Title: Dental calculus metagenomics suggest that ecology, not host phylogeny, shapes the oral microbiome in closely related species

Markella Moraitou[1#], Adrian Forsythe[1#], James A. Fellows Yates[2,3], Jaelle C. Brealey[4], Christina Warinner[2,3,5,6], Katerina Guschanski[1,7,8, *]

[1]Animal Ecology, Department of Ecology and Genetics, Uppsala University, 75236 Uppsala, Sweden

[2]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, 07745 Jena, Germany

[3]Department of Paleobiotechnology, Leibniz Institute for Natural Product Research and Infection Biology Hans Knöll Institute, 07745 Jena, Germany

[4]Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology, 7491 Trondheim, Norway

[5]Faculty of Biological Sciences, Friedrich Schiller University, 07743 Jena, Germany

[6]Department of Anthropology, Harvard University, Cambridge, MA 02138

[7]Science for Life Laboratory, 75237 Uppsala, Sweden

[8]Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FL, United Kingdom

[#]These authors contributed equally

[*]Correspondence email: katerina.guschanski@ed.ac.uk

Key words: ancient DNA, museum specimens, wild animals, gorilla, metagenome-assembled genomes, shotgun metagenomics

Abstract

Hosts and their associated microorganisms, the microbiome, are tightly linked through a multitude of biological processes. Placed at the contact zone between external and internal environments, the oral microbiome has important roles in host health. However, little is known about factors that determine composition and function of the oral microbiome. Here we investigate the contribution of host ecology and evolutionary history in shaping the oral microbiome in three closely related gorilla subspecies that diverged 10,000-250,000 years ago. Shotgun metagenomics of dental calculus samples from 46 museum specimens of mountain (*Gorilla beringei beringei*), Grauer's (*G. b. graueri*), and western lowland gorillas (*G. gorilla gorilla*) allowed us to simultaneously investigate bacterial and dietary profiles. We find that the oral microbiomes of mountain gorillas are functionally and taxonomically distinct from western lowland and Grauer's gorillas, despite close evolutionary relationships and geographic proximity with Grauer's gorillas. Grauer's gorillas show an intermediate bacterial taxonomic composition and functional profile and intermediate dietary composition. Microbiome profiles appear to be related to altitude differences, which reflect habitat differences among the studied gorilla subspecies. We also detect increased abundance in nitrate-reducing oral taxa *Neisseria*, *Rothia* and *Veillonella* in mountain gorillas, which may contribute to high altitude adaptation by modulating blood pressure. Our results suggest that ecology, including diet, rather than evolutionary relationships and geographic proximity, primarily shape the oral microbiome in these closely related species. Further research will be needed to establish if this is a common feature of oral microbiome differentiation at the early stages of host speciation.

Background

The microbial communities that are found on and inside multicellular organisms are not only remarkably diverse, but also play a crucial role in a vast set of biological processes, such as energy uptake (Arora and Sharma, 2011; Nieuwdorp et al., 2014), detoxification (Nichols et al., 2019; Turner and Bucking, 2019), immune responses (Nikitakis et al., 2017; West et al., 2015) and even neurochemical and hormonal processes that eventually influence behaviour (Suzuki, 2017; Vuong et al., 2017). The extensively studied gut microbiome is shaped by a

multitude of factors, including host phylogenetic relationships, social interactions, and dietary adaptations (Moeller et al., 2016; Nishida and Ochman, 2018; Youngblut et al., 2019). In comparison to the gut microbiome, the oral microbiome has received far less attention. Yet, it is of particular importance, as it connects the external and the internal environments and is located directly at the entry point to the digestive and the respiratory tracts. The oral microbiome plays an important role in oral diseases, such as caries and periodontitis, and has been implicated in systemic disorders including cardiovascular disease (Nakano et al., 2011), atherosclerosis (Teles and Wang, 2011), Alzheimer's disease (Olsen and Singhrao, 2015), several cancers (Flynn et al., 2016; Gao et al., 2016) and preterm births (Cobb et al., 2017).

Although studies have linked the oral microbiome composition to both dietary habits (Adler et al., 2016; Hyde et al., 2014b; Janiak et al., 2021) and host evolutionary relationships (Ozga and Ottoni, 2021), it is unclear which of the two predominantly drives the evolution of the oral microbiome and how they interact. Early studies suggested that the oral microbiome is strongly heritable and mostly transferred vertically (Corby et al., 2007; Li et al., 2007), with neonate microbiomes resembling those of their mothers (Dominguez-Bello et al., 2010). Moreover, studies in wild animals have indicated that different species harbour distinct oral microbiota (Brealey et al., 2020; Ozga and Ottoni, 2021). Both observations suggest that host evolutionary relationships have a strong effect on oral microbiome structure. Other studies have highlighted the effect of diet on the oral microbiome (Janiak et al., 2021; Wade, 2013), for instance by uncovering differences between young infants and adults of humans and non-human primates, hence providing a potential link with age-related dietary changes (Cephas et al., 2011). However, distinguishing between host evolutionary and ecological factors as well as their contribution to oral microbiome evolution remains difficult, as most previous studies have focused on rather distantly related host species that occupy distinct ecological niches (Boehlke et al., 2020; Li et al., 2013; Smith et al., 2021; Soares-Castro et al., 2019).

In recent years, dental calculus - the calcified form of dental plaque that forms on the teeth of mammals - has emerged as a useful material for the study of oral microbiomes across diverse mammalian species (Brealey et al., 2020; Fellows Yates et al., 2021; Ottoni et al., 2019; Warinner et al., 2014). Since dental plaque undergoes periodic mineralizations, it effectively fossilises on the living host, reducing postmortem contamination from environmental microorganisms (Warinner et al., 2015). It thus preserves a snapshot of oral and respiratory

microbial communities (Warinner et al., 2014), dietary elements (Adler et al., 2013; Sawafuji et al., 2020) and host cells (Mann et al., 2018). Museum collections can be efficiently used for the study of oral microbiomes from wild animals (Brealey et al., 2020; Fellows Yates et al., 2021), minimising the exposure and disturbances associated with sampling from live hosts. Previous research suggests that microbial communities in historical and modern dental calculus are remarkably similar (Velsko et al., 2019), therefore studying museum specimens provides reliable information about present-day microbiomes. The exceptional preservation in the calcified matrix also allows assembly of relatively accurate metagenome-assembled genomes (MAGs) (Brealey et al., 2020).

In this study, we use dental calculus to uncover ecological and evolutionary factors that drive oral microbiome evolution in a group of closely related species. We focus on gorillas, which consist of two species, western gorilla (*Gorilla gorilla*) and eastern gorilla (*G. beringei*), each of which is further divided into two subspecies. Our sampling covers three of the four subspecies: western lowland gorilla (*G. g. gorilla*), Grauer's gorilla (*G. b. graueri*) and mountain gorilla (*G. b. beringei*). Eastern and western gorillas diverged approximately 250,000 years ago, with evidence for substantial gene flow until more recently (McManus et al., 2015; Xue et al., 2015). Western lowland gorillas are found in western equatorial Africa, primarily at low elevations below 500 m above sea level (masl). Their diet consists mainly of various plant parts and fruits (Maisels et al., 2018). The eastern subspecies - mountain and Grauer's gorillas - occur on the eastern side of the Congo basin. They are estimated to have diverged 10,000 years ago (Roy et al., 2014). Grauer's gorillas occupy the largest altitudinal range of all gorillas, ranging from 500 to 2,900 masl (Plumptre et al., 2016). Mountain gorillas are found at even higher elevations, with most of the range of the Virunga Massif population above 2,200 masl, with gorillas ranging up to 3,800 masl. Consequently, the availability of food at different altitudes leads to different dietary profiles among populations and subspecies (Ganas et al., 2004; Michel et al., 2022; van der Hoek et al., 2021), primarily reflecting availability of fruits, which are highly seasonal and become scarce with increasing altitude (Ganas et al., 2004).

The presence of ecological and dietary differences in these closely related gorilla subspecies provides an opportunity to evaluate the effects of host evolutionary relationships and ecological factors on oral microbiomes. We used shotgun metagenomic sequencing of dental calculus samples from museum specimens (collected between 1910 and 1986) to taxonomically and

functionally characterise the oral microbiota, as well as to assemble several bacterial MAGs and identify dietary components from the three gorilla subspecies.

Results

Data pre-processing and confirmation of oral microbial signature

We produced paired-end shotgun sequencing data from the gorilla dental calculus of 26 newly collected samples, and complemented these with an additional 31 dental calculus samples from previously published studies (Brealey et al., 2020; Fellows Yates et al., 2021) (Table S1). Four dental calculus samples were sequenced independently in this study and by Fellows Yates et al. (2021). At a later stage of the analysis, this set of technical duplicates allowed us to assess the effect of the dataset in our taxonomic analysis, but only the sample with the largest number of reads for each sample pair was retained for further analysis.

Raw sequencing data consisted of, on average, 15,533,319 reads (mean value) per sample (ranging from 1,032 to 95,367,058; Table S1) and 105,653 reads per blank (ranging from 19 to 1,132,867; Table S1). After raw sequence pre-processing (which included removal of poly-G tails, adapter and barcode sequences, merging of forward and reverse reads, and quality filtering), *phiX* and gorilla/human sequences were removed by mapping against respective reference genomes and the resulting unmapped reads were used for taxonomic classification (Methods). Samples with low read counts, low proportion of oral taxa, as well as duplicate samples (those processed at two different facilities), were removed (Methods). The final dataset consisted of 46 dental calculus samples (13 western lowland gorillas, 17 mountain gorilla, 16 Grauer's gorilla; Figure S1), each containing 561,978-77,307,443 reads (mean: 11,365,074, SD: 15,577,988; Table S1). We then applied a multi-step procedure, relying on negative controls and museum environmental samples, to remove contaminant taxa (Methods). The final dataset contained a total of 1007 microbial species (in 430 genera), of which 3.4% (n=34) were members of the core hominid oral microbiome (Fellows Yates et al., 2021) and 4.2% (n=42) members of the Human Oral Microbiome Database (HOMD; (Chen et al., 2010); Table S2). Taxa previously reported as oral accounted for ca. 14% of the total microbial abundance.

The final dataset also included six genera that have been listed as common contaminants by Salter *et al.* (2014) and Weyrich *et al.* (2019), but contained known oral taxa (Chen et al., 2010; Fellows Yates et al., 2021), including *Streptococcus oralis* and *Staphylococcus saprophyticus.* Species belonging to these genera were retained after confirming, where possible, that they exhibit typical post-mortem DNA damage patterns (Methods).

Partial mitochondrial genomes from dental calculus allow subspecies identification

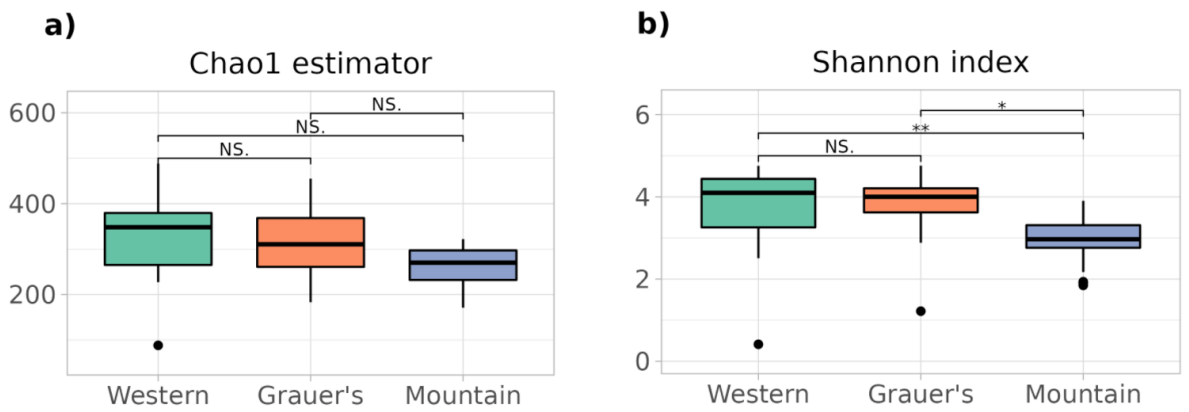
We recovered host mitochondrial sequences from 44 of 46 samples, which we used to confirm gorilla subspecies identity. On average, 31.2% (0.44 - 94.2%; Table S3) of the reference mitochondrial genome was covered by mapped reads, with notable variation in the proportion of genome covered and coverage depth between samples. Sample weight (information available for n=24 samples), sample age (n=36), or dataset (newly generated versus previously published by Fellows Yates et al. (2021)) (n=46) did not have an effect on mtDNA genome completeness (ANOVA, $p>0.54$), whereas host subspecies identity was marginally significant (ANOVA, F-value=3.12, $p=0.054$). Mountain gorillas had a larger proportion of the mtDNA genome covered by mapped reads than other gorilla subspecies. This effect may be due to biological factors specific to the oral environment of this subspecies, or it may be due to a correlation with other factors that were not included in our investigation.

To assist with host taxon assignment, we identified 403 diagnostic sites between western and eastern gorillas and 72 diagnostic sites between the two eastern gorilla subspecies (mountain and Grauer's gorillas) using published gorilla mitochondrial genomes (Methods). In most samples, reads mapped to only a small number of diagnostic sites (Table S3). We accepted molecular taxon assignment for samples with reads mapping to at least six diagnostic sites in six separate reads. For six samples, data was insufficient to distinguish between eastern and western gorilla species and for additional six samples we could not reliably distinguish between mountain and Grauer's gorilla subspecies. In these cases, we assumed that the museum records of host taxon were correct, after consulting collection locality, where possible. All but one sample containing a sufficient number of diagnostic sites could be successfully assigned to their reported subspecies. The exception was sample MTM010 (museum accession: 631168, Swedish Museum of Natural History - NRM), which was confirmed to be a mountain gorilla,

in accordance with the museum records, although previously reported as a western lowland gorilla based on preliminary genomic information (Fellows Yates et al., 2021).

Oral microbiome diversity and composition differ by host subspecies

We evaluated the differences in microbial alpha diversity between gorilla subspecies by measuring species richness and community evenness (Figure 1). Host subspecies had only a marginal effect on the normalised microbial richness (ANOVA, p=0.061; Figure 1a) but a significant effect on the transformed community evenness (ANOVA, p<0.001; Figure 1b), with the microbial communities of mountain gorillas showing the lowest evenness.



Terms	Mean Squares	F-value	p
Read count	54.37	6.05	**0.018**
Dataset	3.90	0.43	0.513
Subspecies	26.83	2.99	0.061
Residuals	8.98		

Terms	Mean Squares	F-value	p
Read count	4181	6.24	**0.016**
Dataset	887	1.325	0.256
Subspecies	5544	8.28	**0.001**
Residuals	670		

Figure 1. Estimates of community richness, using Chao1 index (a) and community evenness, using Shannon index (b), alongside corresponding ANOVA results. Pairwise comparisons in (a) and (b) show FDR-adjusted p-values of the Tukey's test (NS=not significant, *p<0.05, **p<0.01). For the ANOVA models, sqrt(500-x) and exp(x) transformations were implemented on the Chao1 estimator and Shannon index values, respectively.

Microbiome composition differed by gorilla subspecies (Figure 2, Table 1). This effect persisted even after other, potentially confounding factors, such as sequencing depth and dataset, were taken into account. Sequencing depth, included as read count, and dataset were significant factors for the presence-absence-based Jaccard measure, whereas neither of these factors significantly affected relative abundances as estimated with Aitchison distances (Table 1). We confirmed that belonging to different datasets had little to no effect on relative abundance by considering three pairs of samples that were derived from the same museum specimen, but were independently sequenced in this study and in Fellows Yates et al. (2021): the duplicate samples appeared close to each other on ordination plots, in particular on the plot based on Aitchison distances (Figure S2). The fourth sample pair (G0004-IBA002) was excluded, as both replicates had lower than 3% oral component, according to microbial source tracking using FEAST (Shenhav et al., 2019) (Figure S3). Pairwise PERMANOVAs (PERMutational ANalysis Of VAriance) showed strong differentiation between mountain gorillas and the other two subspecies, using both presence-absence and abundance metrics (Table 1). Lastly, we found no significant effect of sex on the oral microbiome using a subset of samples that could be genetically sexed (n=38; Methods; Table S4).

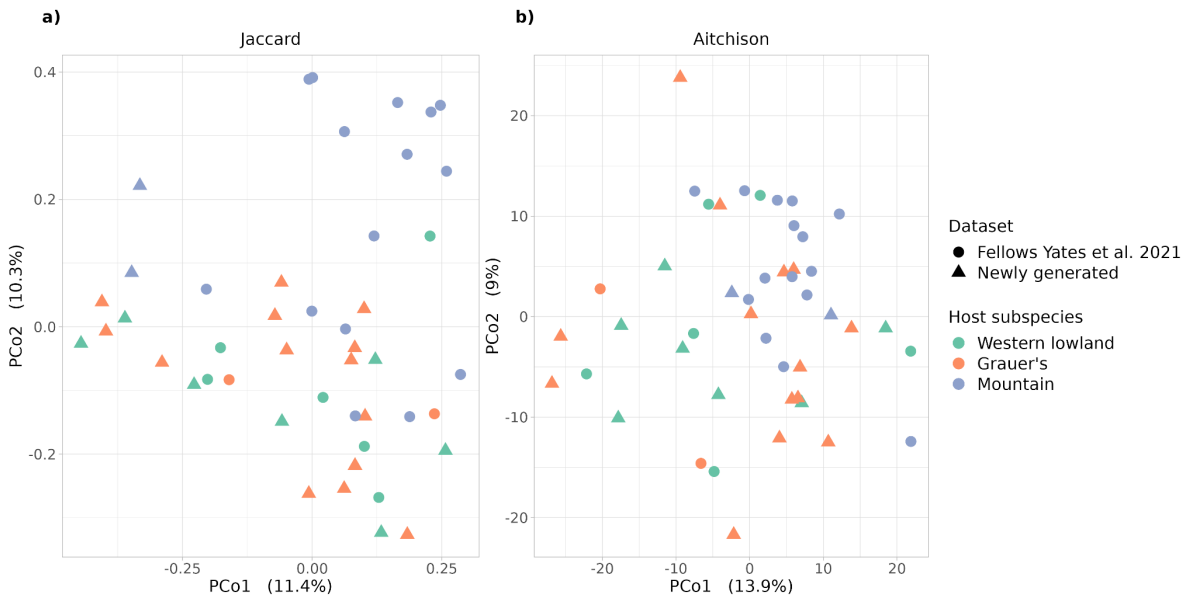


Figure 2. Principal coordinate analysis (PCoA) plots of the individual microbiomes based on a) Jaccard distance and b) Aitchison distance. Host subspecies is displayed in different colours, whereas the dataset is indicated by different shapes (circles for Fellows Yates et al. (2021) and triangles for the newly generated data).

Table 1. PERMANOVA results, showing the effect size (R^2) and p-value of factors putatively contributing to the variance among sampled microbiota, using Jaccard (reflecting presence-absence of taxa) and Aitchison (reflecting relative abundance of taxa) distances. Pairwise PERMANOVA between subspecies show the effect size (R^2) and p-values adjusted for false discovery rate. P-values below 0.05 are shown in bold.

		Jaccard distances		**Aitchison distances**	
Terms	**Within-factor comparison**	**R^2**	**p**	**R^2**	**p**
Read count		0.069	**< 0.001**	0.023	0.297
Dataset		0.037	**0.005**	0.029	0.083
Subspecies		0.067	**< 0.001**	0.074	**0.001**
	Western lowland vs Grauer's	0.029	0.853	0.041	0.223
	Western Lowland vs Mountain	0.086	**0.001**	0.085	**0.001**
	Grauer's vs Mountain	0.073	**0.001**	0.072	**0.001**
Residuals		0.826	-	0.874	-

Among the 1007 microbial (species-level) taxa in our dataset, we detected 91 that differed in abundance among gorilla subspecies and 11 that differed in abundance between the datasets, using ANCOM (Mandal et al., 2015) (Table S5). Ten of these 11 taxa also differed in abundance by subspecies but because they could reflect dataset-specific contamination, we removed the entire genera they belonged to from the list of subspecies-associated taxa. Among the remaining 78 differentially abundant taxa, all but six were absent from at least one of the host subspecies. It is noteworthy that the differentially abundant taxa (n=78) had abundances comparable to those of not differentially abundant taxa (n=929) (t-test: t(123.59)=1.24,

p=0.22): 37 taxa ranked above average and 75 were in the third quartile of abundance ranks, so that their absence from host subspecies is unlikely to be explained by non-detection due to low abundance.

Mountain gorillas, in particular, appeared to be missing many of these differentially abundant microbial taxa (Figure 3) — which is consistent with the general observation of a marginally reduced microbial richness in this subspecies (Figure 1) — even after accounting for the effect of read depth. Microbial taxa that were absent in mountain gorillas included the orders *Rhodobacteriales*, *Pseudonocardiales*, *Corynebacteriales* (represented mainly by *Corynebacterium* and *Mycolicibacterium*), *Bacillales* (including *Staphylococcus*), and *Rhizobiales* (including bacteria associated with the *Fabaceae* rhizosphere, such as *Agrobacterium deltaense, A. fabacearum* (Delamuta et al., 2020; Yan et al., 2017), and three *Rhizobium* species (Poole et al., 2018)). Microbial taxa enriched in mountain gorillas primarily belonged to the orders *Enterobacterales* and certain *Lactobacillales*, like *Streptococcus* sp. and *Lactobacillus gasseri*. The microbiomes of Grauer's gorillas resembled those of western lowland gorillas, both overall (Figure 2, Table 1) and in terms of differentially abundant taxa (Figure 3). However, they also shared some similarities with mountain gorillas (e.g. a presence of *Limosilactobacillus*/*Lactobacillus* species, which were absent in western lowland gorillas), showing an intermediate, or mixed, composition (Figure 3).

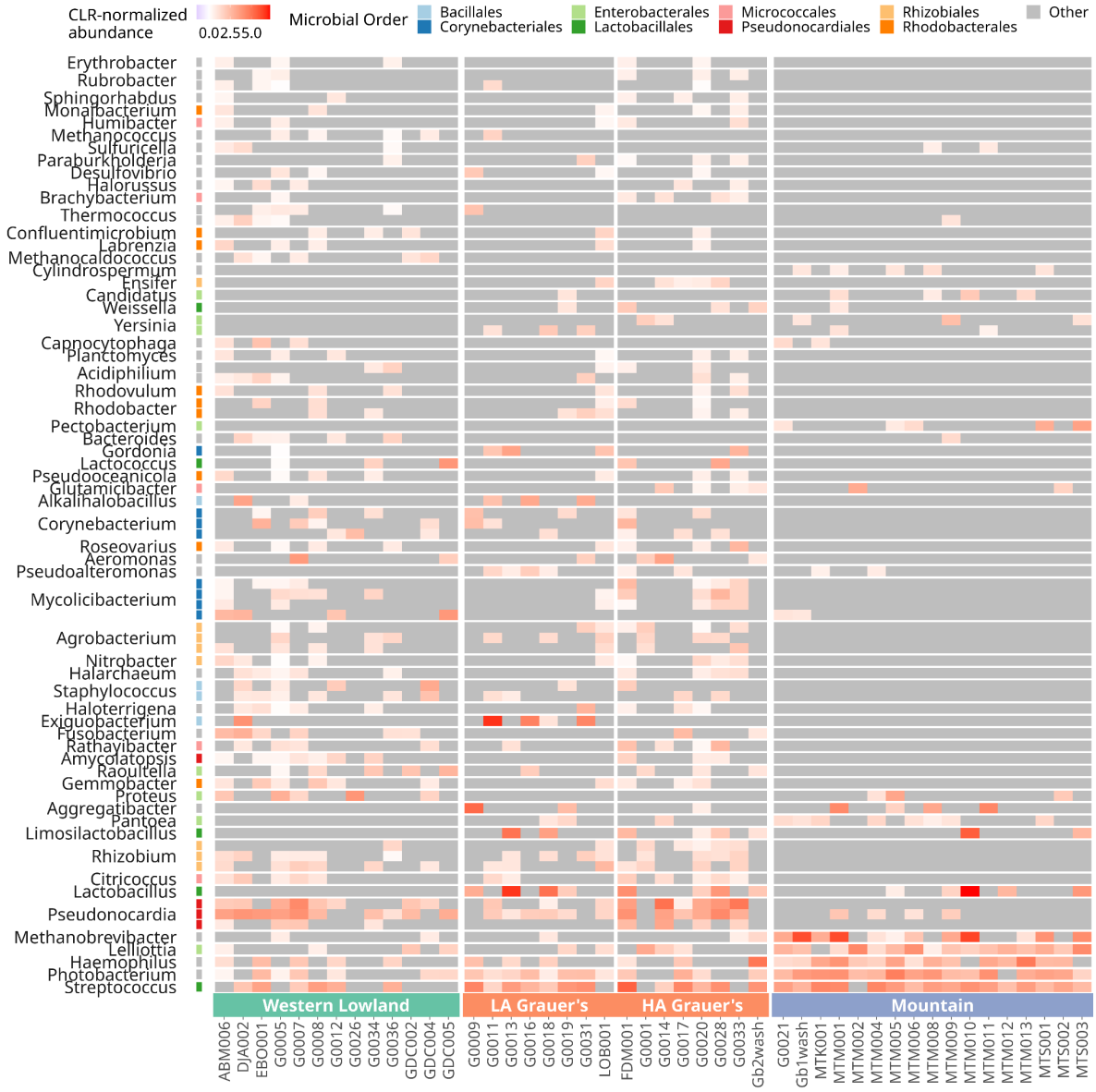


Figure 3. Heatmap depicting centred log-ratio (CLR) normalised abundances of 78 differentially abundant microbial taxa, coloured by taxonomic order (y axis), per sample (x axis). Due to the CLR normalisation, very low abundances or those equal to zero appear as negative values. For clarity, absent taxa (non-normalized abundance equal to zero) are shown in grey.

Gorilla subspecies harbour functionally distinct oral communities

After removing reads mapping to identified contaminants, the retained ("decontaminated") metagenomic reads were functionally characterised using an assembly-free approach using the HUMAnN2 pipeline (Franzosa et al., 2018) (Methods). The normalised abundances of gene families were regrouped under gene ontology (GO) terms, and were aggregated across taxa. Results of functional analyses were broadly consistent with those observed for community-level taxonomy. Specifically, gorilla subspecies was marginally significant (p=0.06) in

explaining functional differences in the oral microbiome, whereas neither sequencing depth nor dataset had a significant effect (Table 2). The functional profiles of mountain gorilla microbiomes differed significantly from the other two subspecies (Table 2).

We used ANCOM to identify differentially abundant biological processes in the dental calculus microbial communities of the gorilla subspecies, requiring that these processes were represented in at least 30% of the samples within each subspecies. Mountain gorillas stood out again, as they missed genes for 236 of the 262 differentially abundant processes (Table S6). In contrast, western lowland and Grauer's gorillas were missing only 13 and 6 processes, respectively. Among six processes that were found in all three gorilla subspecies, mountain gorillas differed significantly from the others in four processes.

Table 2. PERMANOVA results, showing the effect size (R^2) and p-value of factors putatively contributing to the variance in abundance of biological processes among sampled microbiota. Pairwise PERMANOVA between subspecies show the effect size (R^2) and p-values adjusted for false discovery rate. P-values below 0.05 are shown in bold.

Terms	**Within-factor comparison**	**R^2**	**P**
Read count		0.008	0.712
Dataset		0.030	0.220
Subspecies		0.100	0.060
	Western lowland vs Grauer's	0.013	0.695
	Western Lowland vs Mountain	0.124	**0.031**
	Grauer's vs Mountain	0.132	**0.031**
Residuals		0.861	

Metagenome-Assembled Genomes

We applied an iterative assembly and binning approach (Methods) to resolve initial high and medium quality genomic bins using decontaminated sequencing reads from all samples and museum controls. Across all samples, we constructed a total of eight high-quality (>90% completion, <5% contamination) and 27 medium quality MAGs (>50% completion, <10% contamination), belonging to 24 distinct bacterial families (Table S7). Among these 35 MAGs, 25 could be classified to the genus level, with two MAGs assigned to provisional/uncultured genera (UBA8133 and RUG013; Table S7). To place the MAGs within a broader evolutionary context, we constructed a phylogenetic tree including the MAGs and phylogenetically closest bacterial genomes from the Genome Taxonomy Database (GTDB; (Parks et al., 2018); Figure S4). However, for several MAGs assembled in this study we were unable to identify a closely related published genome.

Four MAGs were present in at least 65% of all samples. They were identified as genera *Neisseria* (in 95.7% of samples) and *Rothia* (in 87% of samples), which were found in all Grauer's and mountain gorillas, and as family *Actinomycetaceae* and genus *Lautropia*, both of which were found in all samples of mountain gorillas (Figure S5). Fourteen MAGs reconstructed from dental calculus samples were also present in museum controls, of which eight were present only in the gorilla skin sample and only at low abundances (Figure S5). However, five MAGs were found at higher abundance in at least one of the museum controls compared to any gorilla dental calculus sample (Figure S5). This includes MAGs of *Erysipelothrix* (found in gorilla petrous bone), two different *Planococcaceae* (found in the petrous bone and museum shelf swab), as well as *Propionibacterium* and *Exiguobacterium* (both isolated from museum skull swab). These MAGs belong to genera that are common contaminants in metagenomic studies (i.e. *Propionibacterium*, which is a major genus of common skin bacteria (Salter et al., 2014)) and persistent in environmental reservoirs (i.e. *Erysipelothrix*; (Eriksson et al., 2014)). These five MAGs were excluded from downstream analyses. To further authenticate our reconstructed MAGs, we used a collection of isolation source records from public databases. Taxa with a high proportion (>=25%) of isolation records from sources categorised as environmental/contaminant were considered as potential contaminants (Table S8). Using this approach, we identified 10 MAGs that represent likely contaminants (Figures S5& S6).

A total of 20 MAGs recovered here are members of the oral microbiome and are present in databases of common oral taxa (Chen et al., 2010; Fellows Yates et al., 2021). This includes taxa commonly associated with dental plaque communities: *Rothia* (Tsuzukibashi et al., 2017), *Olsenella* (Abusleme et al., 2013; Socransky et al., 1998), *Corynebacterium* (Welch et al., 2016), *Lautropia* (Gerner-Smidt et al., 1994), *Neisseria (Donati et al., 2016)* and *Actinomyces* (Kolenbrander et al., 2010). *Rothia* species are particularly abundant in gorillas compared to humans and other non-human primates (Fellows Yates et al., 2021).

For two MAGs of high completeness and high prevalence in our samples, we produced an alignment of core genes to further investigate how they relate to the known diversity of these taxa. The *Neisseria* MAG recovered in this study was present in almost all dental calculus samples (Figure S5), but was significantly more abundant in mountain gorillas compared to the other host subspecies (Figure 5b). It clustered with a subset of *Neisseria* species isolated exclusively from humans and was more divergent from *Neisseria* species that formed a clade with isolates from other animals (Figure 5a). Its phylogenetic placement suggests a distinct and not yet identified *Neisseria* taxon. We also recovered a near-complete MAG of *Limosilactobacillus gorillae* (Figure S7). This species has previously been isolated from faeces of mountain and western lowland gorillas (Tsuchida et al., 2018, 2015). The dental calculus *L. gorillae* MAG was sister to the gorilla faecal isolate but distinct from faecal isolates of other primates (Figure S7a).

Rothia, *Neisseria*, *Veillonella*, *Corynebacterium*, *Actinomyces*, *Selenomonas*, *Propionibacterium*, *Fusobacterium*, and *Eikenella* are all common oral bacteria that reduce nitrate in the oral cavity (Rosier et al., 2022) and all of them were present in gorilla dental calculus communities studies here (Figure S5, Table S5). Using MAGs of *Rothia*, *Neisseria*, and *Veillonella* we queried if gorilla oral bacteria contain genes involved in nitrogen metabolism, and specifically nitrate reduction. For *Neisseria*, we could confirm the presence of five nitrate reducates (*nar*), seven genes that provide the cofactor molybdenum to the nitrate reductase enzymes, and four additional genes involved in the regulation of nitrate reduction (*aniA*, *nirB*, *nirD*, *norB*; Figure 5a). Nitric oxide reductase (*norB*) and nitrite reductase (*aniA*) genes are highly conserved in *Neisseria* (Barth et al., 2009) and were consistently found in all *Neisseria* genomes that we analysed (Figure 5a). Only two genes, *narY* and *mopII*, were not identified in the gorilla dental calculus MAG. However, these genes are missing from other

nitrate-reducing oral bacteria (Rosier et al., 2020), thus suggesting that *Neisseria* isolated from gorilla dental calculus likely has the full capacity to reduce nitrate. *Rothia* and *Veillonella* possessed a smaller repertoire of nitrate-reducing genes. The *Rothia* MAG recovered in this study had several nitrate reductases and molybdenum cofactor genes, but they differed from genes present in *R. dentocariosa*, a bacterium well studies for its nitrate-reducing capacity (Figure S8; Rosier et al., 2020). We were also able to identify nine genes involved in nitrate metabolism in *Veillonella*, a species well known as a reducer of nitrate in the oral cavity (Hyde et al., 2014a) (Figure S8). Similar to *Neisseria*, *Veillonella* was significantly more abundant in mountain gorillas than in the other two gorilla subspecies, whereas the abundance of *Rothia* did not differ by gorilla subspecies (Figure S8b and d).

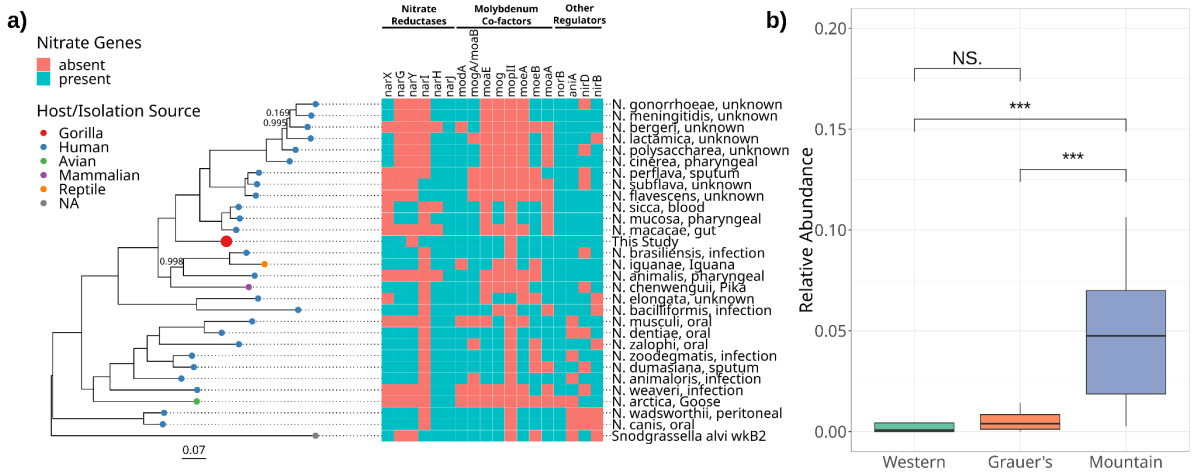


Figure 5: a) Maximum-likelihood phylogeny based on the alignment of core gene sequences from the *Neisseria* MAG and additional published genomes of the same genus. Tip points are coloured by host species identity, and tip labels include the isolation source for each published genome. Scale bar units are the number of substitutions per site and node values represent proportion of node support out of 100 bootstrap replicates, with unlabeled nodes having support of 1. To the right, the presence of key genes involved in the reduction of nitrate in the oral cavity is shown for each *Neisseria* species. **b)** Relative abundances (CLR transformed) of the *Neisseria* MAG in different gorilla subspecies. The results of a Wilcoxon test are denoted by brackets above boxplots (NS = not significant; *** = $p > 0.001$).

Altitude may drive oral microbiome composition

Both taxonomic and functional analyses suggest that Grauer's gorilla oral microbiomes are more similar to western lowland gorillas than to mountain gorillas, with which they share a close evolutionary relationship and an adjacent geographic range. However, the distribution ranges of all three gorilla subspecies differ in elevation, with mountain gorillas occurring at the

highest altitudes. As altitude can influence temperature, humidity and food diversity, which in turn can influence microbial communities, we performed partial Mantel tests between altitudinal distances and taxonomic (Jaccard/Aitchison distances) or functional (Euclidean distance) composition of the oral microbiome, while accounting for log-transformed geographical distance. Geographic location and altitude of the specimens were approximated based on museum records. However, when considering the entire dataset, we did not detect a correlation between altitude and neither taxonomy nor function (Mantel test: R ranging from -0.13 to 0.06, p>0.12).

Since Grauer's gorillas occupy the widest altitudinal range of all gorilla subspecies, we divided Grauer's gorilla samples into high- (>1000 masl, n=8) and low-altitude (≤1000 masl, n=8) groups and tested for differentiation of the oral microbiome both between these groups and among host subspecies. While no compositional differences were identified between the oral microbiomes of western lowland and low altitude Grauer's gorillas, using relative abundance measures western lowland and high altitude Grauer's gorillas displayed marginally significant differences based on pairwise PERMANOVA (FDR-adjusted p-value=0.075, Table 3). The oral microbiome of mountain gorillas consistently differed from other subspecies, independent of altitude grouping (Table 3). Yet, high-altitude Grauer's gorillas showed considerable overlap with mountain gorillas in the first two axes of a PCoA based on Aitchison distances (Figure S9).

Table 3. Results of the pairwise PERMANOVA showing the effect size (R^2) of comparisons of taxonomic composition based on Aitchison distances among four gorilla groups (western lowland, low-latitude Grauer's [≤1000 masl], high-altitude Grauer's [>1000 masl], and mountain gorillas) and the associated p-values (adjusted for false discovery rate). P-values below 0.05 are shown in bold.

Comparison	**R^2**	**p-value (FDR-adjusted)**
Western Lowland vs Low-altitude Grauer's	0.048	0.515
Western Lowland vs High-altitude Grauer's	0.070	0.075

Western Lowland vs Mountain	0.083	**0.002**
Low-altitude Grauer's vs High-altitude Grauer's	0.084	0.164
Low-altitude Grauer's vs Mountain	0.098	**0.002**
High-altitude Grauer's vs Mountain	0.087	**0.002**

Dental calculus reflects dietary differences of gorilla subspecies

After removing reads assigned to bacterial and viral taxa, we performed taxonomic classification of the eukaryotic diversity present within our samples to characterise gorilla dietary components using Kraken2 with the NCBI 'nt' database. Despite applying extensive decontamination procedures (Methods) and removing taxa with few supporting reads (<10), we still recovered erroneous assignments, such as mollusks or other mammals, which gorillas are unlikely to consume. We therefore restricted our analyses to eukaryotic families that are known to be part of the gorilla diet (Michel et al., 2022; Remis et al., 2001; Rogers et al., 2004; Rothman et al., 2014; Yamagiwa et al., 2005); Methods). The majority of the 371 genera (65 families) detected in our dataset (Figure S10) were plants (n=360), but we also identified insects, specifically ants, (N=7) and lichen-forming fungi of the *Parmeliaceae* family (N=6).

We detected 22 genera that differed in abundance across gorilla subspecies, all belonging to plants, except for the fungal genus *Parmotrema* (Figure 6, Table S9). The broad dietary patterns observed here agree with literature reports. For instance, bamboo - represented here by the African genus *Oldeania*, as well as Asian genera *Phyllostachys*, *Bambusa,* and *Ferrocalamus*, which are likely misclassifications of African bamboo species (Brealey et al., 2020) - was mainly detected in mountain and Grauer's gorillas, as expected (Rothman et al., 2014; Yamagiwa et al., 2005); Figure 6). *Marantochloa* and *Thaumatococcus* of the *Marantaceae* family were detected in western lowland gorillas, which are known to consume them (Rogers et al., 2004), but also in Grauer's gorillas, where the family *Marantaceae* has been reported as part of the diet (Michel et al., 2022). Grauer's gorillas appear to have the most diverse diet of

the three subspecies, consuming foodstuffs typical of both western lowland and mountain gorillas (Figure 6). This is in accordance with the literature (Michel et al., 2022; Remis et al., 2001; Rogers et al., 2004; Rothman et al., 2014; Yamagiwa et al., 2005), which reports 61 taxonomic families as being part of their diet, as opposed to 41 and 45 families for western lowland and mountain gorillas, respectively.

Figure 6. Heatmap based on normalised abundances for differentially abundant dietary taxa. For clarity, taxa with abundance equal to 0 are shown in grey. The bars on the right indicate if the family (light green) or the specific genus (dark green) have been reported as part of the diet of each gorilla subspecies (grey: no mention). Horizontal white lines separate plant families.

Discussion

Using three gorilla subspecies we investigated how evolutionary relationships and ecological factors shape the oral microbiome in closely related host taxa. In our study, Grauer's and mountain gorillas are geographically proximate and more closely related to each other than to western lowland gorillas. Ecologically, all three subspecies differ from each other. One of the strongest differences is the altitude of the respective subspecies ranges, which in turn influences diet, reflected e.g. in the level of frugivory. Western lowland gorillas occupy the lowest altitudes and are the most frugivorous (Rogers et al., 2004; Takenoshita and Yamagiwa, 2008). Mountain gorillas occupy the highest altitude and are the least frugivorous (Ganas et al., 2004; Maisels et al., 2018; Plumptre et al., 2016), whereas Grauer's gorillas live across the largest altitudinal range and hence straddle the ecological conditions faced by both western lowland and mountain gorillas (Yamagiwa et al., 2008). Ecological conditions and dietary differences are expected to affect the host-associated microbial communities (Janiak et al., 2021; Kartzinel et al., 2019; Li et al., 2013; Wade, 2013; Youngblut et al., 2019).

By investigating the oral microbiome preserved in dental calculus, we show that microbial communities among closely related gorilla subspecies differ, both taxonomically and functionally. In particular, the oral microbiome of mountain gorillas is significantly different from the other two subspecies in taxonomic composition, relative abundance, and functional profiles of microbial taxa (Tables 1 and 2), despite the close evolutionary relationship and geographic proximity between mountain and Grauer's gorillas. The oral microbiome of Grauer's gorillas shows strong similarity to western lowland gorillas, despite the two species being more distantly related and separated by the Congo Basin. When considering differentially abundant microbial taxa and dietary profiles, Grauer's gorillas appear intermediate, displaying a combination of taxa found in both western lowland and mountain gorillas (Figure 3 and 6). This suggests that host phylogenetic relationships are less important for the structure of the oral microbiome in closely related species than ecological conditions, possibly including diet.

These observations deviate from previous studies that identified a phylogenetic signal in species-specific patterns of the oral microbiome in wild animals. However, these studies compared distantly related species like chimpanzees, humans (Li et al., 2013), other great apes (Boehlke et al., 2020), and even more diverse groups of mammals (Brealey et al., 2020; Ozga

and Ottoni, 2021). Often the effects of host phylogenetic relationships and ecology could not be decoupled. For instance, Smith *et al.* (2021) found differences among three snake species that each occupied a distinct ecological niche. Similarly, Soares-Castro et al. (2019) studied three marine mammals, among which the most evolutionary divergent species also lived in a different habitat, so the effects of ecology were indistinguishable from those of evolutionary relationships.

Yet, ecological effects on a short evolutionary scale, as observed in this study, do not preclude co-diversification of the host and its associated microbiome over longer evolutionary time periods. They rather point to the possibility of rapid and dynamic responses of the host-associated microbial communities to environmental conditions and/or the possibility of seeding of the dental calculus oral microbiome at least to some extent with environmental taxa. In addition, our results are in accordance with other studies that propose external factors, such as diet, influencing the oral microbiome (Adler et al., 2016; Hyde et al., 2014b; Janiak et al., 2021), including Li et al. (2013) who observed that the oral microbiome of captive apes is distinct from their wild counterparts and more similar to other captive apes, highlighting the importance of environmental factors/diet.

Oral microbiome diversity might reflect ecology or diet

Our analyses provide a tentative suggestion that ecological differences, approximated here by altitude, may be the primary factor shaping oral microbiome diversity. This is further exemplified by comparisons within the host species. Specifically, grouping Grauer's gorillas into low- and high-altitude populations, we showed that low-altitude populations harbour an oral microbiome that is indistinguishable from western lowland gorillas, whereas high-altitude populations show near-significant differences (Table 3). Although the oral microbiome of Grauer's gorillas is significantly different from mountain gorillas independent of the altitude, this may be because most mountain gorillas in our dataset were collected above 3000m, notably higher than any of the Grauer's gorilla samples (600-2100m; Table S1). Furthermore, the dietary profiles of Grauer's gorillas appear to contain dietary items present in both western lowland and mountain gorillas (Figures 6 and S10). Our observation from dental calculus is in line with previous studies (Table S10), which report considerable overlap in the diet of Grauer's

gorillas with western lowland (30 of the 41 plant families) and mountain (30 of the 45 plant families) gorillas. In contrast, only 16 dietary families are consumed by both western lowland and mountain gorillas.

It is also intriguing to speculate that the local environment itself could affect the composition of the dental calculus community by serving as a source of at least some colonising taxa, for example via the consumed phyllosphere and rhizosphere. A recent study of the human oral microbiome found that specific taxa within the dental plaque community are evolutionarily close to environmental bacteria, whereas those living on the tongue surface are more closely related to other host-associated taxa (Shaiber et al., 2020). This finding suggested that dental plaque microbial communities could be more affected by the environment than other host-associated microbiomes. Within the gorilla oral microbiome, we detected microbial taxa that likely have a dietary origin and differ in abundance among host subspecies, possibly reflecting dietary differences. In particular, bacteria associated with the *Fabaceae* rhizosphere, including *Agrobacterium deltaense, A. fabacearum* (Delamuta et al., 2020; Yan et al., 2017), and three *Rhizobium* species (Poole et al., 2018) may be derived from the consumed dietary items or accidentally ingested soil. These bacteria were significantly more abundant in the dental calculus of western lowland and Grauer's gorillas than in mountain gorillas (Figure 3, Table S5). Two lines of evidence provide a link between these microbial taxa and dietary differences among the gorilla subspecies. First, our data suggest a higher abundance of the family *Fabaceae* and specifically the genus *Tamarindus* in western lowland and Grauer's gorillas (Figures 6 and S10, Table S9) compared to mountain gorillas. Second, although no comparative data exists on the prevalence of roots and rhizomes in the diet of different gorilla species, indirect evidence suggests that western lowland gorillas consume more roots than mountain gorillas. Behavioural observations report frequent root and rhizome consumption in western lowland gorillas of all ages (Fletcher and Nowell, 2008) and comparative dentition analyses show reduced tooth wear in mountain gorillas compared to western lowland gorillas, whereas tooth wear in mountain gorillas is well explained by the time spent feeding on roots (Galbany et al., 2016).

Metagenome-assembled genomes add to the understanding of wild microbiomes

The recovery of high quality MAGs from metagenomic data can provide an important perspective on the evolutionary relationships among microbial taxa and between microorganisms and their hosts. As reference databases are often biassed towards organisms of medical, agricultural and industrial importance (Marcelino et al., 2020; Velsko et al., 2018), MAGs recovered from novel environments, such as dental calculus oral microbiome of non-human animals, may represent poorly described, understudied or completely unknown lineages of microbial diversity (Jiao et al., 2021). From the collection of high and medium quality MAGs which were constructed in this study, several were recovered in higher abundance from environmental controls than from any dental calculus samples and were considered to represent environmental contaminants. However, other MAGs were highly abundant in the dental calculus samples and have previously been reported as members of the oral microbiome (Figures S5 & 6). Many of these MAGs were distinct from the evolutionary closest sequenced genomes available in GTDB (Figures S4, S7, S8, Figure 5). This suggests that the MAGs in this study may represent undescribed bacterial lineages, in line with recent studies that reported undescribed microbial lineages in unstudies host species (Levin et al., 2021; Youngblut et al., 2020). The closest evolutionary relatives of several bacterial taxa with high quality MAGs were found among isolates from primates, including humans (Figures 5a, S7, S8).

In this study, we recovered a near-complete MAG of *Limosilactobacillus gorillae*, which was more closely related to a faecal isolate from a captive gorilla than to a faecal isolate from another primate (Figure S7a). Usually associated with gorilla faeces (Tsuchida et al., 2014), the presence of *L. gorillae* within the oral microbiome may be the result of coprophagy, a common behaviour in wild gorillas (Graczyk and Cranfield, 2003). However, questions remain regarding the persistence of faecal-associated bacteria within dental calculus and their potential function within the plaque biofilm. Coprophagic behaviour is suggested to serve as a route for vertical or horizontal transmission of gastrointestinal microbiota between individuals (Abusleme et al., 2020) and may have a stabilising impact on host-associated microbial communities overall (Bo et al., 2020).

We also reconstructed MAGs of *Neisseria*, *Rothia*, and *Veillonella* and confirmed the presence of nitrate-reducing genes in these members of the gorilla oral microbiome (Figures 5a, S8). The reduction of nitrate to nitrite from dietary sources by oral bacteria provides the precursor to nitric oxide (NO), which is an important signalling and effector molecule (Jones et al., 2021). Along with multiple benefits to circulatory health (Lundberg et al., 2018), nitric oxide is increased in response to hypoxic stress (Feelisch, 2018; Levett et al., 2011), and results in increased blood oxygen levels in the host (Beall et al., 2012). Human populations adapted to high altitude exhale higher concentrations of nitric oxide compared to low-land populations (Beall et al., 2001; Erzurum et al., 2007). For these reasons, regulation of nitric oxide metabolism is thought to be beneficial to high-altitude adaptation (Beall et al., 2012).

Mountain gorillas live at elevation of up to 3,800 masl (Williamson et al., 2013) and hence are exposed to hypoxic stress (Ruff et al., 2022). They also show higher abundance of *Neisseria* and *Veillonella* compared to the other two gorilla subspecies, which live at considerably lower altitudes (Figures 5b and S8d). Although it is not known if dietary nitrate amounts differ among gorilla subspecies, the primarily herbivorous diet of mountain gorillas is naturally rich in nitrate. Taken together, it is likely that high abundance of nitrate-reducing oral taxa may aid mountain gorillas cope with the physiological demands of their high-altitude habitat.

Mitochondrial DNA in dental calculus can aid taxonomic assignments of the host

The mitochondrial sequences recovered from the dental calculus metagenomes were used to confirm gorilla subspecies identity. Six of the samples yielded mitochondrial genomes with completeness >80% and coverage >3X (maximum 94% completeness and 77X coverage, Table S3). However, even a few mitochondrial fragments were sufficient to allow subspecies assignment. We could genetically distinguish gorilla species (eastern versus western gorillas) with mitochondrial genomes coverage as low as 2.6% and eastern gorilla subspecies (mountain versus Grauer's) with coverage as low as 6.6%. Grauer's and mountain gorilla mitochondrial genomes could be successfully distinguished despite differing by only 0.5% (72 positions) of the mitochondrial genome (excluding the hypervariable D-loop region; (van der Valk et al., 2018)). This observation supports the usefulness of dental calculus as material for obtaining

genetic information about the host (Mann et al., 2018; Ozga et al., 2016; Warinner, 2016; Warinner et al., 2015).

Interpreting the results of metagenomic analyses has caveats

Dental calculus is a highly suitable material for the study of the oral microbiome of long-deceased hosts. However, contamination and reference database biases are considerable challenges in the field of ancient metagenomics, including the analysis of dental calculus. To tackle this issue, we implemented a stringent multi-layered protocol, which included both wet-lab procedures and bioinformatic methods to identify and remove likely contaminants (see Methods). We also explicitly modelled the effect of the dataset (Uppsala University, Sweden and the Max Planck Institute for the Science of Human History, Germany) in all our analyses to account for a potential batch effect. Nonetheless, erroneous taxonomic identification could originate from sources other than contamination. Notably, biases related to reduced retention of ultrashort AT-rich fragments (Mann et al., 2018), or the severe underrepresentation on microbial communities from non-model host species in databases can lead to undescribed taxa being undetected or misassigned to close relatives. However, within our dataset, these sources of error are equally likely for all studied gorilla subspecies and hence should not affect performed comparisons.

Furthermore, several previous studies have identified caveats associated with dietary analyses from dental calculus samples (Haas et al., 2022; Mann et al., 2020; Ottoni et al., 2019). Eukaryotic taxonomic classifications tend to contain many false positive assignments due to reference genomes being contaminated with microbial DNA, among other reasons (Mann et al., 2020). In our study, the authenticity of putatively dietary taxa could not be assessed using damage patterns due to sparsity of eukaryotic reads. Therefore, interpreting these results was only possible by relying on previous knowledge of gorilla diets (Michel et al., 2022; Remis et al., 2001; Rogers et al., 2004; Rothman et al., 2014; Yamagiwa et al., 2005). Nevertheless, we could confirm dietary differences among gorilla subspecies (Figure 6), which are mostly in agreement with the published literature.

Conclusions

Our study is, to our knowledge, the first to investigate the structure of the dental calculus microbiome early in the speciation process of the host, and one of the largest metagenomic dental calculus datasets available for wild animals to date. It adds to a new but growing field of research on the microbiomes of wild animals, which remain severely understudied. We find that ecology is more important than evolutionary relationships in explaining taxonomy and function of the oral microbiome in closely related species. Our findings indicate that the oral microbiome may contribute to host adaptation by allowing the host to occupy a specific ecological niche. Specifically, taxa enriched in mountain gorilla oral microbiome may contribute to their high-altitude lifestyle through increased nitrate reduction potential and the associated physiological benefits. Our finding of distinct phylo- and rhizosphere taxa in the dental calculus of gorillas suggests a close connection between this host-associated microbiome and the local environment. This may open opportunities for the colonisation of the host with environmental taxa with functions in host local adaptation, health and disease - an exciting evolutionary route that could be explored in the future.

Methods

Sample collection

The study dataset consisted of dental calculus samples from 57 specimens of three gorilla subspecies: 16 western lowlands gorillas, 22 Grauer's gorillas, and 19 mountain gorillas (Table S1). Dental calculus samples of western lowland gorillas were collected at the Royal Museum for Central Africa (RMCA, Tervuren, Belgium) and the Cleveland Museum of Natural History (CMNH, USA). Samples of mountain gorillas came from the Swedish Museum of Natural History (Naturhistoriska Riksmuseet - NRM, Stockholm, Sweden) and the RMCA. Samples of Grauer's gorillas were collected primarily at RMCA, with a few samples from NRM and the Royal Belgian Institute of Natural Sciences (RBINS, Brussels, Belgium). Our dataset consisted of newly generated shotgun data from 26 specimens and published gorilla dental calculus sequences: two samples from Brealey et al. (2020) and 29 samples from Fellows Yates et al. (2021). We also included sequences from 35 published and newly generated extraction blanks, twelve library preparation blanks, and four museum controls. One extraction and one library blank had too few reads for taxonomic classification and were therefore excluded from downstream analysis (Table S1). The museum controls consisted of two swabs taken at NRM from a museum shelf surface holding reindeer (*Rangifer tarandus*) specimens and from the surface of a brown bear (*Ursus arctos*) skull. In addition, we included data from two gorilla specimens from previous studies (van der Valk et al., 2019, 2017): A petrous bone sample from a specimen at NRM and a skin sample from RMCA (ENA accession numbers: ERR2503700 and ERR2868193, respectively). These sequences were generated to study the gorilla host, however, we removed host reads and retained microbial reads for our analyses.

Preparation of Genomic Libraries and Metagenomic Shotgun Sequencing

All samples were processed in cleanroom facilities following appropriate methods for working with ancient and historical DNA. DNA was extracted according to the protocol by Dabney et al. (2013) with slight changes, as described in Brealey et al. (2020) and Fellows Yates et al. (2021). The datasets differed in library preparation protocol, indexing strategy, sequencing platform, read length, and sequencing depth (Table S1). Samples from four specimens were processed in both laboratory facilities: Uppsala University, Uppsala, Sweden and the Max Planck Institute for the Science of Human History, Jena, Germany (Fellows Yates et al., 2021).

We used these technical duplicates to assess putative batch effects (Fig. S3), but only retained one sample per pair for biological analyses, selecting the one with the largest number of reads.

For the newly generated data, we followed the protocol as detailed in Brealey et al. (2020). Briefly, dental calculus samples ranging in weight from < 5 mg and up to 20 mg were surface-decontaminated using UV light (10 min at 254 nm) and washing in 500 µl of 0.5M ethylenediaminetetraacetate (EDTA) for 1 min (Brealey et al., 2021, 2020; Ozga et al., 2016). DNA was extracted using a silica-based method (Dabney et al., 2013) in batches of at most 16 samples with two negative controls. DNA was eluted in 45 µl of EB buffer (10 mM tris-hydrochloride, pH 8.0; QIAGEN, Netherlands) supplemented with 0.05% (v/v) Tween-20.

Double-stranded genomic libraries for all newly generated and previously sequenced samples were prepared following the double indexing protocol (Dabney and Meyer, 2012; Meyer and Kircher, 2010). Newly generated libraries and those previously published by (Brealey et al., 2020) included double in-line barcodes to guard against index hopping (van der Valk et al., 2020). A detailed protocol for library preparation is provided in Brealey et al. (2020). Library blanks were included for each batch of approximately 20 samples. Adapter-ligated libraries were quantified by a quantitative PCR assay, allowing us to estimate an appropriate number of index PCR cycles, which ranged from 8 to 20. Following indexing, individual libraries were purified with Qiagen MinElute columns and quantified using a quantitative PCR assay. All extraction and library blanks consistently showed lower DNA content than calculus samples (Table S1). We pooled 1.5 µl of each indexed library (including blanks and museum controls) and performed size selection with AMPure XP beads (Beckman Coulter, IN, USA) for fragments of approximately 100-500 bp in length. The pooled library was sequenced by SciLifeLab Uppsala on two Illumina NovaSeq S2 flowcells using paired-end 100 bp read length and V1 sequencing chemistry.

Preprocessing of sequencing data

For each sample, the data generated from different sequencing runs was concatenated into a single file (separately for reverse and forward reads) and poly-G tails, which are artefacts of the sequencing chemistry of the NextSeq and NovaSeq Illumina platforms, were removed using fastp (V0.20.0; (Chen et al., 2018). We removed unpaired reads with BBTools `repair.sh` (V38.61b; (Bushnell, 2014). For newly generated data, we used AdapterRemoval (V2.2.2;

(Schubert et al., 2016) to clip adapters, trim reads based on minimum phred quality (>=30) and length (30bp), and merge forward and reverse reads. The unmerged reads were excluded from downstream analysis (Table S1). In-line barcodes in newly generated data were trimmed from both ends using a Python script (Brealey et al., 2020). Read quality filtering was performed using PrinSeq-Lite (V0.20.4; (Schmieder and Edwards, 2011)) with a mean base quality threshold of 30. PCR duplicates were removed using a Python script (Brealey et al., 2020), which randomly kept one read among those with the same sequence. All commands and scripts used in our bioinformatic analysis are available at https://github.com/markella-moraitou/Gorilla_dental_calculus.

Sequences of the *phi*X bacteriophage, which is used as an internal control for Illumina sequencing (Mukherjee et al., 2015), were removed from the dataset by mapping against the *phiX* genome (GenBank: GCA_000819615.1) using BWA-MEM (0.7.17; (Li and Durbin, 2009). Unmapped reads were retained using SAMtools (V1.12; (Danecek et al., 2021) and converted to FASTQ for downstream analysis, using the 'bamtofastq' function of BEDtools (V2.29.2; (Quinlan and Hall, 2010). To remove host reads and potential human contamination, we mapped all the reads to a combined file containing western lowland gorilla and human reference genomes (GenBank: GCF_000151905.2 and GCF_000001405.38, respectively), repeating the steps above. Extraction blanks, library blanks, and museum controls were only mapped to the human genome. The unmapped reads were used in oral microbiome and dietary analyses, whereas the mapped reads were extracted, sorted, and indexed using SAMtools, and used for the host genomic analyses.

Host genome analysis

Mitochondrial genome analysis. From the host mapped BAM files, we extracted all reads mapping to the gorilla mitochondrial genome from position 1 to position 15,446, with a mapping quality of 30 or higher. Nucleotide positions 15,447-16,364, which contain the D-loop, were excluded to prevent the hypervariable regions from introducing biases in our analyses. We calculated sample coverage and the total number of reads mapping to the mitochondrial genome with SAMtools and generated a consensus sequence in FASTA format with ANGSD (V0.933; (Korneliussen et al., 2014).

Subspecies assignment verification. The subspecies identity of each sample was initially assigned using museum records. We confirmed these assignments using diagnostic sites within the host mitochondrial genome. Using 102 published mitochondrial genomes (Das et al., 2014; Hallast et al., 2016; Hu and Gao, 2016; van der Valk et al., 2018; Xu and Arnason, 1996); Table S11), we identified two sets of diagnostic sites that were fixed for different alleles in each of the contrasted groups. The first set was species-specific, identifying membership in western versus eastern gorillas. The second set was subspecies-specific, distinguishing between mountain and Grauer's gorillas.

Molecular sexing of samples. We used sexassign (Gower et al., 2019) to determine the sex of each individual (Table S1). To this end, we mapped filtered reads, prior to removing host and human reads, to the gorilla reference genome (GCF_000151905.2) and counted reads mapping to the X chromosome and the autosomes. Sexassign conducts a likelihood ratio test comparing the observed X-to-autosome ratio to expected ratios for males (R_X ~0.5) and females ($0.8 > R_X < 1.2$).

Initial taxonomic classification and decontamination

Taxonomic classification with Kraken2/Bracken. After removing host and human reads, the retained unmapped reads were taxonomically classified using Kraken2 (V2.1.1; (Wood et al., 2019). We used the standard Kraken2 database (accessed by the software on 1st September 2021), which includes all bacterial, archaeal and viral genomes from NCBI. Abundances were re-estimated at the species level using Bracken (V2.6.2; (Lu et al., 2017), by redistributing reads from higher taxonomic levels to the species level. The average read length was set to 55 bp, which corresponded to the average of the mean read length of our data. The raw species table and the corresponding metadata were combined into a phyloseq object (V1.34.0; (McMurdie and Holmes, 2013) in R (R Core Team, 2015). Taxonomic information was assigned based on taxonomic IDs from the NCBI taxonomy database using the 'classification' function of the taxize package (V0.9.99; (Chamberlain and Szöcs, 2013).

Removal of low-quality samples and contaminant taxa. Samples containing less than 300,000 processed reads were excluded from downstream analysis (for the distribution of read counts per sample see Figure S11). We then evaluated the composition of the microbial communities of the retained samples with FEAST (V0.1.0; (Shenhav et al., 2019). FEAST makes use of

user-provided reference microbiomes to partition the composition of the study samples and to estimate the proportional contribution of the reference "sources". We provided 29 such sources sampled from six different environments that were chosen to reflect both the oral cavity and potential contaminants. Specifically, we included human calculus (n=5 (Mann et al., 2018)) and human plaque (n=5 (Human Microbiome Project Consortium, 2012; Lloyd-Price et al., 2017)) as oral sources, and human gut (n=5 (Human Microbiome Project Consortium, 2012; Lloyd-Price et al., 2017)), human skin (n=5 (Oh et al., 2014)), tundra soil (n=5 (Johnston et al., 2016)) and laboratory contaminants excluding human sequences (n=4 (Salter et al., 2014)) as potential contaminants (Table S12). Gorilla dental calculus samples for which the proportion of the metagenomic community similar to the human oral microbiome (human calculus and human plaque considered jointly) was lower than 3%, according to the output of FEAST, were excluded from further analyses. The threshold was decided based on the distribution of the oral proportions in the study samples (Figure S3).

For the retained samples, we then carried out a multi-step approach to remove putative contaminant microbial taxa. First, we used the R package decontam (V1.10.0; (Davis et al., 2018) with the 'isContaminant' function (method="combined", normalize = TRUE). The package employs two main approaches: a prevalence-based approach, which identifies contaminants based on their increased prevalence in blanks, and a frequency-based approach, which relies on the observation of a reverse relationship between taxon abundance and DNA quantity in the sample. We used a combination of both decontamination approaches, performed separately for the subset of newly generated data, which also included two previously published samples processed in the same facility (Uppsala University, Sweden (Brealey et al., 2020)), and data from gorilla dental calculus published by Fellows Yates et al. (2021) and processed at the Max Planck Institute for the Science of Human History, Germany. We refer to these datasets as newly generated and Fellows Yates et al. (2021), respectively. Thresholds were allowed to differ by dataset (0.2 for newly generated and 0.3 for Fellows Yates et al. (2021)) and selected to minimise the number of oral taxa (taxa reported as oral in Chen et al. (2010) and Fellows Yates et al. (2021)) among all taxa identified as contaminants (Figure S12). All taxa identified as contaminants were then removed from the full dataset, regardless of which subset they were identified in.

Second, we performed abundance filtering on the entire dataset, setting each taxon to 0 if it had relative abundance <0.005% in a given sample. Third, we considered the relative abundances of taxa in the museum controls to remove likely environmental contaminants. A microbial taxon was removed as a likely contaminant if it was found to have a higher relative abundance in at least one of the four museum controls than in any of the samples. Finally, we considered bacterial genera identified as common contaminants in typical molecular and specialised ancient DNA laboratories (Salter et al., 2014; Weyrich et al., 2019). Some of the listed genera included taxa found in the Human Oral Microbiome Database (HOMD; accessed July 14 2021; (Chen et al., 2010) and in the core hominid oral microbiome (Fellows Yates et al., 2021). Therefore, we used a two-step approach to remove contaminant taxa. We directly removed genera that were listed only as contaminants and did not appear in the oral databases. Taxa that were present in both contaminant lists and the oral databases and had a sufficient number of reads in our dataset were investigated further by running mapDamage2 (V2.0.9; (Jónsson et al., 2013). To this end, we identified the sample with the highest abundance of a given taxon (according to Kraken2 output, requiring at least 10,000 reads) and mapped the reads to the respective reference genome. Taxa with too few reads to be tested were automatically retained. Since many taxa needed to be investigated in this way, making a visual inspection impractical, we assessed the presence of damage using a custom R script (Data and Materials). The presence of typical post-mortem DNA damages was assumed if deamination (C-to-T in the 3' end and G-to-A in the 5' end) was the most frequent change for at least two of the three terminal positions for each fragment end and had the frequency above 0.02 for at least one position. This rule was employed because in-line barcodes can create atypical DNA damage patterns in metagenomic data, showing lower damage at 1st terminal position (Brealey et al., 2020). Bacterial taxa with no damage were assumed to be modern contaminants and were therefore removed. The resulting dataset was used for downstream taxonomic analyses.

Sequence-level decontamination. We applied additional filtering by removing sequencing reads from taxa identified as contaminants (n=167) to allow for functional characterization of microbial communities and reconstructing metagenome-assembled genomes (MAGs). To this end, we used two approaches. First, we removed reads assigned to contaminants using the 'extract_kraken_reads.py' script from the KrakenTools suite (V1.2; (Wood et al., 2019). Using the read-level taxonomic assignments by Kraken2, we removed reads assigned to the 167

contaminant taxa identified above. This approach retained sequencing reads of all taxa not identified as contaminants (i.e. those included in the final taxonomic dataset), but also a large number of unclassified reads.

The reads which remained unclassified by Kraken2 could contain reads belonging to contaminant taxa. Therefore, we performed an additional decontamination step. We constructed a dataset containing the genomes of the identified contaminant taxa and a selection of abundant non-contaminant taxa from our dataset. To identify non-contaminant taxa, we selected all taxa with more than 10,000 reads in at least one sample (as long as they were not identified as contaminants), for a total of 72 unique taxa. The non-contaminant taxa were included to avoid forcing the reads to map to the contaminants. For both contaminants and non-contaminants, one genome per taxon was selected from the NCBI assembly summary files for bacteria, archaea, and viruses (NCBI ftp accessed 14 April 2021, available at https://github.com/markella-moraitou/Gorilla_dental_calculus/RD2_mapping/) with the following order of priority: genome representation (full > partial), RefSeq (yes > no), assembly level (full > chromosome > scaffold > contig), publication date (most recent). Twenty contaminant, and 11 non-contaminant taxa did not have a reference genome available. For them, we obtained a different genome from the same genus, unless that genus was already represented, using the taxize R package, and selected one genome per missing taxon as a replacement.

Finally, all references (consisting of 155 contaminant genomes and 61 non-contaminant genomes) were concatenated into a combined database. We mapped the FASTQ files from the previous decontamination step to this combined reference using BWA-MEM with default settings and then used SAMtools to retain the unmapped reads and reads that mapped primarily to the non-contaminant genomes.

Taxonomic analyses

All taxonomic analyses were performed at the species level. To account for the compositional nature of the data, a centred-log ratio (CLR) normalisation was applied (Gloor et al., 2017) using the 'transform' function of the microbiome R package (V1.12.0; (Lahti and Shetty, 2018).

Oral composition. To assess differences in oral microbiome composition by subspecies and sex, we used two distance measures. First, we evaluated presence-absence-based Jaccard distances (Jaccard, 1901) on the non-normalized data. Second, Aitchison's distances (Aitchison and Aitchison, 1986), which take into account relative abundances, were calculated as Euclidean distances on the CLR-transformed data. Using these distance metrics, we performed PERMANOVAs with the 'adonis' function of the vegan package (V2.5-7; (Oksanen et al., 2011). Because of the unequal distribution of host subspecies among the two datasets (newly generated versus Fellows Yates et al. (2021); Figure S1), we included dataset as a factor in each model and considered it prior to evaluating the contribution of biological variables. We used the 'adonis.pair' function of the EcolUtils R package (V0.1; (Salazar, 2019) to perform pairwise PERMANOVAs for significant factors. Principal Coordinate Analysis (PCoA) plots were generated using the 'plot_ordination' function of the phyloseq package. Because sex assignment was missing for some samples (Table S1), we re-ran PERMANOVAs for each distance measure for the subset of samples that could be sexed.

Alpha diversity. We estimated alpha diversity with two different metrics, applied to the non-normalized data: the Chao1 estimator (Chao et al., 2009), which evaluates species richness while accounting for the number of species that are missing from the dataset, and the Shannon index (Shannon, 1948), which reflects evenness. To assess the effect of the host subspecies on alpha diversity, we ran an ANOVA model and confirmed the model met assumptions of normality, using a Shapiro-Wilk's test on the Studentized residuals, and visually assessed the fit against the residuals (Quinn and Keough, 2002). Both Chao1 and Shannon models were transformed to account for skew ('sqrt(500-x)' and 'exp(x)', respectively). Finally, we used a Tukey post-hoc test to compare the different levels of the factor(s) that were statistically significant.

Identification of differentially abundant taxa. We identified bacterial taxa that significantly differed in abundance between host subspecies using ANCOM-II (Mandal et al., 2015), a method appropriate for compositional data, which accounts for the sparsity of the taxon table (Kaul et al., 2017). To account for the association between host subspecies and dataset (newly generated versus Fellows Yates *et al.* (2021); Figure S1) and to reduce the possibility that we identify dataset-specific taxa as differentially abundant, we ran the analysis twice. First, we identified microbial taxa that differed in abundance between host subspecies, while accounting

for read depth. Second, we identified microbial taxa that differed in abundance between datasets, while accounting for the host subspecies. Only genera that differed among subspecies and did not differ by dataset were considered relevant for evaluating microbiome differences among gorilla subspecies. Differentially abundant taxa were identified as those above the 0.9 quantile of the ANCOM W-statistic and Kruskal-Wallace post-hoc test (Bonferroni adjusted p-values < 0.05), or those that were structurally absent from one or more subspecies. Only microbial taxa identified at the genus or species level were considered.

Functional analysis

Decontaminated metagenomic reads were functionally characterised using an assembly-free approach implemented in the HUMAnN2 pipeline (V0.11.2; (Franzosa et al., 2018). The output tables containing the abundances of gene families with specific function were regrouped to show gene ontology (GO) terms (Ashburner et al., 2000; Gene Ontology Consortium, 2021), normalised to copies per million reads and feature abundances across taxa were aggregated in a single feature table, which was imported into R for further analysis. We subset our dataset to GO terms referring to biological processes and used PERMANOVA on their Euclidean distances to investigate functional differences of the oral microbiome by host subspecies. We then compared the abundances of biological processes among host subspecies using ANCOM, using the same approach as above.

Metagenome-assembled genomes (MAGs)

We used the metaWRAP pipeline (V1.3.2; (Uritskiy et al., 2018) to construct MAGs from decontaminated sequencing reads. To increase the amount of data available for building the initial assemblies with MEGAHIT (Li et al., 2016), we concatenated the reads across all samples (Li et al., 2016). The initial assembly was split into genomic bins using three different metagenomic binning tools (metabat2; (Kang et al., 2019), maxbin2; (Wu et al., 2016), CONCOCT; (Alneberg et al., 2013)) and we retained only those bins that were consistently defined by all three tools. For each MAG, we assigned taxonomic identity with GTDB-TK (V1.5.0; (Chaumeil et al., 2019) and constructed a phylogeny containing our MAGs and closely-related reference genomes identified by GTDB-TK. The abundance of MAGs in each sample was estimated using salmon (Patro et al., 2017) and the CLR-transformed abundances were visualised using pheatmap (Kolde and Others, 2012). For each assembled and

taxonomically identified MAG, we extracted and curated records of isolation sources (sources listed in Table S8) following an approach described by Madin et al. (2020). This information was used to categorise the isolation sources for taxa into one of the following categories: “contaminant”, “environmental”, “host-associated”, “oral”, or “unknown”.

To evaluate evolutionary relationships for select MAGs, reference genomes or gene sequences for the identified species/genus were downloaded from GenBank and Genome Taxonomy Database (GTDB; V202; (Parks et al., 2018). We created alignments using the amino acid sequences of core genes with PhyloPhlAn (Segata et al., 2013) and constructed maximum-likelihood phylogenies using FastTree (V2.1.10; (Price et al., 2010), which we visualised using ggtree (V3.1.0; (Yu et al., 2017).

We used two different approaches to determine the presence of genes involved in the nitrate reduction pathway in MAGs belonging to oral genera with nitrate reduction ability (see references provided in Rosier et al. (2020)). First, we used the pangenomic tool Panaroo (Tonkin-Hill et al., 2020) to determine presence of genes across a set of fragmented MAG assemblies. Second, we performed a translated BLAST search (tblastn, V2.11.0 (Camacho et al., 2009)) using protein sequences obtained from UniProt against MAG contigs. Genes in MAGs and reference genomes were considered as present if detected with either method.

Estimating ecological variation in oral microbiome composition

Using altitude as a proxy for ecology and diet, we estimated the correlation between altitudinal, geographical and microbiome distances. First, we approximated altitude and geographic location (latitude, and longitude) for each specimen based on the collection locality as retrieved from museum records. We constructed Euclidean distance matrices for altitudinal and geographical distance using R package vegan. Using the ade4 R package (1.7-16; (Dray et al., 2007), we performed partial Mantel tests (10,000 permutations) between microbiome distance matrices (Pearson correlation used for Jaccard distances and Spearman's rank coefficient correlation for Aitchison taxonomic and Euclidean functional distances) and altitudinal distances, while accounting for the effect of log-transformed geographical distances.

To further investigate the effects of altitude on Grauer’s gorilla microbiomes, which is the subspecies with the largest altitudinal range (Plumptre et al., 2016), we grouped the Grauer’s

gorilla samples into low- (500-1000 m above sea level [masl]) and high-altitude (>1000 masl) populations. We evaluated the contribution of altitude by performing a PERMANOVA, after accounting for potential dataset and host subspecies effects. Post hoc tests were performed using pairwise PERMANOVA, as described above.

Dietary Analysis

Eukaryotic reads in dental calculus were used to identify taxa potentially consumed by gorillas. Before taxonomically classifying these reads, we removed reads assigned to prokaryotes and viruses from the metagenomic dataset using the 'extract_kraken_reads.py' script from KrakenTools. The resulting FASTQ files were used as input in a second run of Kraken2 using the NCBI 'nt' database (accessed 27 September 2021). A joint feature table containing species and genus-level assignments was produced using kraken-biom and imported into R for further analysis.

Preprocessing of feature table. The feature table produced by kraken-biom was used in a phyloseq object along with a corresponding taxonomic matrix, created using the taxize R package. Using this matrix, we aggregated all identified taxa to the genus level, and used the decontam (Davis et al., 2018) R package to remove likely contaminants, as described above for the oral microbiome analyses, including using different cut-offs for each dataset. Then, all genera present in the museum controls or with fewer than 10 reads per sample were removed, to reduce contaminants, spurious and unreliable classifications. Despite this filtering, spurious classifications or contaminants were still present (species that are unlikely to be consumed by gorillas, e.g. mollusks), so we restricted it to genera belonging to taxonomic families that were previously reported to be consumed by gorillas (Michel et al., 2022; Remis et al., 2001; Rogers et al., 2004; Rothman et al., 2014; Yamagiwa et al., 2005). Since only few reference genomes of tropical plants are available and classically-used barcoding genes represent only a small part of the genome, accurate taxonomic classification from a sparse metagenomic sample is difficult and can likely only be achieved on genus or family level (Mann et al., 2020). For this reason, we focused our analysis not only on genera known to be consumed by gorillas, but also included all detected genera in the same family. The resulting feature table was used to detect differentially abundant dietary components with ANCOM using the same approach as described above, while accounting for differences in the number of reads per sample.

Collecting gorilla diet reference data. We produced a reference dataset of gorilla foods (Table S10) by searching the literature for plant species known to be consumed by the three gorilla subspecies based either on direct observations, analyses of food remains and faeces or more recently on molecular evidence (Michel et al., 2022; Remis et al., 2001; Rogers et al., 2004; Rothman et al., 2014; Yamagiwa et al., 2005). We retrieved the corresponding taxonomic IDs from NCBI using the taxize R package. For the taxa that had no match in the NCBI Taxonomy database, we manually searched the database and corrected the names that were obviously misspelt (difference of 1-2 characters from the NCBI entry). For the remaining taxa, we manually searched the GBIF species database (Wheeler, 2004) and obtained GBIF taxonomic IDs. However, several taxa (26 out of 314) could not be found in either database and therefore were excluded from our analysis. The taxonomic IDs obtained from either NCBI or GBIF were used with the 'classification' function of taxize to obtain the genus and family name for each taxon.

Availability of Data and Materials

Sequencing data generated in this study have been deposited on ENA, under Project Accession Number: PRJEB49638. Scripts for both preprocessing and downstream analysis are available at https://github.com/markella-moraitou/Gorilla_dental_calculus.

Funding

This work was supported by the Swedish Research Council (Formas) grant 2019-00275 and the Science for Life Laboratory National Sequencing Projects grant (NP00039) to KG. AF was supported by the Carl Tryggers Stiftelse grant 19:123. JFY and CW were supported by the Max Planck Society and the Werner Siemens-Stiftung (Paleobiotechnology, awarded to CW). None of the funders had any role in study design, collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

JCB, MM, AF extracted and prepared genomic libraries for the newly generated data. MM processed sequencing data, performed bioinformatic and statistical analyses (microbiome taxonomic and functional analyses, dietary and host genome analyses, altitudinal analyses) and wrote the manuscript draft. AF performed bioinformatic and statistical analyses (metagenome-assembled genomes, functional analyses, altitudinal analyses), wrote the respective parts of the manuscript and contributed to others. JCB performed initial bioinformatics analyses on a smaller dataset. JAFY and CW provided genomic data. KG conceived the research plan, interpreted the results, provided financial support and infrastructure, and contributed significantly to manuscript writing. All authors contributed to writing the final version of this manuscript.

Acknowledgments

We would like to thank Henrique Leitão, Thijs Hofstede, Franziska Aron, and Cody Parker, who contributed to processing the samples in the ancient DNA lab. Tom van der Valk, Axel Jensen and Samantha López Clinton for their contribution to subspecies verification using mitochondrial DNA. Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The SNP&SEQ Technology Platform is part of the National Genomics Infrastructure Sweden and Science of Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. We also acknowledge the National Bioinformatics Infrastructure for providing computational resources to this project.

References

Abusleme L, Dupuy AK, Dutzan N, Silva N, Burleson JA, Strausbaugh LD, Gamonal J, Diaz PI. 2013. The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *ISME J* **7**:1016–1025. doi:10.1038/ismej.2012.174

Abusleme L, O'Gorman H, Dutzan N, Greenwell-Wild T, Moutsopoulos NM. 2020. Establishment and Stability of the Murine Oral Microbiome. *J Dent Res* **99**:721–729. doi:10.1177/0022034520915485

Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, Bradshaw CJA,

Townsend G, Sołtysiak A, Alt KW, Parkhill J, Cooper A. 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat Genet* **45**:450–5, 455e1. doi:10.1038/ng.2536

Adler CJ, Malik R, Browne GV, Norris JM. 2016. Diet may influence the oral microbiome composition in cats. *Microbiome* **4**:23. doi:10.1186/s40168-016-0169-y

Aitchison J, Aitchison JW. 1986. The Statistical Analysis of Compositional Data. Springer Netherlands.

Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Loman NJ, Andersson AF, Quince C. 2013. CONCOCT: Clustering cONtigs on COverage and ComposiTion. *arXiv [q-bioGN]*.

Arora T, Sharma R. 2011. Fermentation potential of the gut microbiome: implications for energy homeostasis and weight management. *Nutr Rev* **69**:99–106. doi:10.1111/j.1753-4887.2010.00365.x

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Michael Cherry J, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene Ontology: tool for the unification of biology. *Nature Genetics*. doi:10.1038/75556

Barth KR, Isabella VM, Clark VL. 2009. Biochemical and genomic analysis of the denitrification pathway within the genus Neisseria. *Microbiology* **155**:4093–4103. doi:10.1099/mic.0.032961-0

Beall CM, Laskowski D, Erzurum SC. 2012. Nitric oxide in adaptation to altitude. *Free Radic Biol Med* **52**:1123–1134. doi:10.1016/j.freeradbiomed.2011.12.028

Beall CM, Laskowski D, Strohl KP, Soria R, Villena M, Vargas E, Alarcon AM, Gonzales C, Erzurum SC. 2001. Pulmonary nitric oxide in mountain dwellers. *Nature* **414**:411–412. doi:10.1038/35106641

Boehlke C, Rupf S, Tenniswood M, Chittur SV, Hannig C, Zierau O. 2020. Caries and periodontitis associated bacteria are more abundant in human saliva compared to other great apes. *Arch Oral Biol* **111**:104648. doi:10.1016/j.archoralbio.2020.104648

Bo T-B, Zhang X-Y, Kohl KD, Wen J, Tian S-J, Wang D-H. 2020. Coprophagy prevention alters microbiome, metabolism, neurochemistry, and cognitive behavior in a small mammal. *The ISME Journal*. doi:10.1038/s41396-020-0711-6

Brealey JC, Leitão HG, Hofstede T, Kalthoff DC, Guschanski K. 2021. The oral microbiota of wild bears in Sweden reflects the history of antibiotic use by humans. *Curr Biol* **31**:4650–4658.e6. doi:10.1016/j.cub.2021.08.010

Brealey JC, Leitão HG, van der Valk T, Xu W, Bougiouri K, Dalén L, Guschanski K. 2020. Dental Calculus as a Tool to Study the Evolution of the Mammalian Oral Microbiome. *Mol Biol Evol* **37**:3003–3022. doi:10.1093/molbev/msaa135

Bushnell B. 2014. BBMap: A fast, accurate, splice-aware aligner (No. LBNL-7065E). Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**:421. doi:10.1186/1471-2105-10-421

Cephas KD, Kim J, Mathai RA, Barry KA, Dowd SE, Meline BS, Swanson KS. 2011. Comparative Analysis of Salivary Bacterial Microbiome Diversity in Edentulous Infants and Their Mothers or Primary Care Givers Using Pyrosequencing. *PLoS ONE*. doi:10.1371/journal.pone.0023503

Chamberlain SA, Szöcs E. 2013. taxize: taxonomic search and retrieval in R. *F1000Res* **2**:191. doi:10.12688/f1000research.2-191.v2

Chao A, Colwell RK, Lin C-W, Gotelli NJ. 2009. Sufficient sampling for asymptotic minimum species richness estimators. *Ecology* **90**:1125–1133. doi:10.1890/07-2147.1
Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. doi:10.1093/bioinformatics/btz848
Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**:i884–i890. doi:10.1093/bioinformatics/bty560
Chen T, Yu W-H, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. 2010. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* **2010**:baq013. doi:10.1093/database/baq013
Cobb CM, Kelly PJ, Williams KB, Babbar S, Angolkar M, Derman RJ. 2017. The oral microbiome and adverse pregnancy outcomes. *Int J Womens Health* **9**:551–559. doi:10.2147/IJWH.S142730
Corby PM, Bretz WA, Hart TC, Schork NJ, Wessel J, Lyons-Weiler J, Paster BJ. 2007. Heritability of oral microbial species in caries-active and caries-free twins. *Twin Res Hum Genet* **10**:821–828. doi:10.1375/twin.10.6.821
Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, García N, Pääbo S, Arsuaga J-L, Meyer M. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A* **110**:15758–15763. doi:10.1073/pnas.1314445110
Dabney J, Meyer M. 2012. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *Biotechniques* **52**:87–94. doi:10.2144/000113809
Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* **10**. doi:10.1093/gigascience/giab008
Das R, Hergenrother SD, Soto-Calderón ID, Dew JL, Anthony NM, Jensen-Seaman MI. 2014. Complete mitochondrial genome sequence of the Eastern gorilla (Gorilla beringei) and implications for african ape biogeography. *J Hered* **105**:752–761. doi:10.1093/jhered/esu056
Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**:226. doi:10.1186/s40168-018-0605-2
Delamuta JRM, Scherer AJ, Ribeiro RA, Hungria M. 2020. Genetic diversity of Agrobacterium species isolated from nodules of common bean and soybean in Brazil, Mexico, Ecuador and Mozambique, and description of the new species Agrobacterium fabacearum sp. nov. *Int J Syst Evol Microbiol* **70**:4233–4244. doi:10.1099/ijsem.0.004278
Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R. 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A* **107**:11971–11975. doi:10.1073/pnas.1002601107
Donati C, Zolfo M, Albanese D, Tin Truong D, Asnicar F, Iebba V, Cavalieri D, Jousson O, De Filippo C, Huttenhower C, Segata N. 2016. Uncovering oral Neisseria tropism and persistence using metagenomic sequencing. *Nat Microbiol* **1**:16070. doi:10.1038/nmicrobiol.2016.70
Dray S, Dufour A-B, Others. 2007. The ade4 package: implementing the duality diagram for

ecologists. *J Stat Softw* **22**:1–20.
Eriksson H, Bagge E, Båverud V, Fellström C, Jansson DS. 2014. Erysipelothrix rhusiopathiae contamination in the poultry house environment during erysipelas outbreaks in organic laying hen flocks. *Avian Pathol* **43**:231–237. doi:10.1080/03079457.2014.907485
Erzurum SC, Ghosh S, Janocha AJ, Xu W, Bauer S, Bryan NS, Tejero J, Hemann C, Hille R, Stuehr DJ, Feelisch M, Beall CM. 2007. Higher blood flow and circulating NO products offset high-altitude hypoxia among Tibetans. *Proc Natl Acad Sci U S A* **104**:17593–17598. doi:10.1073/pnas.0707462104
Feelisch M. 2018. Enhanced nitric oxide production is a universal response to hypoxic stress. *National Science Review*. doi:10.1093/nsr/nwy041
Fellows Yates JA, Velsko IM, Aron F, Posth C, Hofman CA, Austin RM, Parker CE, Mann AE, Nägele K, Arthur KW, Arthur JW, Bauer CC, Crevecoeur I, Cupillard C, Curtis MC, Dalén L, Díaz-Zorita Bonilla M, Díez Fernández-Lomana JC, Drucker DG, Escribano Escrivá E, Francken M, Gibbon VE, González Morales MR, Grande Mateu A, Harvati K, Henry AG, Humphrey L, Menéndez M, Mihailović D, Peresani M, Rodríguez Moroder S, Roksandic M, Rougier H, Sázelová S, Stock JT, Straus LG, Svoboda J, Teßmann B, Walker MJ, Power RC, Lewis CM, Sankaranarayanan K, Guschanski K, Wrangham RW, Dewhirst FE, Salazar-García DC, Krause J, Herbig A, Warinner C. 2021. The evolution and changing ecology of the African hominid oral microbiome. *Proc Natl Acad Sci U S A* **118**. doi:10.1073/pnas.2021655118
Fletcher A, Nowell A. 2008. The development of feeding behaviour in wild western lowland gorillas (Gorilla gorilla gorilla). *Behaviour*. doi:10.1163/156853907783244747
Flynn KJ, Baxter NT, Schloss PD. 2016. Metabolic and Community Synergy of Oral Bacteria in Colorectal Cancer. *mSphere* **1**. doi:10.1128/mSphere.00102-16
Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, Lipson KS, Knight R, Caporaso JG, Segata N, Huttenhower C. 2018. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* **15**:962–968. doi:10.1038/s41592-018-0176-y
Galbany J, Imanizabayo O, Romero A, Vecellio V, Glowacka H, Cranfield MR, Bromage TG, Mudakikwa A, Stoinski TS, McFarlin SC. 2016. Tooth wear and feeding ecology in mountain gorillas from Volcanoes National Park, Rwanda. *Am J Phys Anthropol* **159**:457–465. doi:10.1002/ajpa.22897
Ganas J, Robbins MM, Nkurunungi JB, Kaplin BA, McNeilage A. 2004. Dietary variability of mountain gorillas in bwindi impenetrable National Park, Uganda. *Int J Primatol* **25**:1043–1072. doi:10.1023/b:ijop.0000043351.20129.44
Gao S, Li S, Ma Z, Liang S, Shan T, Zhang M, Zhu X, Zhang P, Liu G, Zhou F, Yuan X, Jia R, Potempa J, Scott DA, Lamont RJ, Wang H, Feng X. 2016. Presence of Porphyromonas gingivalis in esophagus and its association with the clinicopathological characteristics and survival in patients with esophageal cancer. *Infectious Agents and Cancer*. doi:10.1186/s13027-016-0049-x
Gene Ontology Consortium. 2021. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res* **49**:D325–D334. doi:10.1093/nar/gkaa1113
Gerner-Smidt P, Keiser-Nielsen H, Dorsch M, Stackebrandt E, Ursing J, Blom J, Christensen AC, Christensen JJ, Frederiksen W, Hoffmann S. 1994. Lautropia mirabilis gen. nov., sp. nov., a gram-negative motile coccus with unusual morphology isolated from the human mouth. *Microbiology* **140 (Pt 7)**:1787–1797. doi:10.1099/13500872-140-7-1787
Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. 2017. Microbiome Datasets Are

Compositional: And This Is Not Optional. *Front Microbiol* **8**:2224. doi:10.3389/fmicb.2017.02224
Gower G, Fenderson LE, Salis AT, Helgen KM, van Loenen AL, Heiniger H, Hofman-Kamińska E, Kowalczyk R, Mitchell KJ, Llamas B, Cooper A. 2019. Widespread male sex bias in mammal fossil and museum collections. *Proc Natl Acad Sci U S A* **116**:19019–19024. doi:10.1073/pnas.1903275116
Graczyk TK, Cranfield MR. 2003. Coprophagy and intestinal parasites: implications to human-habituated mountain gorillas (Gorilla gorilla beringei) of the Virunga Mountains and Bwindi Impenetrable Forest. *Primate Conserv* **9**:58–64.
Haas FB, Eisenhofer R, Schreiber M, Weyrich LS, Rensing SA. 2022. Neanderthals did not likely consume Physcomitrium patens – a model moss species. doi:10.1101/2022.01.26.472964
Hallast P, Maisano Delser P, Batini C, Zadik D, Rocchi M, Schempp W, Tyler-Smith C, Jobling MA. 2016. Great ape Y Chromosome and mitochondrial DNA phylogenies reflect subspecies structure and patterns of mating and dispersal. *Genome Res* **26**:427–439. doi:10.1101/gr.198754.115
Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* **486**:207–214. doi:10.1038/nature11234
Hu X-D, Gao L-Z. 2016. The complete mitochondrial genome of eastern lowland gorilla, Gorilla beringei graueri, and comparative mitochondrial genomics of Gorilla species. *Mitochondrial DNA A DNA Mapp Seq Anal* **27**:1484–1485. doi:10.3109/19401736.2014.953103
Hyde ER, Andrade F, Vaksman Z, Parthasarathy K, Jiang H, Parthasarathy DK, Torregrossa AC, Tribble G, Kaplan HB, Petrosino JF, Bryan NS. 2014a. Metagenomic analysis of nitrate-reducing bacteria in the oral cavity: implications for nitric oxide homeostasis. *PLoS One* **9**:e88645. doi:10.1371/journal.pone.0088645
Hyde ER, Luk B, Cron S, Kusic L, McCue T, Bauch T, Kaplan H, Tribble G, Petrosino JF, Bryan NS. 2014b. Characterization of the rat oral microbiome and the effects of dietary nitrate. *Free Radic Biol Med* **77**:249–257. doi:10.1016/j.freeradbiomed.2014.09.017
Jaccard P. 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vaud sci nat* **37**:547–579.
Janiak MC, Montague MJ, Villamil CI, Stock MK, Trujillo AE, DePasquale AN, Orkin JD, Bauman Surratt SE, Gonzalez O, Platt ML, Martínez MI, Antón SC, Dominguez-Bello MG, Melin AD, Higham JP. 2021. Age and sex-associated variation in the multi-site microbiome of an entire social group of free-ranging rhesus macaques. *Microbiome* **9**:68. doi:10.1186/s40168-021-01009-w
Jiao J-Y, Liu L, Hua Z-S, Fang B-Z, Zhou E-M, Salam N, Hedlund BP, Li W-J. 2021. Microbial dark matter coming to light: challenges and opportunities. *Natl Sci Rev* **8**:nwaa280. doi:10.1093/nsr/nwaa280
Johnston ER, Rodriguez-R LM, Luo C, Yuan MM, Wu L, He Z, Schuur EAG, Luo Y, Tiedje JM, Zhou J, Konstantinidis KT. 2016. Metagenomics Reveals Pervasive Bacterial Populations and Reduced Community Diversity across the Alaska Tundra Ecosystem. *Front Microbiol* **7**:579. doi:10.3389/fmicb.2016.00579
Jones AM, Vanhatalo A, Seals DR, Rossman MJ, Piknova B, Jonvik KL. 2021. Dietary Nitrate and Nitric Oxide Metabolism: Mouth, Circulation, Skeletal Muscle, and Exercise Performance. *Med Sci Sports Exerc* **53**:280–294. doi:10.1249/MSS.0000000000002470
Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*

29:1682–1684. doi:10.1093/bioinformatics/btt193
Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**:e7359. doi:10.7717/peerj.7359
Kartzinel TR, Hsing JC, Musili PM, Brown BRP, Pringle RM. 2019. Covariation of diet and gut microbiome in African megafauna. *Proc Natl Acad Sci U S A* **116**:23588–23593. doi:10.1073/pnas.1905666116
Kaul A, Mandal S, Davidov O, Peddada SD. 2017. Analysis of Microbiome Data in the Presence of Excess Zeros. *Front Microbiol* **8**:2114. doi:10.3389/fmicb.2017.02114
Kolde R, Others. 2012. Pheatmap: pretty heatmaps. *R package version* **1**:726.
Kolenbrander PE, Palmer RJ Jr, Periasamy S, Jakubovics NS. 2010. Oral multispecies biofilm development and the key role of cell-cell distance. *Nat Rev Microbiol* **8**:471–480. doi:10.1038/nrmicro2381
Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**:356. doi:10.1186/s12859-014-0356-4
Lahti L, Shetty S. 2018. Introduction to the microbiome R package. http://bioconductor.statistik.tu-dortmund.de/packages/3.6/bioc/vignettes/microbiome/inst/doc/vignette.html
Levett DZ, Fernandez BO, Riley HL, Martin DS, Mitchell K, Leckstrom CA, Ince C, Whipp BJ, Mythen MG, Montgomery HE, Grocott MP, Feelisch M, Caudwell Extreme Everest Research Group. 2011. The role of nitrogen oxides in human adaptation to hypoxia. *Sci Rep* **1**:109. doi:10.1038/srep00109
Levin D, Raab N, Pinto Y, Rothschild D, Zanir G, Godneva A, Mellul N, Futorian D, Gal D, Leviatan S, Zeevi D, Bachelet I, Segal E. 2021. Diversity and functional landscapes in the microbiota of animals in the wild. *Science* **372**. doi:10.1126/science.abb5352
Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, Yamashita H, Lam T-W. 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**:3–11. doi:10.1016/j.ymeth.2016.02.020
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**:1754–1760. doi:10.1093/bioinformatics/btp324
Li J, Nasidze I, Quinque D, Li M, Horz H-P, André C, Garriga RM, Halbwax M, Fischer A, Stoneking M. 2013. The saliva microbiome of Pan and Homo. *BMC Microbiol* **13**:204. doi:10.1186/1471-2180-13-204
Li Y, Ismail AI, Ge Y, Tellez M, Sohn W. 2007. Similarity of Bacterial Populations in Saliva from African-American Mother-Child Dyads. *Journal of Clinical Microbiology*. doi:10.1128/jcm.00771-07
Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, Brady A, Creasy HH, McCracken C, Giglio MG, McDonald D, Franzosa EA, Knight R, White O, Huttenhower C. 2017. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**:61–66. doi:10.1038/nature23889
Lu J, Breitwieser FP, Thielen P, Salzberg SL. 2017. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput Sci* **3**:e104. doi:10.7717/peerj-cs.104
Lundberg JO, Carlström M, Weitzberg E. 2018. Metabolic Effects of Dietary Nitrate in Health and Disease. *Cell Metab* **28**:9–22. doi:10.1016/j.cmet.2018.06.007
Madin JS, Nielsen DA, Brbic M, Corkrey R, Danko D, Edwards K, Engqvist MKM, Fierer N, Geoghegan JL, Gillings M, Kyrpides NC, Litchman E, Mason CE, Moore L, Nielsen SL, Paulsen IT, Price ND, Reddy TBK, Richards MA, Rocha EPC, Schmidt TM,

Shaaban H, Shukla M, Supek F, Tetu SG, Vieira-Silva S, Wattam AR, Westfall DA, Westoby M. 2020. A synthesis of bacterial and archaeal phenotypic trait data. *Sci Data* **7**:170. doi:10.1038/s41597-020-0497-4

Maisels F, Bergl RA, Williamson EA. 2018. Gorilla gorilla (amended version of 2016 assessment). *The IUCN Red List of Threatened Species* 2018–2012.

Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R, Peddada SD. 2015. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial Ecology in Health & Disease*. doi:10.3402/mehd.v26.27663

Mann AE, Fellows Yates JA, Fagernäs Z, Austin RM, Nelson EA, Hofman CA. 2020. Do I have something in my teeth? The trouble with genetic analyses of diet from archaeological dental calculus. *Quat Int*. doi:10.1016/j.quaint.2020.11.019

Mann AE, Sabin S, Ziesemer K, Vågene ÅJ, Schroeder H, Ozga AT, Sankaranarayanan K, Hofman CA, Fellows Yates JA, Salazar-García DC, Frohlich B, Aldenderfer M, Hoogland M, Read C, Milner GR, Stone AC, Lewis CM Jr, Krause J, Hofman C, Bos KI, Warinner C. 2018. Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Sci Rep* **8**:9822. doi:10.1038/s41598-018-28091-9

Marcelino, R, Vanessa, Holmes EC, Sorrell TC. 2020. The use of taxon-specific reference databases compromises metagenomic classification. *BMC Genomics* **21**:184. doi:10.1186/s12864-020-6592-2

McManus KF, Kelley JL, Song S, Veeramah KR, Woerner AE, Stevison LS, Ryder OA, Ape Genome Project G, Kidd JM, Wall JD, Bustamante CD, Hammer MF. 2015. Inference of gorilla demographic and selective history from whole-genome sequence data. *Mol Biol Evol* **32**:600–612. doi:10.1093/molbev/msu394

McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*. doi:10.1371/journal.pone.0061217

Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010**:db.prot5448. doi:10.1101/pdb.prot5448

Michel A, Minocher R, Niehoff P, Li Y, Nota K, Gadhvi MA, Su J, Iyer N, Porter A, Ngobobo-As-Ibungu U, Binyinyi E, Pekeyake RN, Parducci L, Caillaud D, Guschanski K. 2022. Isolated Grauer's gorilla populations differ in diet and gut microbiome. *bioRxiv*. doi:10.1101/2022.01.04.474987

Moeller AH, Foerster S, Wilson ML, Pusey AE, Hahn BH, Ochman H. 2016. Social behavior shapes the chimpanzee pan-microbiome. *Sci Adv* **2**:e1500997. doi:10.1126/sciadv.1500997

Mukherjee S, Huntemann M, Ivanova N, Kyrpides NC, Pati A. 2015. Large-scale contamination of microbial isolate genomes by Illumina PhiX control. *Stand Genomic Sci* **10**:18. doi:10.1186/1944-3277-10-18

Nakano K, Wada K, Nomura R, Nemoto H, Inaba H, Kojima A, Naka S, Hokamura K, Mukai T, Nakajima A, Umemura K, Kamisaki Y, Yoshioka H, Taniguchi K, Amano A, Ooshima T. 2011. Characterization of aortic aneurysms in cardiovascular disease patients harboring Porphyromonas gingivalis. *Oral Dis* **17**:370–378. doi:10.1111/j.1601-0825.2010.01759.x

Nichols RG, Peters JM, Patterson AD. 2019. Interplay Between the Host, the Human Microbiome, and Drug Metabolism. *Hum Genomics* **13**:27. doi:10.1186/s40246-019-0211-9

Nieuwdorp M, Gilijamse PW, Pai N, Kaplan LM. 2014. Role of the microbiome in energy

regulation and metabolism. *Gastroenterology* **146**:1525–1533. doi:10.1053/j.gastro.2014.02.008
Nikitakis NG, Papaioannou W, Sakkas LI, Kousvelari E. 2017. The autoimmunity-oral microbiome connection. *Oral Dis* **23**:828–839. doi:10.1111/odi.12589
Nishida AH, Ochman H. 2018. Rates of gut microbiome divergence in mammals. *Mol Ecol* **27**:1884–1897. doi:10.1111/mec.14473
Oh J, Byrd AL, Deming C, Conlan S, NISC Comparative Sequencing Program, Kong HH, Segre JA. 2014. Biogeography and individuality shape function in the human skin metagenome. *Nature* **514**:59–64. doi:10.1038/nature13786
Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2011. vegan: Community ecology package. *R package version* 117–118.
Olsen I, Singhrao SK. 2015. Can oral infection be a risk factor for Alzheimer's disease? *J Oral Microbiol* **7**:29143. doi:10.3402/jom.v7.29143
Ottoni C, Guellil M, Ozga AT, Stone AC, Kersten O, Bramanti B, Porcier S, Van Neer W. 2019. Metagenomic analysis of dental calculus in ancient Egyptian baboons. *Scientific Reports*. doi:10.1038/s41598-019-56074-x
Ozga AT, Nieves-Colón MA, Honap TP, Sankaranarayanan K, Hofman CA, Milner GR, Lewis CM Jr, Stone AC, Warinner C. 2016. Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. *Am J Phys Anthropol* **160**:220–228. doi:10.1002/ajpa.22960
Ozga AT, Ottoni C. 2021. Dental calculus as a proxy for animal microbiomes. *Quat Int*. doi:10.1016/j.quaint.2021.06.012
Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* **36**:996–1004. doi:10.1038/nbt.4229
Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**:417–419. doi:10.1038/nmeth.4197
Plumptre A, Nixon S, Caillaud D, Hall JS, Hart JA, Nishuli R, Williamson EA. 2016. Gorilla beringei ssp. graueri. The IUCN Red List of Threatened Species 2016: e. T39995A102328430 DOI 10.2305/IUCN. UK. 2016-2. *RLTS T39995A17989838 en*.
Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol* **16**:291–303. doi:10.1038/nrmicro.2017.171
Price MN, Dehal PS, Arkin AP. 2010. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**:e9490. doi:10.1371/journal.pone.0009490
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842. doi:10.1093/bioinformatics/btq033
Quinn GP, Keough MJ. 2002. Experimental Design and Data Analysis for Biologists. doi:10.1017/cbo9780511806384
R Core Team. 2015. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
Remis MJ, Dierenfeld ES, Mowry CB, Carroll RW. 2001. Nutritional aspects of western lowland gorilla (Gorilla gorilla gorilla) diet during seasons of fruit scarcity at Bai Hokou, Central African Republic. *Int J Primatol* **22**:807–836.
Rogers ME, Abernethy K, Bermejo M, Cipolletta C, Doran D, McFarland K, Nishihara T, Remis M, Tutin CEG. 2004. Western gorilla diet: a synthesis from six sites. *Am J Primatol* **64**:173–192. doi:10.1002/ajp.20071

Rosier BT, Moya-Gonzalvez EM, Corell-Escuin P, Mira A. 2020. Isolation and Characterization of Nitrate-Reducing Bacteria as Potential Probiotics for Oral and Systemic Health. *Front Microbiol* **11**:555465. doi:10.3389/fmicb.2020.555465

Rosier BT, Takahashi N, Zaura E, Krom BP, MartÍnez-Espinosa RM, van Breda SGJ, Marsh PD, Mira A. 2022. The Importance of Nitrate Reduction for Oral Health. *J Dent Res* 220345221080982. doi:10.1177/00220345221080982

Rothman JM, Nkurunungi JB, Shannon BF, Bryer MAH. 2014. High Altitude Diets: Implications for the Feeding and Nutritional Ecology of Mountain Gorillas In: Grow NB, Gursky-Doyen S, Krzton A, editors. High Altitude Primates. New York, NY: Springer New York. pp. 247–264. doi:10.1007/978-1-4614-8175-1_14

Roy J, Arandjelovic M, Bradley BJ, Guschanski K, Stephens CR, Bucknell D, Cirhuza H, Kusamba C, Kyungu JC, Smith V, Robbins MM, Vigilant L. 2014. Recent divergences and size decreases of eastern gorilla populations. *Biol Lett* **10**:20140811. doi:10.1098/rsbl.2014.0811

Ruff CB, Junno J-A, Burgess ML, Canington SL, Harper C, Mudakikwa A, McFarlin SC. 2022. Body proportions and environmental adaptation in gorillas. *American Journal of Biological Anthropology* **177**:501–529. doi:10.1002/ajpa.24443

Salazar G. 2019. EcolUtils: Utilities for community ecology analysis. *R Package Version 0 1 Available online: https://github com/GuillemSalazar/EcolUtils (accessed on 22 October 2019)*.

Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* **12**:87. doi:10.1186/s12915-014-0087-z

Sawafuji R, Saso A, Suda W, Hattori M, Ueda S. 2020. Ancient DNA analysis of food remains in human dental calculus from the Edo period, Japan. *PLoS One* **15**:e0226654. doi:10.1371/journal.pone.0226654

Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**:863–864. doi:10.1093/bioinformatics/btr026

Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* **9**:88. doi:10.1186/s13104-016-1900-2

Segata N, Börnigen D, Morgan XC, Huttenhower C. 2013. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun* **4**:2304. doi:10.1038/ncomms3304

Shaiber A, Willis AD, Delmont TO, Roux S, Chen L-X, Schmid AC, Yousef M, Watson AR, Lolans K, Esen ÖC, Lee STM, Downey N, Morrison HG, Dewhirst FE, Mark Welch JL, Eren AM. 2020. Functional and genetic markers of niche partitioning among enigmatic members of the human oral microbiome. *Genome Biol* **21**:292. doi:10.1186/s13059-020-02195-w

Shannon CE. 1948. A mathematical theory of communication. *The Bell System Technical Journal* **27**:379–423. doi:10.1002/j.1538-7305.1948.tb01338.x

Shenhav L, Thompson M, Joseph TA, Briscoe L, Furman O, Bogumil D, Mizrahi I, Pe'er I, Halperin E. 2019. FEAST: fast expectation-maximization for microbial source tracking. *Nat Methods* **16**:627–632. doi:10.1038/s41592-019-0431-x

Smith SN, Colston TJ, Siler CD. 2021. Venomous Snakes Reveal Ecological and Phylogenetic Factors Influencing Variation in Gut and Oral Microbiomes. *Front Microbiol* **12**:657754. doi:10.3389/fmicb.2021.657754

Soares-Castro P, Araújo-Rodrigues H, Godoy-Vitorino F, Ferreira M, Covelo P, López A,

Vingada J, Eira C, Santos PM. 2019. Microbiota fingerprints within the oral cavity of cetaceans as indicators for population biomonitoring. *Sci Rep* **9**:13679. doi:10.1038/s41598-019-50139-7

Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL. 1998. Microbial complexes in subgingival plaque. *Journal of Clinical Periodontology*. doi:10.1111/j.1600-051x.1998.tb02419.x

Suzuki TA. 2017. Links between Natural Variation in the Microbiome and Host Fitness in Wild Mammals. *Integr Comp Biol* **57**:756–769. doi:10.1093/icb/icx104

Takenoshita Y, Yamagiwa J. 2008. Estimating gorilla abundance by dung count in the northern part of Moukalaba-Doudou National Park, Gabon. *African Study Monographs Supplementary Issue* **39**:41–54.

Teles R, Wang C-Y. 2011. Mechanisms involved in the association between peridontal diseases and cardiovascular disease. *Oral Diseases*. doi:10.1111/j.1601-0825.2010.01784.x

Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, Gladstone RA, Lo S, Beaudoin C, Floto RA, Frost SDW, Corander J, Bentley SD, Parkhill J. 2020. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol* **21**:180. doi:10.1186/s13059-020-02090-4

Tsuchida S, Kakooza S, Mbehang Nguema PP, Wampande EM, Ushida K. 2018. Characteristics of Gorilla-Specific Lactobacillus Isolated from Captive and Wild Gorillas. *Microorganisms* **6**. doi:10.3390/microorganisms6030086

Tsuchida S, Kitahara M, Nguema PPM, Norimitsu S, Fujita S, Yamagiwa J, Ngomanda A, Ohkuma M, Ushida K. 2014. Lactobacillus gorillae sp. nov., isolated from the faeces of captive and wild western lowland gorillas (Gorilla gorilla gorilla). *Int J Syst Evol Microbiol* **64**:4001–4006. doi:10.1099/ijs.0.068429-0

Tsuchida S, Nezuo M, Tsukahara M, Ogura Y, Hayashi T, Ushida K. 2015. Draft Genome Sequence of Lactobacillus gorillae Strain KZ01T, Isolated from a Western Lowland Gorilla. *Genome Announc* **3**. doi:10.1128/genomeA.01196-15

Tsuzukibashi O, Uchibori S, Kobayashi T, Umezawa K, Mashimo C, Nambu T, Saito M, Hashizume-Takizawa T, Ochiai T. 2017. Isolation and identification methods of Rothia species in oral cavities. *J Microbiol Methods* **134**:21–26. doi:10.1016/j.mimet.2017.01.005

Turner LA, Bucking C. 2019. The role of intestinal bacteria in the ammonia detoxification ability of teleost fish. *J Exp Biol* **222**. doi:10.1242/jeb.209882

Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**:158. doi:10.1186/s40168-018-0541-1

van der Hoek Y, Pazo WD, Binyinyi E, Ngobobo U, Stoinski TS, Caillaud D. 2021. Diet of Grauer's Gorillas (Gorilla beringei graueri) in a Low-Elevation Forest. *Folia Primatol* **92**:126–138. doi:10.1159/000515377

van der Valk T, Díez-Del-Molino D, Marques-Bonet T, Guschanski K, Dalén L. 2019. Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas. *Curr Biol* **29**:165–170.e6. doi:10.1016/j.cub.2018.11.055

van der Valk T, Lona Durazo F, Dalén L, Guschanski K. 2017. Whole mitochondrial genome capture from faecal samples and museum-preserved specimens. *Mol Ecol Resour* **17**:e111–e121. doi:10.1111/1755-0998.12699

van der Valk T, Sandoval-Castellanos E, Caillaud D, Ngobobo U, Binyinyi E, Nishuli R, Stoinski T, Gilissen E, Sonet G, Semal P, Kalthoff DC, Dalén L, Guschanski K. 2018.

Significant loss of mitochondrial diversity within the last century due to extinction of peripheral populations in eastern gorillas. *Sci Rep* **8**:6551. doi:10.1038/s41598-018-24497-7
van der Valk T, Vezzi F, Ormestad M, Dalén L, Guschanski K. 2020. Index hopping on the Illumina HiseqX platform and its consequences for ancient DNA studies. *Mol Ecol Resour* **20**:1171–1181. doi:10.1111/1755-0998.13009
Velsko IM, Fellows Yates JA, Aron F, Hagan RW, Frantz LAF, Loe L, Martinez JBR, Chaves E, Gosden C, Larson G, Warinner C. 2019. Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. *Microbiome* **7**:102. doi:10.1186/s40168-019-0717-3
Velsko IM, Frantz LAF, Herbig A, Larson G, Warinner C. 2018. Selection of Appropriate Metagenome Taxonomic Classifiers for Ancient Microbiome Research. *mSystems* **3**. doi:10.1128/mSystems.00080-18
Vuong HE, Yano JM, Fung TC, Hsiao EY. 2017. The Microbiome and Host Behavior. *Annu Rev Neurosci* **40**:21–49. doi:10.1146/annurev-neuro-072116-031347
Wade WG. 2013. The oral microbiome in health and disease. *Pharmacol Res* **69**:137–143. doi:10.1016/j.phrs.2012.11.006
Warinner C. 2016. Dental Calculus and the Evolution of the Human Oral Microbiome. *J Calif Dent Assoc* **44**:411–420.
Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, Radini A, Hancock Y, Tito RY, Fiddyment S, Speller C, Hendy J, Charlton S, Luder HU, Salazar-García DC, Eppler E, Seiler R, Hansen LH, Castruita JAS, Barkow-Oesterreicher S, Teoh KY, Kelstrup CD, Olsen JV, Nanni P, Kawai T, Willerslev E, von Mering C, Lewis CM Jr, Collins MJ, Gilbert MTP, Rühli F, Cappellini E. 2014. Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* **46**:336–344. doi:10.1038/ng.2906
Warinner C, Speller C, Collins MJ. 2015. A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. *Philos Trans R Soc Lond B Biol Sci* **370**:20130376. doi:10.1098/rstb.2013.0376
Welch JLM, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. 2016. Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences* **113**:E791–E800.
West CE, Rydén P, Lundin D, Engstrand L, Tulic MK, Prescott SL. 2015. Gut microbiome and innate immune response patterns in IgE-associated eczema. *Clin Exp Allergy* **45**:1419–1429. doi:10.1111/cea.12566
Weyrich LS, Farrer AG, Eisenhofer R, Arriola LA, Young J, Selway CA, Handsley-Davis M, Adler CJ, Breen J, Cooper A. 2019. Laboratory contamination over time during low-biomass sample analysis. *Molecular Ecology Resources*. doi:10.1111/1755-0998.13011
Wheeler QD. 2004. What if GBIF? *Bioscience* **54**:718. doi:10.1641/0006-3568(2004)054[0718:wig]2.0.co;2
Williamson EA, Maisel FG, Groves CP, Fruth BI, Humle T, Morton FB, Richardson MC, Russon AE, Singleton I. 2013. Family Hominidae (Great Apes) In: Mittermeier RA, Rylands AB, Wilson DE, editors. Handbook of the Mammals of the World: Primates. Lynx Edicions in association with Conservation International and IUCN.
Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. *Genome Biol* **20**:257. doi:10.1186/s13059-019-1891-0
Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**:605–607. doi:10.1093/bioinformatics/btv638

Xue Y, Prado-Martinez J, Sudmant PH, Narasimhan V, Ayub Q, Szpak M, Frandsen P, Chen Y, Yngvadottir B, Cooper DN, de Manuel M, Hernandez-Rodriguez J, Lobon I, Siegismund HR, Pagani L, Quail MA, Hvilsom C, Mudakikwa A, Eichler EE, Cranfield MR, Marques-Bonet T, Tyler-Smith C, Scally A. 2015. Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science* **348**:242–245. doi:10.1126/science.aaa3952

Xu X, Arnason U. 1996. A complete sequence of the mitochondrial genome of the western lowland gorilla. *Mol Biol Evol* **13**:691–698. doi:10.1093/oxfordjournals.molbev.a025630

Yamagiwa J, Basabose AK, Kaleme KP, Yumoto T. 2008. Phenology of fruits consumed by a sympatric population of gorillas and chimpanzees in Kahuzibiega National Park, Democratic Republic of Congo. *African Study Monographs Supplementary Issue* **39**:3–22.

Yamagiwa J, Basabose AK, Kaleme K, Yumoto T. 2005. Diet of grauer's gorillas in the Montane forest of kahuzi, democratic republic of Congo. *Int J Primatol* **26**:1345–1373. doi:10.1007/s10764-005-8856-8

Yan J, Li Y, Han XZ, Chen WF, Zou WX, Xie Z, Li M. 2017. Agrobacterium deltaense sp. nov., an endophytic bacteria isolated from nodule of Sesbania cannabina. *Arch Microbiol* **199**:1003–1009. doi:10.1007/s00203-017-1367-0

Youngblut ND, de la Cuesta-Zuluaga J, Reischer GH, Dauser S, Schuster N, Walzer C, Stalder G, Farnleitner AH, Ley RE. 2020. Large-Scale Metagenome Assembly Reveals Novel Animal-Associated Microbial Genomes, Biosynthetic Gene Clusters, and Other Genetic Diversity. *mSystems* **5**. doi:10.1128/mSystems.01045-20

Youngblut ND, Reischer GH, Walters W, Schuster N, Walzer C, Stalder G, Ley RE, Farnleitner AH. 2019. Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. *Nat Commun* **10**:2200. doi:10.1038/s41467-019-10191-3

Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. 2017. Ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* **8**:28–36. doi:10.1111/2041-210x.12628