ascanius schmidti*) [60,61], vervet monkeys (*Cercopithecus aethiops pygerythrus*) [60,61], baboons (*Papio anubis*), and chimpanzees [60,61]. The high prevalence of *Entamoeba coli* among non-human primates might be attributed to their social behavior and communal living, as Entamoeba species are known to be transmitted by intake of a mature cyst through either ingestion of contaminated water or food, or direct oral-fecal contact [62,63], in addition to the commensal relationship between the parasite and its hosts.

Among the most abundant taxa identified in baboons was *Blastocystis* (56%). *Blastocystis* is a common chromista found in humans, and non-human primates such as baboons. gorillas, chimpanzees, and other mammals, as well as in birds [64–66]. Whether *Blastocystis* is commensal

or pathogenic is still under debate [17]. In humans, members of the *Blastocystis* genus are common and stable colonizers of the gut of healthy individuals [67,68]. However, *Blastocystis* has also been associated with some gastrointestinal disorders such as inflammatory bowel disease (IBD) and irritable bowel syndrome (IBS) in humans [69], indicating possible pathogenicity. The role of *Blastocystis* in non-human primates remains unclear [66]. The high prevalence of *Blastocystis* among the baboons included in this study, which all appeared healthy upon visual examination when sampled, suggests that *Blastocystis* might be a commensal member of the gut microbiome of wild baboons.

The most common *Blastocystis* identified in the gut of the baboons in our study was *Blastocystis* subtype 3, found in 38.7% of the samples. *Blastocystis* subtype 3 has been previously found in humans and livestock [64], and is among the *Blastocystis* subtypes previously connected to human gut colonization [64,70]. Together, these results indicate the possibility for transmission of *Blastocystis* between wild baboons, livestock, and human populations, though further research is warranted to test this possibility. Transmission between species would not be surprising as both wild baboons and livestock share water holes during the dry season, and *Blastocystis* are commonly transmitted through shared water sources. Other taxa were less common in our samples (found in less than 10.7% of samples), and included *Candida blattae, Malassezia restricta* (which is a member of the normal primate skin microbiota, occasionally contributing to skin conditions), *Ogatea naganashii, Preussia sp.* (a non-pathogenic genus), and *Aspergillus sydowii* (which can cause respiratory infections in immunocompromised individuals; **Supplementary Table 4**). The latter was exclusively found in Mica's group at low prevalence (2%), suggesting that, in contrast

to recent observations in wild macaques (*Tibetan macaque*, *Macaca fascicularis*, and *Macaca namestrina*) [71,72], *Aspergillus* is not a common member of the wild baboon gut microbiome.

Among the variables we investigated, social group membership was the strongest predictor of eukaryotic diversity, explaining approximately 14.8% of the variation in eukaryotic community composition. Specifically, the larger social group (Viola) was characterized by greater eukaryotic diversity. This pattern may be due to the fact that individuals in larger groups interact with more hosts in larger as compared to smaller groups, and as a result they may be exposed to more diverse eukaryote communities. This result is consistent with a previous study from the Amboseli ecosystem [73] and another on geladas [74], showing that members of a larger social group exhibited higher diversity in their gut microbiota than individuals belonging to smaller groups. However, another possibility is that the increased eukaryotic diversity observed in Viola's group is related to their home range occupancy. Viola's home range was bigger than Mica's at the time the samples we analyzed were collected [29], suggesting that Viola's group could have been exposed to a wider variety of resources, substrates, and, ultimately, microbes. Other variables such as sex and age were less relevant in explaining eukaryotic diversity in our sample (1.6 % and 1.7%, respectively). Combined, these results suggest that social group membership plays an important role in eukaryotes' ability to colonize hosts, and are in line with previous studies that reported a clear association between social group membership and the composition of bacterial communities in the gut of humans, non-human primates, carnivores, rodents, insects, and birds [29,75–83]. Physical contact, shared environment, and group behaviors might therefore influence both prokaryotic and eukaryotic diversity and composition via similar mechanisms.

We found that eukaryotic diversity was slightly higher during the wet season compared to the dry season, and that the eukaryotic composition in the gut varied across seasons. However, these trends were not large nor statistically significant. This result contrasts with previous studies that identified seasonality as highly relevant in shaping the bacterial component of the gut microbiome of human [37] and animal populations [33,35,36], including baboons [35–37,84]. However, this result may be due to the limited sample size of the second dataset used to investigate the impact of seasonality. Further investigation in larger cohorts is warranted to assess whether seasonality significantly impacts eukaryotic diversity in the gut microbiome of wild baboons. An additional limitation of our study is that the analysis of season on gut microbiome eukaryotic composition was performed on a cohort with highly heterogeneous group membership, and we could not use them to analyze the effects of social group membership due to low statistical power for this variable. Likewise, the social group membership sample set was collected entirely in the dry season and could not be used to assess seasonal variability. Therefore, the importance of these two variables in relation to each other is not clear from this study.

Conclusions

To our knowledge, this is the first metagenomic study to characterize the eukaryotic gut microbiota of wild baboons. Taken together, our results indicate that eukaryotes are an important part of the microbial communities inhabiting the gut of wild baboons, and that social group membership plays a role in shaping gut eukaryotic composition and diversity over time. Understanding how social factors affect microbiome composition could therefore be informative about the evolution of social behavior and its health implications, and reinforces the importance of considering social dynamics in microbiome research.

Data availability

The raw metagenomic sequences for the social data set dataset and the seasonal data set presented in this study are available on NCBI Sequence Read Archive (SRA) under the BioProject accession numbers PRJNA271618 and PRJEB81717, respectively. Comprehensive metadata for the samples introduced in this study are available as the Supplementary Material. Code is available on github: https://github.com/ArchieLab/chege_etal_2024.

Acknowledgements

We thank Jeanne Altmann for her essential role in stewarding the Amboseli Baboon Project. In Kenya, we thank the Kenya Wildlife Service, the Wildlife Training and Research Institute, the National Council for Science, Technology, and Innovation, and the National Environment Management Authority for permission to conduct research and collect biological samples. We also thank the University of Nairobi, Kenya Institute of Primate Research, National Museums of Kenya, the Amboseli-Longido pastoralist communities, the Enduimet Wildlife Management Area, Ker & Downey Safaris, Air Kenya, and Safarilink for their cooperation and assistance in the field. We thank Karl Pinc for managing and designing the database. We thank Raphael Mututua, Serah Sayialel, Kinyua Warutere, Lilian Musembei, and Long'ida Siodi for collecting the field data and fecal samples. Complete acknowledgments for the Amboseli Baboon Research Project can be found online at https://amboselibaboons.nd.edu/acknowledgements/.

Author contributions

Conceptualization, EAA, MC, MYA, JT; Sample and metadata collection, EAA, JT, SCA; Samples processing and data analysis, MC, SW, PF; Writing – original draft, MC.; Writing – review & editing, MC, PF, SW, EAA, MYA, RM, GO, JK, JT, SCA; Supervision, RWM, GO,

MYA, EAA; Funding acquisition, EAA, JT, SCA. All authors read and approved the final manuscript.

Funding

This work was supported by the National Institutes of Health and National Science Foundation, especially the National Institute on Aging for R01 AG071684 (EAA), R21 AG055777 (EAA, RB), NIH R01 AG053330 (EAA), NIH R35 GM128716 (RB), NSF DBI 2109624 (MRD), and NSF DEB 1840223 (EAA, JAG). We also thank the Duke University Population Research Institute P2C-HD065563 (pilot to JT), the University of Notre Dame's Eck Institute for Global Health (EAA), and the Notre Dame Environmental Change Initiative (EAA). We also thank Duke University, Princeton University, the University of Notre Dame, the Chicago Zoological Society, the Max Planck Institute for Demographic Research, the L.S.B. Leakey Foundation and the National Geographic Society for support at various times over the years.

References

1. Berg G, Rybakova D, Fischer D, Cernava T, Vergès MCC, Charles T, et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome. 2020; 8:1–22.

2. Davenport ER, Sanders JG, Song SJ, Amato KR, Clark AG, Knight R. The human microbiome in evolution. BMC Biol. 2017; 15:1–12.

3. Mas-Lloret J, Obón-Santacana M, Ibáñez-Sanz G, Guinó E, Pato ML, Rodriguez-Moranta F, et al. Gut microbiome diversity detected by high-coverage 16S and shotgun sequencing of paired stool and colon sample. Scientific Data [Internet]. 2020;7. Available from: http://dx.doi.org/10.1038/s41597-020-0427-5

4. Wei F, Wu Q, Hu Y, Huang G, Nie Y, Yan L. Conservation metagenomics: a new branch of conservation biology. Sci China Life Sci. 2019; 62:168–78.

5. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, et al. Host-gut microbiota metabolic interactions. Science. 2012; 336:1262–7.

6. Marchesi JR, Adams DH, Fava F, Hermes GDA, Hirschfield GM, Hold G, et al. The gut microbiota and host health: a new clinical frontier. Gut. 2016; 65:330–9.

7. Koropatkin NM, Cameron EA, Martens EC. How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol. 2012; 10:323–35.

8. Sylvia KE, Demas GE. A gut feeling: Microbiome-brain-immune interactions modulate social and affective behaviors. Horm Behav. 2018; 99:41–9.

9. Amato KR, Metcalf JL, Song SJ, Hale VL, Clayton J, Ackermann G, et al. Using the gut microbiota as a novel tool for examining colobine primate GI health. Global Ecology and Conservation. 2016; 7:225–37.

10. Archie EA, Tung J. Social behavior and the microbiome. Current Opinion in Behavioral Sciences. 2015; 6:28–34.

11. Jasinska AJ, Dong TS, Lagishetty V, Katzka W, Jacobs JP, Schmitt CA, et al. Shifts in microbial diversity, composition, and functionality in the gut and genital microbiome during a natural SIV infection in vervet monkeys. Microbiome. 2020; 8:1–19.

12. Manara S, Asnicar F, Beghini F, Bazzani D, Cumbo F, Zolfo M, et al. Microbial genomes from non-human primate gut metagenomes expand the primate-associated bacterial tree of life with over 1000 novel species. Genome Biol [Internet]. 2019;20. Available from: http://dx.doi.org/10.1186/s13059-019-1923-9

13. Hamad I, Raoult D, Bittar F. Repertory of eukaryotes (eukaryome) in the human gastrointestinal tract: Taxonomy and detection methods. Parasite Immunol. 2016; 38:12–36.

14. Laforest-Lapointe I, Arrieta M-C. Microbial Eukaryotes: a Missing Link in Gut Microbiome

462 Studies. mSystems. 2018; 3:1–5.

463 15. Parfrey LW, Walters WA, Lauber CL, Clemente JC, Berg-Lyons D, Teiling C, et al.
464 Communities of microbial eukaryotes in the mammalian gut within the context of environmental
465 eukaryotic diversity. Front Microbiol. 2014; 5:1–13.

466 16. Chabé M, Lokmer A, Ségurel L. Gut Protozoa: Friends or Foes of the Human Gut
467 Microbiota? Trends Parasitol. 2017; 33:925–34.

468 17. Lukeš J, Stensvold CR, Jirků-Pomajbíková K, Wegener Parfrey L. Are Human Intestinal
469 Eukaryotes Beneficial or Commensals? PLoS Pathog. 2015; 11:7–12.

470 18. Parfrey LW, Walters WA, Knight R. Microbial eukaryotes in the human microbiome:
471 Ecology, evolution, and future directions. Front Microbiol. 2011; 2:1–6.

472 19. David M. Underhill and Iliyan D. Iliev. Host Immune System. Nat Rev Immunol. 2014;
473 14:405–16.

474 20. Badparva E, Kheirandish F. Blastocystis hominis: A Pathogenic Parasite. Arch Clin Infect
475 Dis [Internet]. 2020;15. Available from: http://dx.doi.org/10.5812/archcid.97388

476 21. Earle K, Valero C, Conn DP, Vere G, Cook PC, Bromley MJ, et al. Pathogenicity and
477 virulence of Aspergillus fumigatus. Virulence. 2023; 14:2172264.

478 22. Audebert C, Even G, Cian A, Blastocystis Investigation Group, Loywick A, Merlin S, et al.
479 Colonization with the enteric protozoa Blastocystis is associated with increased diversity of
480 human gut bacterial microbiota. Sci Rep. 2016; 6:1–11.

481 23. Tito RY, Chaffron S, Caenepeel C, Lima-Mendez G, Wang J, Vieira-Silva S, et al.
482 Population-level analysis of Blastocystis subtype prevalence and variation in the human gut
483 microbiota. Gut. 2019; 68:1180–9.

484 24. Aleksey Chudnovskiy et al. Host-protozoan interactions protect from mucosal infections
485 through activation of the inflammasome. Physiol Behav. 2016; 176:139–48.

486 25. del Campo J, Bass D, Keeling PJ. The eukaryome: Diversity and role of microeukaryotic
487 organisms associated with animal hosts. Funct Ecol. 2020; 34:2045–54.

488 26. Mann AE, Mazel F, Lemay MA, Morien E, Billy V, Kowalewski M, et al. Biodiversity of
489 protists and nematodes in the wild nonhuman primate gut. ISME J. 2020; 14:609–22.

490 27. Ramayo-Caldas Y, Prenafeta-Boldú F, Zingaretti LM, Gonzalez-Rodriguez O, Dalmau A,
491 Quintanilla R, et al. Gut eukaryotic communities in pigs: diversity, composition and host
492 genetics contribution. Animal Microbiome [Internet]. 2020;2. Available from:
493 http://dx.doi.org/10.1186/s42523-020-00038-4

494 28. Alberts SC, Altmann J. The Amboseli Baboon Research Project: 40 Years of Continuity and
495 Change. In: Kappeler PM, Watts DP, editors. Long-Term Field Studies of Primates. Berlin,

Heidelberg: Springer Berlin Heidelberg; 2012. p. 261–87.

29. Tung J, Barreiro LB, Burns MB, Grenier JC, Lynch J, Grieneisen LE, et al. Social networks predict gut microbiome composition in wild baboons. Elife. 2015; 2015:1–18.

30. Vernier CL, Chin IM, Adu-Oppong B, Krupp JJ, Levine J, Dantas G, et al. The gut microbiome defines social group membership in honey bee colonies. Sci Adv. 2020;6: eabd3431.

31. Antwis RE, Lea JMD, Unwin B, Shultz S. Gut microbiome composition is associated with spatial structuring and social interactions in semi-feral Welsh Mountain ponies. Microbiome. 2018; 6:207.

32. Leclaire S, Jacob S, Greene LK, Dubay GR, Drea CM. Social odours covary with bacterial community in the anal secretions of wild meerkats. Sci Rep. 2017; 7:1–13.

33. Long C, Scheun J, Sauther ML, Cuozzo FP, Millette J, Tordiffe ASW. Seasonal Effects on the Fecal Microbial Composition of Wild Greater Thick-Tailed Galagos (Otolemur crassicaudatus). Int J Primatol [Internet]. 2023; Available from: https://doi.org/10.1007/s10764-023-00407-1

34. Pafčo B, Sharma AK, Petrželková KJ, Vlčková K, Todd A, Yeoman CJ, et al. Gut microbiome composition of wild western lowland gorillas is associated with individual age and sex factors. Am J Phys Anthropol. 2019; 169:575–85.

35. Góngora E, Elliott KH, Whyte L. Gut microbiome is affected by inter-sexual and inter-seasonal variation in diet for thick-billed murres (Uria lomvia). Sci Rep. 2021; 11:1–12.

36. Baniel A, Amato KR, Beehner JC, Bergman TJ, Mercer A, Perlman RF, et al. Seasonal shifts in the gut microbiome indicate plastic responses to diet in wild geladas. Microbiome. 2021; 9:26.

37. Davenport ER, Mizrahi-Man O, Michelini K, Barreiro LB, Ober C, Gilad Y. Seasonal Variation in Human Gut Microbiome Composition. PLoS One. 2014;9: e90731.

38. Norton GW, Rhine RJ, Wynn GW, Wynn RD. Baboon Diet: A Five-Year Study of Stability and Variability in the Plant Feeding and Habitat of the Yellow Baboons (Papio cynocephalus) of Mikumi National Park, Tanzania. Folia Primatol. 1987;48:78–120.

39. Kartzinel TR, Hsing JC, Musili PM, Brown BRP, Pringle RM. Covariation of diet and gut microbiome in African megafauna. Available from: https://doi.org/10.5061/dryad.c119gm5

40. Amato KR, Leigh SR, Kent A, Mackie RI, Yeoman CJ, Stumpf RM, 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. 2014;155:652–64.

41. Reese AT, Phillips SR, Owens LA, Venable EM, Langergraber KE, Machanda ZP, et al. Age Patterning in Wild Chimpanzee Gut Microbiota Diversity Reveals Differences from Humans in Early Life. Curr Biol. 2021;31:613–20.e3.

42. Videvall E, Song SJ, Bensch HM, Strandh M, Engelbrecht A, Serfontein N, et al. Major shifts in gut microbiota during development and its relationship to growth in ostriches. Mol Ecol. 2019;28:2653–67.

43. Björk JR, Dasari MR, Roche K, Grieneisen L, Gould TJ, Grenier J-C, et al. Synchrony and idiosyncrasy in the gut microbiome of wild baboons. Nat Ecol Evol. 2022;6:955–64.

44. Alberts SC, Altmann J. Immigration and hybridization patterns of yellow and anubis baboons in and around Amboseli, Kenya. Am J Primatol. 2001;53:139–54.

45. Tung J, Charpentier MJE, Garfield DA, Altmann J, Alberts SC. Genetic evidence reveals temporal change in hybridization patterns in a wild baboon population. Mol Ecol. 2008;17:1998–2011.

46. Vilgalys TP, Fogel AS, Anderson JA, Mututua RS, Warutere JK, Siodi IL, et al. Selection against admixture and gene regulatory divergence in a long-term primate field study. Science. 2022;377:635–41.

47. Okello MM. An assessment of the large mammal component of the proposed wildlife sanctuary site in Maasai Kuku Group Ranch near Amboseli, Kenya. S Afr J Wildl Res. 2005;35:63–76.

48. Andrews S. FastQC: a quality control tool for high throughput sequence data [Internet]. 2010. Available from: https://www.bibsonomy.org/bibtex/2b6052877491828ab53d3449be9b293b3/ozborn

49. Xu H, Luo X, Qian J, Pang X, Song J, Qian G, et al. FastUniq: a fast de novo duplicates removal tool for paired short reads. PLoS One. 2012;7:e52249.

50. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

51. Lind AL, Pollard KS. Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. Microbiome. 2021;9:1–18.

52. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. Bioinformatics. 2012;28:2520–2.

53. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

54. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. Mol Biol Evol. 2016;33:1635–8.

55. R Core Team. R: A language and environment for statistical computing. 2014;

56. Oksanen J. Vegan: ecological diversity. R Package Version 24-4. 2017;1:11.

563 57. Monroe LS. Gastrointestinal parasites. Bockus gastroenterology, 5th ed The WB Saunders
564 Co, Philadelphia, Pa. 1995;3181–3.

565 58. Sleeman JM, Meader LL, Mudakikwa AB, Foster JW, Patton S. Gastrointestinal parasites of
566 mountain gorillas (Gorilla gorilla beringei) in the Parc National des Volcans, Rwanda. J Zoo
567 Wildl Med. 2000; 31:322–8.

568 59. Betts EL, Gentekaki E, Tsaousis AD. Exploring Micro-Eukaryotic Diversity in the Gut: Co-
569 occurrence of Blastocystis Subtypes and Other Protists in Zoo Animals. Front Microbiol
570 [Internet]. 2020;11. Available from: http://dx.doi.org/10.3389/fmicb.2020.00288

571 60. Kooriyama T, Hasegawa H, Shimozuru M, Tsubota T, Nishida T, Iwaki T. Parasitology of
572 five primates in Mahale Mountains National Park, Tanzania. Primates. 2012;53:365–75.

573 61. Larbi JA, Akyeampong S, Abubakari A, Offei Addo S, Okoto D, Hanson H. Zoonotic
574 Gastrointestinal Parasites of Baboons (Papio anubis) in the Shai Hill Reserve in Ghana. Biomed
575 Res Int. 2020;2020:1083251.

576 62. Ragazzo LJ, Zohdy S, Velonabison M, Herrera J, Wright PC, Gillespie TR. Entamoeba
577 histolytica infection in wild lemurs associated with proximity to humans. Vet Parasitol.
578 2018;249:98–101.

579 63. Ravdin JI. Entamoeba histolytica: from adherence to enteropathy. J Infect Dis.
580 1989;159:420–9.

581 64. Alfellani MA, Jacob AS, Perea NO, Krecek RC, Taner-Mulla D, Verweij JJ, et al. Diversity
582 and distribution of Blastocystis sp. subtypes in non-human primates. Parasitology.
583 2013;140:966–71.

584 65. Hublin JSY, Maloney JG, Santin M. Blastocystis in domesticated and wild mammals and
585 birds. Res Vet Sci. 2021;135:260–82.

586 66. Hamad I, Keita MB, Peeters M, Delaporte E, Raoult D, Bittar F. Pathogenic eukaryotes in
587 gut microbiota of western lowland gorillas as revealed by molecular survey. Sci Rep. 2014;4:1–
588 9.

589 67. Scanlan PD, Marchesi JR. Micro-eukaryotic diversity of the human distal gut microbiota:
590 qualitative assessment using culture-dependent and -independent analysis of faeces. ISME J.
591 2008;2:1183–93.

592 68. Scanlan PD, Stensvold CR, Rajilić-Stojanović M, Heilig HGHJ, De Vos WM, O'Toole PW,
593 et al. The microbial eukaryote Blastocystis is a prevalent and diverse member of the healthy
594 human gut microbiota. FEMS Microbiol Ecol. 2014;90:326–30.

595 69. Tan KSW, Mirza H, Teo JDW, Wu B, Macary PA. Current Views on the Clinical Relevance
596 of Blastocystis spp. Curr Infect Dis Rep. 2010;12:28–35.

597 70. Tan KSW. New insights on classification, identification, and clinical relevance of

598 Blastocystis spp. Clin Microbiol Rev. 2008;21:639–65.

599 71. Sun B, Huffman M, Li J-H. The gut microbiome of Tibetan macaques: Composition,
600 influencing factors and function in feeding ecology. The Behavioral Ecology of the Tibetan
601 Macaque [Internet]. 2019; Available from:
602 https://library.oapen.org/bitstream/handle/20.500.12657/22940/1/1007221.pdf#page=210

603 72. Mohd-Taib FS, Ahmad SN, Aqma WS, Sitam FT. Fungal communities in Macaca
604 fascicularis and Macaca nemestrina associates with forest matrix types in West Malaysia. J
605 Microbiol Exp. 2020;8:7–18.

606 73. Grieneisen LE, Livermore J, Alberts S, Tung J, Archie EA. Group living and male dispersal
607 predict the core gut microbiome in wild baboons. Integr Comp Biol. 2017;57:770–85.

608 74. Moeller AH, Foerster S, Wilson ML, Pusey AE, Hahn BH, Ochman H. Social behavior
609 shapes the chimpanzee pan-microbiome. Sci Adv. 2016;2:e1500997.

610 75. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, et al.
611 Cohabiting family members share microbiota with one another and with their dogs. Elife.
612 2013;2013:1–22.

613 76. Lax S, Smith DP, Hampton-marcell J, Owens SM, Handley M, Scott NM, et al. Longitudinal
614 analysis of microbial interaction between humans and the indoor environment. Science.
615 2014;345:1048–52.

616 77. Leclaire S, Nielsen JF, Drea CM. Bacterial communities in meerkat anal scent secretions
617 vary with host sex, age, and group membership. Behav Ecol. 2014;25:996–1004.

618 78. Chiyo PI, Grieneisen LE, Wittemyer G, Moss CJ, Lee PC, Douglas-hamilton I, et al. The
619 Influence of Social Structure , Habitat , and Host Traits on the Transmission of Escherichia coli
620 in Wild Elephants. 2014;9. Available from: http://dx.doi.org/10.1371/journal.pone.0093408

621 79. Koch H, Schmid-Hempel P. Socially transmitted gut microbiota protect bumble bees against
622 an intestinal parasite. Proc Natl Acad Sci U S A. 2011;108:19288–92.

623 80. Theis KR, Venkataraman A, Dycus JA, Koonter KD, Schmitt-Matzen EN, Wagner AP, et al.
624 Symbiotic bacteria appear to mediate hyena social odors. Proc Natl Acad Sci U S A.
625 2013;110:19832–7.

626 81. Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, et al. Gut microbiota
627 composition correlates with diet and health in the elderly. Nature. 2012;488:178–84.

628 82. Degnan PH, Pusey AE, Lonsdorf EV, Goodall J, Wroblewski EE, Wilson ML, et al. Factors
629 associated with the diversification of the gut microbial communities within chimpanzees from
630 Gombe National Park. Proc Natl Acad Sci U S A. 2012;109:13034–9.

631 83. Finnegan PM, Garber PA, McKenney AC, Bicca-Marques JC, De la Fuente MF, Abreu F, et
632 al. Group membership, not diet, structures the composition and functional potential of the gut

633 microbiome in a wild primate. mSphere. 2024;e0023324.

634 84. Ren T, Grieneisen LE, Alberts SC, Archie EA, Wu M. Development, diet and dynamism:
635 longitudinal and cross-sectional predictors of gut microbial communities in wild baboons.
636 Environ Microbiol. 2016;18:1312–25.

637