1 **Ecology, not host phylogeny, shapes the oral microbiome in closely related**
2 **species of gorillas**

3

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

6

7 [1]Animal Ecology, Department of Ecology and Genetics, Uppsala University, 75236 Uppsala,
8 Sweden

9 [2]Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, 04103
10 Leipzig, Germany

11 [3]Department of Paleobiotechnology, Leibniz Institute for Natural Product Research and
12 Infection Biology Hans Knöll Institute, 07745 Jena, Germany

13 [4]Department of Natural History, NTNU University Museum, Norwegian University of Science
14 and Technology, 7491 Trondheim, Norway

15 [5]Faculty of Biological Sciences, Friedrich Schiller University, 07743 Jena, Germany

16 [6]Department of Anthropology, Harvard University, Cambridge, MA 02138

17 [7]Science for Life Laboratory, 75237 Uppsala, Sweden

18 [8]Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh,
19 Edinburgh EH9 3FL, United Kingdom

20 [#]These authors contributed equally

21 [*]Correspondence   email: katerina.guschanski@ed.ac.uk;   markella.mor@gmail.com;
22 adrian.e.forsythe@gmail.com

23 Key words: ancient DNA, dental calculus, museum specimens, gorilla, metagenome-
24 assembled genomes, shotgun metagenomics

25

26

Abstract

Host-associated microbiomes are essential for a multitude of biological processes. Placed at the contact zone between external and internal environments, the little-studied oral microbiome has important roles in host physiology and health. Here we investigate the contribution of host evolutionary relationships and ecology in shaping the oral microbiome in three closely related gorilla subspecies (mountain, Grauer's, and western lowland gorillas) using shotgun metagenomics of 46 museum-preserved dental calculus samples. We find that the oral microbiomes of mountain gorillas are functionally and taxonomically distinct from the other two subspecies, despite close evolutionary relationships and geographic proximity with Grauer's gorillas. Grauer's gorillas show intermediate bacterial taxonomic and functional, and dietary profiles. Altitudinal differences in gorilla subspecies ranges appear to explain these patterns, proposing a close connection between dental calculus microbiome and the environment, which is further supported by the presence of gorilla subspecies-specific phyllosphere/rhizosphere taxa. Mountain gorillas show high abundance of nitrate-reducing oral taxa, which may contribute to high altitude adaptation by modulating blood pressure. Our results suggest that ecology, rather than evolutionary relationships and geographic proximity, primarily shape the oral microbiome in these closely related species.

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

2

49   al., 2019; Turner and Bucking, 2019), immune responses (Nikitakis et al., 2017; West et al.,

50   2015) and even neurochemical and hormonal processes that eventually influence behaviour

51   (Suzuki, 2017; Vuong et al., 2017). The extensively studied gut microbiome is shaped by a

52   multitude of factors, including host phylogenetic relationships, social interactions, and dietary

53   adaptations (Moeller et al., 2016; Nishida and Ochman, 2018; Youngblut et al., 2019). In

54   comparison to the gut microbiome, the oral microbiome has received far less attention. Yet, it

55   is of particular importance, as it connects the external and the internal environments and is

56   located directly at the entry point to the digestive and the respiratory tracts. The oral

57   microbiome plays an important role in oral diseases, such as caries and periodontitis, and has

58   been implicated in systemic disorders including cardiovascular disease (Nakano et al., 2011),

59   atherosclerosis (Teles and Wang, 2011), Alzheimer's disease (Olsen and Singhrao, 2015),

60   several cancers (Flynn et al., 2016; Gao et al., 2016) and preterm births (Cobb et al., 2017).

61   Although studies have linked oral microbiome composition to both dietary habits (Adler et al.,

62   2016; Hyde et al., 2014b; Janiak et al., 2021; Nearing et al., 2020) and host evolutionary

63   relationships (Brealey et al., 2020; Ozga and Ottoni, 2021), it is unclear which of the two

64   predominantly drives the evolution of the oral microbiome and how they interact. Early studies

65   suggested that the oral microbiome is strongly heritable and mostly transferred vertically

66   (Corby et al., 2007; Li et al., 2007), with neonate microbiomes resembling those of their

67   mothers (Dominguez-Bello et al., 2010). Moreover, studies in wild animals have indicated that

68   different species harbour distinct oral microbiota (Brealey et al., 2020; Ozga and Ottoni, 2021).

69   Both observations suggest that host evolutionary relationships have a strong effect on oral

70   microbiome structure. Other studies have highlighted the effect of diet on the oral microbiome

71   (Janiak et al., 2021; Wade, 2013), for instance by uncovering differences between young

72   infants and adults of humans and non-human primates, hence providing a potential link with

3

73 age-related dietary changes (Cephas et al., 2011). However, distinguishing between host

74 evolutionary and ecological factors as well as their contribution to oral microbiome evolution

75 remains difficult, as most previous studies have focused on rather distantly related host species

76 that occupy distinct ecological niches (Boehlke et al., 2020; Li et al., 2013; Smith et al., 2021;

77 Soares-Castro et al., 2019).

78 In recent years, dental calculus - the calcified form of dental plaque that forms on the teeth of

79 mammals - has emerged as a useful material for the study of oral microbiomes across diverse

80 mammalian species (Brealey et al., 2020; Fellows Yates et al., 2021; Ottoni et al., 2019;

81 Warinner et al., 2014). Since dental plaque undergoes periodic mineralizations, it effectively

82 fossilises on the living host, reducing postmortem contamination from environmental

83 microorganisms (Warinner et al., 2015). It thus preserves a snapshot of oral and respiratory

84 microbial communities (Warinner et al., 2014), dietary components (Adler et al., 2013; Radini

85 et al., 2017; Sawafuji et al., 2020) and host DNA (Mann et al., 2018). Museum collections can

86 be efficiently used for the study of oral microbiomes from wild animals (Brealey et al., 2020;

87 Fellows Yates et al., 2021), minimising the exposure and disturbances associated with sampling

88 from live hosts. Furthermore, the detection of damage patterns, a common verification method

89 for ancient DNA (Briggs et al., 2007), can also be implemented on such historical specimens

90 to distinguish endogenous taxa from modern contaminants (Fellows Yates et al., 2021).

91 Previous research suggests that microbial communities in historical and modern dental calculus

92 are remarkably similar (Velsko et al., 2019), therefore studying museum specimens provides

93 reliable information about present-day microbiomes. The exceptional preservation in the

94 calcified matrix also allows assembly of near complete metagenome-assembled genomes

95 (MAGs) (Brealey et al., 2020).

4

96   In this study, we use dental calculus to uncover ecological and evolutionary factors that drive

97   oral microbiome evolution in a group of closely related species. We focus on gorillas, which

98   consist of two species, western gorilla (*Gorilla gorilla*) and eastern gorilla (*G. beringei*), each

99   of which is further divided into two subspecies. Our sampling covers three of the four

100   subspecies: western lowland gorilla (*G. g. gorilla*), Grauer's gorilla (*G. b. graueri*) and

101   mountain gorilla (*G. b. beringei*). Eastern and western gorillas diverged approximately 250,000

102   years ago, with evidence for substantial gene flow until more recently (McManus et al., 2015;

103   Xue et al., 2015). Western lowland gorillas are found in western equatorial Africa, primarily at

104   low elevations below 500 m above sea level (masl). Their diet consists mainly of various plant

105   parts and fruits (Rogers et al., 2004). The eastern subspecies - mountain and Grauer's gorillas

106   - occur on the eastern side of the Congo basin. They are estimated to have diverged from each

107   other 10,000 years ago (Roy et al., 2014). Grauer's gorillas occupy the largest altitudinal range

108   of all gorillas, ranging from 500 to 2,900 masl (Plumptre et al., 2016). Mountain gorillas are

109   found at even higher elevations, with most of the range of the Virunga Massif population above

110   2,200 masl and extending to up to 3,800 masl. Consequently, these altitudinal differences lead

111   to different dietary profiles among populations and subspecies (Doran et al., 2002; Ganas et

112   al., 2004; Michel et al., 2022; Rogers et al., 2004; van der Hoek et al., 2021), primarily

113   reflecting availability of fruits that are highly seasonal and become scarce with increasing

114   altitude (Ganas et al., 2004).

115   The presence of ecological and dietary differences in these closely related gorilla subspecies

116   provides an opportunity to evaluate the effects of host evolutionary relationships and ecological

117   factors on oral microbiomes. We used shotgun metagenomic sequencing of dental calculus

118   samples from museum specimens (collected between 1910 and 1986) to taxonomically and

119   functionally characterise the oral microbiota, as well as to assemble bacterial MAGs and

5

120   identify dietary components from the three gorilla subspecies. Between closely related gorilla

121   subspecies, ecology has a large impact on the taxonomic composition and function capacities

122   of the oral microbiome. Our analyses highlight subspecies-specific differences that can be

123   attributed ecological/dietary differences, and bacterial members of the oral microbiome that

124   may contribute to adaptation to high-altitude lifestyle.

125

Results
126

127   **Data pre-processing and confirmation of oral microbial signature**

128   We produced paired-end shotgun sequencing data from the gorilla dental calculus of 26 newly

129   collected samples, and complemented these with an additional 31 dental calculus samples from

130   previously published studies (Brealey et al., 2020; Fellows Yates et al., 2021) (Table S1). Four

131   dental calculus samples were sequenced independently in this study and by Fellows Yates et

132   al. (2021). At a later stage of the analysis, this set of technical duplicates allowed us to assess

133   the effect of the dataset in our taxonomic analysis, but only the sample with the largest number

134   of reads for each sample pair was retained for further analyses.

135   Raw sequencing data consisted of, on average, 15,533,319 reads (mean value) per sample

136   (ranging from 1,032 to 95,367,058; Table S1) and 105,653 reads per negative control (both

137   extraction and library preparation, ranging from 19 to 1,132,867; Table S1). After raw sequence

138   pre-processing (which included removal of poly-G tails, adapter and barcode sequences,

139   merging of forward and reverse reads, and quality filtering), *phiX* and gorilla/human sequences

140   were removed by mapping against respective reference genomes and the resulting unmapped

141   reads were used for taxonomic classification using Kraken2 (Wood et al., 2019) with the

6

142   standard database, which includes all bacterial, archaeal and viral genomes from NCBI.

143   Samples with low read counts, low proportion of oral taxa, as well as duplicate samples (those

144   processed at two different facilities), were removed (Methods). The final dataset consisted of

145   46 dental calculus samples (13 western lowland gorillas, 17 mountain gorilla, 16 Grauer's

146   gorilla; Figure S1), each containing 561,978-77,307,443 reads (mean: 11,365,074, SD:

147   15,577,988; Table S1). We then applied a multi-step procedure, relying on negative controls

148   and museum environmental samples, to remove contaminant taxa (Methods). The final dataset

149   contained a total of 1007 microbial species (in 430 genera), of which 3.4% (n=34) were

150   members of the core hominid oral microbiome (Fellows Yates et al., 2021) and 4.2% (n=42)

151   members of the Human Oral Microbiome Database (HOMD; (Chen et al., 2010); Table S2).

152   These microbial species accounted for ca. 14% of the total microbial non-normalised

153   abundance.

154   The final dataset also included six genera that have been listed as common contaminants by

155   Salter *et al.* (2014) and Weyrich *et al.* (2019), but contain known oral taxa (Chen et al., 2010;

156   Fellows Yates et al., 2021), including *Streptococcus oralis* and *Staphylococcus saprophyticus.*

157   Species belonging to these genera were retained after confirming, where possible, that they

158   exhibit typical post-mortem DNA damage patterns (Methods).

159

160   **Partial mitochondrial genomes from dental calculus allow host subspecies identification**

161   We recovered host mitochondrial sequences from 44 of 46 samples, which we used to confirm

162   gorilla subspecies identity. On average, 31.2% (0.44 - 94.2%; Table S3) of the reference

163   mitochondrial genome was covered by mapped reads, with notable variation in the proportion

7

164   of genome covered and coverage depth (0.01 - 77.04%, average: 3.27%) between samples.

165   Sample weight (information available for n=24 samples), sample age (n=36), or dataset (newly

166   generated versus previously published by Fellows Yates et al. (2021)) (n=46) did not have an

167   effect on mtDNA genome completeness (ANOVA, p>0.54), whereas host subspecies identity

168   was marginally significant (ANOVA, F-value=3.12, p=0.054). Mountain gorillas had a larger

169   proportion of the mtDNA genome covered by mapped reads than other gorilla subspecies. The

170   cause of this difference is unknown but could be related to factors affecting the shedding of

171   host DNA into the oral environment, such as differences in dietary abrasiveness, salivary flow,

172   or inflammation.

173   To assist with host taxon assignment, we identified 403 diagnostic sites between western and

174   eastern gorillas and 72 diagnostic sites between the two eastern gorilla subspecies (mountain

175   and Grauer's gorillas) using published gorilla mitochondrial genomes (Methods). In most

176   samples, reads mapped to only a small number of diagnostic sites (Table S3). We accepted

177   molecular taxon assignment for samples with reads mapping to at least six diagnostic sites in

178   six separate reads. For six samples, data was insufficient to distinguish between eastern and

179   western gorilla species and for additional six samples we could not reliably distinguish between

180   mountain and Grauer's gorilla subspecies. In these cases, we accepted the host subspecies

181   identification based on museum records, after consulting collection locality, where possible.

182   All but one sample containing a sufficient number of diagnostic sites were successfully

183   assigned to their reported subspecies. The exception was sample MTM010 (museum accession:

184   631168, Swedish Museum of Natural History - NRM), which was confirmed to be a mountain

185   gorilla, in accordance with the museum records, although it was previously reported as a

186   western lowland gorilla based on preliminary genomic information (Fellows Yates et al., 2021).

8

187

Oral microbiome diversity and composition differ by host subspecies

189 We evaluated the differences in microbial alpha diversity between gorilla subspecies by

190 measuring species richness and community evenness (Figure 1). Host subspecies had only a

191 marginal effect on the normalised microbial richness (ANOVA, p=0.061; Figure 1a) but a

192 significant effect on the transformed community evenness (ANOVA, p<0.001; Figure 1b), with

193 the microbial communities of mountain gorillas showing the lowest evenness.
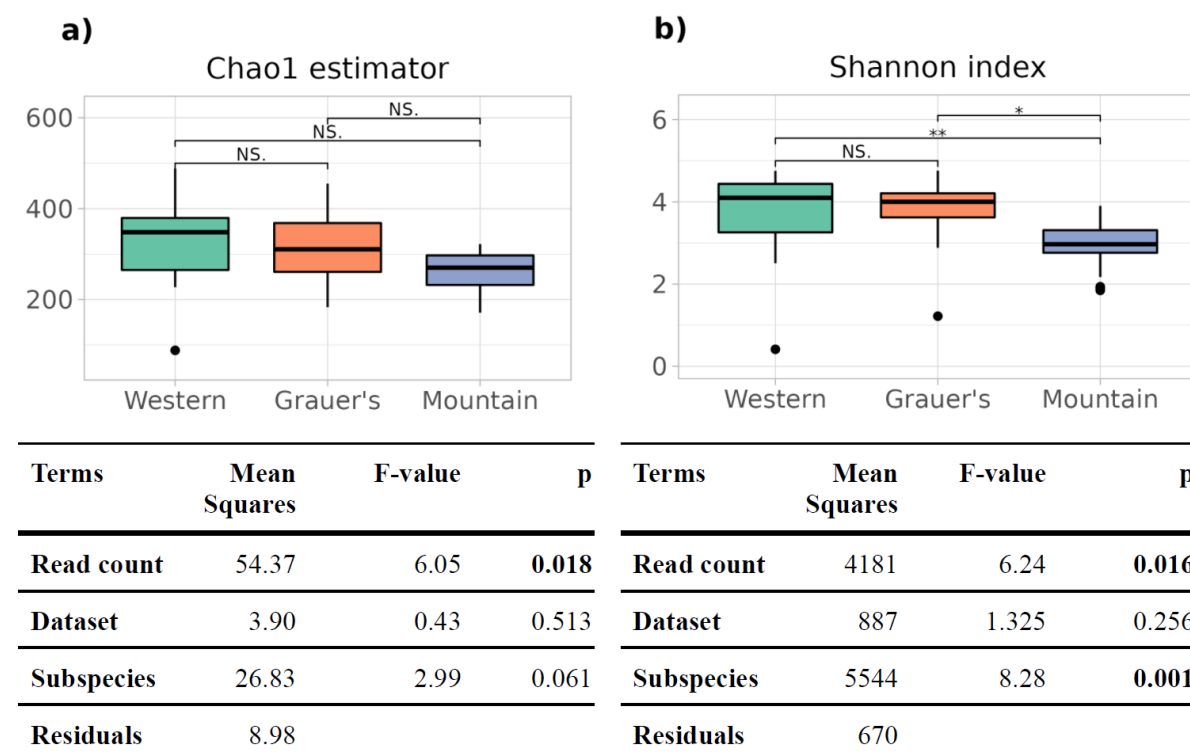

a)

Chao1 estimator

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		

b)

Shannon index

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		

194

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

199

9

200   Microbiome composition differed by gorilla subspecies (Figure 2, Table 1). This effect

201   persisted even after other potentially confounding factors, such as sequencing depth and

202   dataset, were taken into account. Sequencing depth, included as read count, and dataset were

203   significant factors for the presence-absence-based Jaccard measure, whereas neither of these

204   factors significantly affected relative abundances as estimated with Aitchison distances (Table

205   1). We confirmed that belonging to different datasets had little to no effect on relative

206   abundance by considering three pairs of samples that were derived from the same museum

207   specimen, but were independently sequenced in this study and in Fellows Yates et al. (2021):

208   the duplicate samples appeared close to each other on ordination plots, in particular on the plot

209   based on Aitchison distances (Figure S2). The fourth sample pair (G0004-IBA002) was

210   excluded, as both replicates had lower than 3% oral component, according to microbial source

211   tracking using FEAST (Shenhav et al., 2019) (Figure S3). Pairwise PERMANOVAs

212   (PERMutational ANalysis Of VAriance) (Anderson, 2001) showed strong differentiation

213   between mountain gorillas and the other two subspecies, using both presence-absence and

214   abundance metrics (Table 1). Lastly, we found no significant effect of sex on the oral

215   microbiome using a subset of samples that could be genetically sexed (n=38; Methods; Table
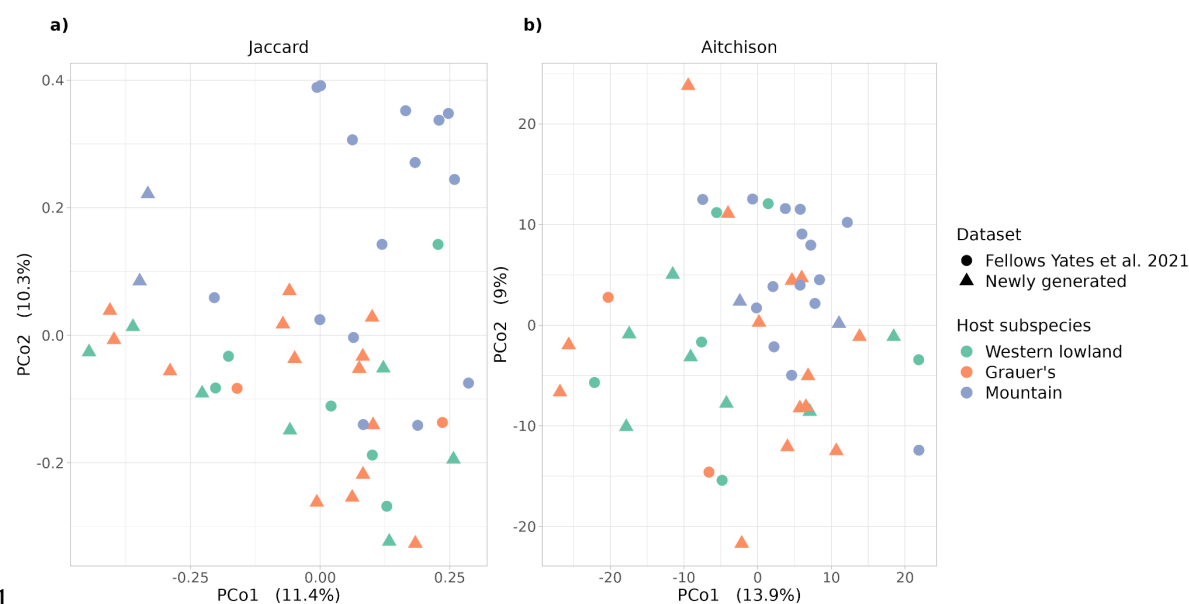
216   S4).

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.

Terms	Within-factor comparison	Jaccard distances		Aitchison distances	
		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 significantly differed in abundance among gorilla subspecies and 11 that significantly 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 artefacts, 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

12

233   one host subspecies. The abundances of these taxa were comparable to not differentially

234   abundant taxa (n=929) (t-test: $t(123.59)=1.24$, $p=0.22$): 37 taxa ranked above average and 75

235   were in the third quartile of abundance ranks, so that their absence from some host subspecies

236   is unlikely to be explained by non-detection due to low abundance.

237   Mountain gorillas, in particular, appeared to be missing many of these differentially abundant

238   microbial taxa (Figure 3) — which is consistent with the general observation of a marginally

239   reduced microbial richness in this subspecies (Figure 1) — even after accounting for the effect

240   of read depth. Microbial taxa that were absent in mountain gorillas included the orders

241   *Rhodobacteriales*, *Pseudonocardiales*, *Corynebacteriales* (represented mainly by

242   *Corynebacterium* and *Mycolicibacterium*), *Bacillales* (including *Staphylococcus*), and

243   *Rhizobiales* (including bacteria associated with the *Fabaceae* rhizosphere, such as

244   *Agrobacterium deltaense, A. fabacearum* (Delamuta et al., 2020; Yan et al., 2017), and three

245   *Rhizobium* species (Poole et al., 2018)). The presence of rhizosphere- and phyllosphere-

246   associated taxa in the other two subspecies, western lowland and Grauer's gorillas, may reflect

247   habitat or dietary differences among the subspecies. Microbial taxa enriched in mountain

248   gorillas primarily belonged to the orders *Enterobacterales* and certain *Lactobacillales*, like

249   *Streptococcus* sp. and *Lactobacillus gasseri*. The microbiomes of Grauer's gorillas resembled

250   those of western lowland gorillas, both overall (Figure 2, Table 1) and in terms of differentially

251   abundant taxa (Figure 3). However, they also shared some similarities with mountain gorillas

252   (e.g. a presence of *Limosilactobacillus*/*Lactobacillus* species, which were absent in western

253   lowland gorillas), showing an intermediate or mixed composition (Figure 3).

13

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 with an assembly-free approach using the

14

263   HUMAnN2 pipeline (Franzosa et al., 2018) (Methods). The normalised abundances of gene

264   families were regrouped under gene ontology (GO) terms, and were aggregated across

265   microbial taxa. Results of functional analyses were broadly consistent with those observed for

266   community-level taxonomy. Specifically, gorilla subspecies was marginally significant

267   (p=0.06) in explaining functional differences in the oral microbiome, whereas neither

268   sequencing depth nor dataset had a significant effect (Table 2). The functional profiles of

269   mountain gorilla microbiomes differed significantly from the other two subspecies (Table 2).

270   We used ANCOM to identify differentially abundant biological processes in the dental calculus

271   microbial communities of the gorilla subspecies, requiring that these processes were

272   represented in at least 30% of the samples in at least one subspecies. Mountain gorillas stood

273   out again, as they missed genes for 236 of the 262 differentially abundant processes (Table S6).

274   In contrast, western lowland and Grauer's gorillas were missing only 13 and 6 processes,

275   respectively. Among six processes that were found in all three gorilla subspecies, mountain

276   gorillas differed significantly from the others in four processes.

277

278 **Table 2**. PERMANOVA results, showing the effect size (R^2) and p-value of factors putatively contributing to the
279 variance in abundance of biological processes among sampled microbiota. Pairwise PERMANOVA between
280 subspecies show the effect size (R^2) and p-values adjusted for false discovery rate. P-values below 0.05 are shown
281 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	

282

283 **Metagenome-Assembled Genomes**

284 We applied an iterative assembly and binning approach (Methods) to resolve initial MAGs

285 using decontaminated sequencing reads from all samples and museum controls. Across all

286 samples, we constructed a total of eight high-quality (>90% completion, <5% contamination)

287 and 27 medium quality MAGs (>50% completion, <10% contamination), belonging to 24

288 distinct bacterial families (Table S7). Among these 35 MAGs, 25 could be classified to the

289 genus level, with two MAGs assigned to provisional/uncultured genera (UBA8133 and

290 RUG013; Table S7). To place the MAGs within a broader evolutionary context, we constructed

16

291   a phylogenetic tree including the MAGs and phylogenetically closest bacterial genomes from

292   the Genome Taxonomy Database (GTDB; (Parks et al., 2018); Figure S4). However, for

293   several MAGs assembled in this study we were unable to identify a closely related published

294   genome.

295   Fourteen MAGs reconstructed from dental calculus samples were also present in museum

296   controls (a gorilla skin sample, a gorilla petrous bone, a skull swab and a museum shelf swab).

297   Eight MAGs were present only in the skin sample and only at low abundances (Figure S5).

298   However, five MAGs were found at higher abundance in at least one of the museum controls

299   compared to any gorilla dental calculus sample (Figure S5). This includes MAGs of

300   *Erysipelothrix* (found in gorilla petrous bone), two different *Planococcaceae* (found in the

301   petrous bone and museum shelf swab), as well as *Propionibacterium* and *Exiguobacterium*

302   (both isolated from museum skull swab). These MAGs belong to genera that are common

303   contaminants in metagenomic studies (i.e. *Propionibacterium*, which is a major genus of

304   common skin bacteria (Salter et al., 2014)) and persistent in environmental reservoirs (i.e.

305   *Erysipelothrix*; (Eriksson et al., 2014)). These five MAGs were excluded from downstream

306   analyses. To further authenticate our reconstructed MAGs, we used a collection of isolation

307   source records from public databases of bacterial ecological metadata (Table S8). Taxa with a

308   high proportion (>=25%) of isolation records from sources categorised as

309   environmental/contaminant were considered as potential contaminants (Table S8). Using this

310   approach, we identified 10 additional MAGs that represent likely contaminants (Figures S5&

311   S6).

312   The remaining 20 MAGs recovered here are members of the oral microbiome and are present

313   in databases of common oral taxa (Chen et al., 2010; Fellows Yates et al., 2021). This includes

17

314   taxa commonly associated with dental plaque communities: *Rothia* (Tsuzukibashi et al., 2017),

315   *Olsenella* (Abusleme et al., 2013; Socransky et al., 1998), *Corynebacterium* (Mark Welch et

316   al., 2016), *Lautropia* (Gerner-Smidt et al., 1994), *Neisseria (Donati et al., 2016)* and

317   *Actinomyces* (Kolenbrander et al., 2010). *Rothia* species are particularly abundant in gorillas

318   compared to humans and other non-human primates (Fellows Yates et al., 2021). Three other

319   MAGs were present in at least 65% of all samples: a MAG most closely related to *Neisseria*

320   (in 95.7% of samples) was found in all Grauer's and mountain gorillas, and the MAGs

321   characterised to the family *Actinomycetaceae* and the genus *Lautropia*, both of which were

322   found in all samples of mountain gorillas (Figure S5).

323   For two MAGs of high completeness and high prevalence in our samples (*Neisseria* and

324   *Limosilactobacillus gorillae*), we produced an alignment of core genes to further investigate

325   how they relate to the known diversity of these taxa. The *Neisseria* MAG clustered with a

326   subset of *Neisseria* species isolated exclusively from humans and was more divergent from

327   *Neisseria* species that formed a clade with isolates from other animals (Figure 4a). Its

328   phylogenetic placement suggests a distinct and not yet identified *Neisseria* taxon. We also

329   recovered a near-complete MAG of *Limosilactobacillus gorillae* (Figure S7), a species

330   previously isolated from faeces of mountain and western lowland gorillas (Tsuchida et al.,

331   2018, 2015). The dental calculus *L. gorillae* MAG was sister to the gorilla faecal isolate but

332   distinct from faecal isolates of other primates (Figure S7a).

333   Oral bacteria play an important role in the metabolism of inorganic nitrate and its reduction to

334   nitrite and nitric oxide, which is essential for the regulation of cardiovascular, metabolic, and

335   neurological processes (Bryan et al., 2017; Hezel and Weitzberg, 2015). Several oral nitrate

336   reducing bacteria, including *Rothia*, *Neisseria*, *Veillonella*, *Corynebacterium*, *Actinomyces*,

18

337   *Selenomonas*, *Propionibacterium*, *Fusobacterium*, and *Eikenella (Rosier et al., 2022),* were

338   present in gorilla dental calculus samples (Figure S5, Table S5). We used MAGs of *Rothia*,

339   *Neisseria*, and *Veillonella*, which were prevalent in the study samples (Figure S5) to analyse

340   the presence of nitrate-reducing genes in these members of the gorilla oral microbiome.

341   Although we recovered a medium quality MAG of *Corynebacterium* from some samples of

342   western and Grauer's gorillas (Figure S5), it was less prevalent and abundant in gorilla dental

343   calculus and is lacking well-characterised genes involved in nitrate reduction. For *Neisseria*,

344   we could confirm the presence of five nitrate reducates (*nar*), seven genes that provide the

345   cofactor molybdenum to the nitrate reductase enzymes, and four additional genes involved in

346   the regulation of nitrate and nitrite reduction (*aniA*, *nirB*, *nirD*, *norB*; Figure 4a). Only two

347   genes involved in nitrate reduction, *narY* and *mopII*, were not found in the gorilla dental

348   calculus MAG. However, these genes are also absent from other nitrate-reducing oral bacteria

349   (Rosier et al., 2020). Genomic data thus suggests that *Neisseria* isolated from gorilla dental

350   calculus likely has the full capacity to reduce nitrate. The *Neisseria* MAG recovered in this

351   study was significantly more abundant in mountain gorillas compared to the other host

352   subspecies (Figure 4b).

353   *Rothia* and *Veillonella* possessed a smaller repertoire of nitrate-reducing genes compared to

354   *Neisseria*. The *Rothia* MAG recovered in this study had several nitrate reductases and

355   molybdenum cofactor genes, but they differed from genes present in *R. dentocariosa*, a

356   bacterium well studied for its nitrate-reducing capacity (Figure S8; (Rosier et al., 2020). We

357   were also able to identify nine genes involved in nitrate metabolism in *Veillonella*, a well

358   known reducer of nitrate in the oral cavity (Hyde et al., 2014a) (Figure S8). Similar to

359   *Neisseria*, *Veillonella* was significantly more abundant in mountain gorillas than in the other

19

360 two gorilla subspecies, whereas the abundance of *Rothia* did not differ among gorilla

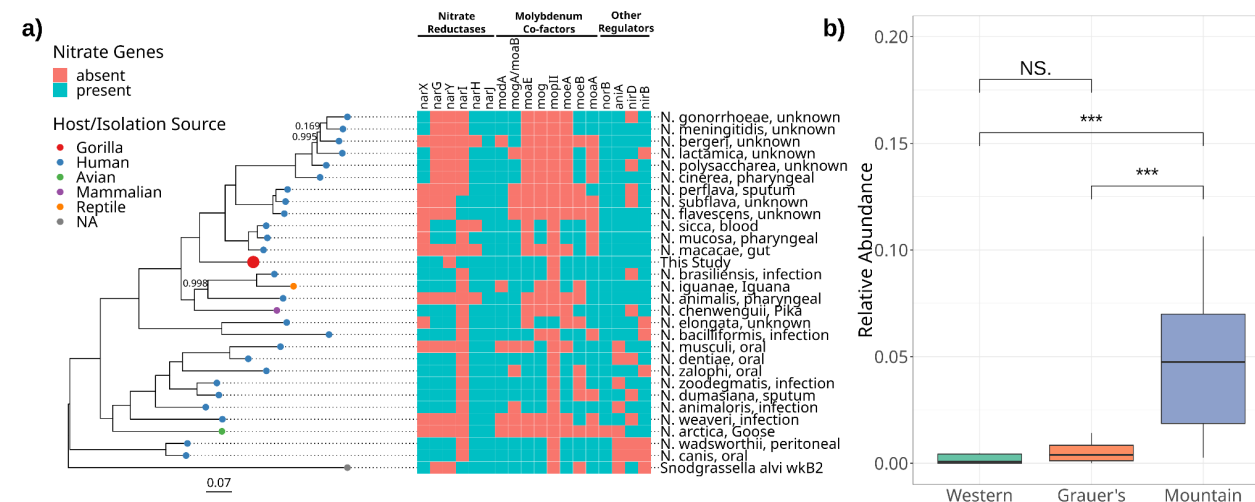361 subspecies (Figure S8b and d).

362

Figure 4: 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 the proportion of node support out of 100 bootstrap replicates, with unlabeled
nodes having complete support (1.0). 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).

371

Altitude may drive oral microbiome composition

373 Both taxonomic and functional analyses suggest that Grauer's gorilla oral microbiomes are

374 more similar to western lowland gorillas than to mountain gorillas, despite sharing a close

375 evolutionary relationship and an adjacent geographic range with the latter. However, the

376 distribution ranges of all three gorilla subspecies differ in elevation, with mountain gorillas

377 occurring at the highest altitudes. As altitude can influence temperature, humidity and food

378 diversity, which in turn can influence microbial communities, we performed partial Mantel

379 tests between altitudinal distances and taxonomic (Jaccard/Aitchison distances) or functional

380 (Euclidean distance) composition of the oral microbiome, while accounting for log-

381   transformed geographical distance. Geographic location and altitude of the specimens were

382   approximated based on museum records. However, when considering the entire dataset, we did

383   not detect a correlation between altitude and either taxonomy or function (Mantel test: R

384   ranging from -0.13 to 0.06, p>0.12).

385   Since Grauer's gorillas occupy the widest altitudinal range of all gorilla subspecies, we divided

386   Grauer's gorilla samples into high- (>1000 masl, n=8) and low-altitude (≤1000 masl, n=8)

387   groups and tested for differentiation of the oral microbiome both between these groups and

388   among host subspecies. While no compositional differences were identified between the oral

389   microbiomes of western lowland and low altitude Grauer's gorillas using relative abundance

390   measures, western lowland and high altitude Grauer's gorillas displayed marginally significant

391   differences based on pairwise PERMANOVA (FDR-adjusted p-value=0.075, Table 3). The

392   oral microbiome of mountain gorillas consistently differed from other subspecies, independent

393   of altitude grouping (Table 3). Yet, high-altitude Grauer's gorillas showed considerable overlap

394   with mountain gorillas in the first two axes of a PCoA based on Aitchison distances (Figure

395   S9).

396

21

397 **Table 3.** Results of the pairwise PERMANOVA showing the effect size (R^2) of comparisons of taxonomic
398 composition based on Aitchison distances among four gorilla groups (western lowland, low-latitude Grauer's
399 [$\leq$1000 masl], high-altitude Grauer's [$>$1000 masl], and mountain gorillas) and the associated p-values (adjusted
400 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**

401

402 To further investigate the effect of altitude on the oral microbiome, we identified a set of

403 differentially abundant taxa between mountain and western lowland gorillas (n=41) and asked

404 if Grauer's gorillas from different altitudes show similarities with these two subspecies. We

405 observe no significant difference between low-altitude Grauer's and western lowland gorilla

406 and between high-altitude Grauer's and mountain gorillas, whereas all other comparisons were

407 significant (p<0.003, Table S9).

22

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 (Mann et al., 2020). 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 5, Table S10). 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 5). *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). The family Fabaceae is consumed by all gorilla subspecies and the genus *Tamarindus* detected here is likely a misidentification of

23

431   another member of this family, as *Tamarindus* itself has not been reported to be part of gorilla

432   diet. Grauer's gorillas appear to have the most diverse diet of the three subspecies, consuming

433   foodstuffs typical of both western lowland and mountain gorillas (Figure 5). This is in

434   accordance with the literature (Michel et al., 2022; Remis et al., 2001; Rogers et al., 2004;

435   Rothman et al., 2014; Yamagiwa et al., 2005), which reports 61 plant families as being part of

436   their diet, as opposed to 41 and 45 families for western lowland and mountain gorillas,

437   respectively.

438   Similar to our analyses of differentially abundant bacteria, dietary components which were

439   differentially abundant between mountain and western lowland gorillas showed no significant

440   difference between low-altitude Grauer's and western lowland gorilla and between high-

441   altitude Grauer's and mountain gorillas, while significantly differentiating low- and high-

442   altitude Grauer's gorillas (Table S9).

Figure 5. Heatmap based on normalised abundances for differentially abundant dietary taxa at the genus level. 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 dental calculus oral microbiome in closely related host taxa. Ecologically, all three subspecies differ from each other, but Grauer's and mountain gorillas are geographically proximate and more closely related to each other than to western lowland gorillas. One of the strongest differences between the three subspecies is the altitude of their ranges, which in turn influences diet, reflected, e.g., in the level of frugivory. Western lowland gorillas occupy the

25

456   lowest altitudes and are the most frugivorous (Rogers et al., 2004; Takenoshita and Yamagiwa,

457   2008). Mountain gorillas occupy the highest altitude and are the least frugivorous (Ganas et

458   al., 2004; Maisels et al., 2018; Plumptre et al., 2016), whereas Grauer's gorillas live across the

459   largest altitudinal range and hence straddle the ecological conditions faced by both western

460   lowland and mountain gorillas (Yamagiwa et al., 2008). Ecological conditions and dietary

461   differences are expected to affect the host-associated microbial communities (Janiak et al.,

462   2021; Kartzinel et al., 2019; Li et al., 2013; Wade, 2013; Youngblut et al., 2019).

463   By investigating the oral microbiome preserved in dental calculus, we show that oral microbial

464   communities among closely related gorilla subspecies differ, both taxonomically and

465   functionally. In particular, multiple lines of evidence show that the oral microbiome of

466   mountain gorillas differs from the other gorilla subspecies in taxonomic composition, relative

467   abundance, and functional profiles of microbial taxa (Tables 1 and 2), despite the close

468   evolutionary relationship and geographic proximity between mountain and Grauer's gorillas.

469   The oral microbiome of Grauer's gorillas shows strong similarity to western lowland gorillas,

470   although the two species are more distantly related and separated by the Congo Basin. When

471   considering differentially abundant microbial taxa and dietary profiles, Grauer's gorillas

472   appear intermediate, displaying a combination of taxa found in both western lowland and

473   mountain gorillas (Figure 3 and 6). This suggests that host phylogenetic relationships are less

474   important for the structure of the oral microbiome in closely related species than ecological

475   conditions, possibly including diet.

476   These observations deviate from previous studies that identified a phylogenetic signal in the

477   oral microbiome in wild animals. However, these studies compared distantly related species

478   like chimpanzees, humans (Li et al., 2013), other great apes (Boehlke et al., 2020), and even

26

479   more diverse groups of mammals (Brealey et al., 2020; Ozga and Ottoni, 2021). Often the

480   effects of host phylogenetic relationships and ecology could not be decoupled. For instance,

481   Smith et al. (2021) found differences among three snake species that each occupied a distinct

482   ecological niche. Similarly, Soares-Castro et al. (2019) studied three marine mammals, among

483   which the most evolutionary divergent species also lived in a different habitat, so the effects of

484   ecology were indistinguishable from those of evolutionary relationships.

485   Yet, ecological effects on a short evolutionary scale, as observed in this study, do not preclude

486   co-diversification of the host and its associated microbiome over longer evolutionary time.

487   They rather point to the presence of rapid and dynamic responses of the host-associated

488   microbial communities to environmental conditions, potentially through the seeding of the

489   dental calculus oral microbiome at least to some extent with environmental taxa (see below).

490   In addition, our results are in accordance with other studies that propose external factors, such

491   as diet, influencing the oral microbiome (Adler et al., 2016; Hyde et al., 2014b; Janiak et al.,

492   2021), including Li et al. (2013) who observed that the oral microbiome of captive apes is

493   distinct from their wild counterparts and more similar to other captive apes, highlighting the

494   importance of environmental factors/diet.

495

496   **Oral microbiome diversity might reflect ecology or diet**

497   Our analyses provide a tentative suggestion that ecological differences, approximated here by

498   altitude, may be the primary factor shaping oral microbiome diversity. Specifically, grouping

499   Grauer's gorillas into low- and high-altitude populations, we showed that low-altitude

500   populations harbour an oral microbiome that is indistinguishable from western lowland

501 gorillas, whereas high-altitude populations show near-significant differences (Table 3).

502 Although the oral microbiome of Grauer's gorillas is significantly different from mountain

503 gorillas independent of the altitude, this may be because most mountain gorillas in our dataset

504 were collected above 3000m, notably higher than any of the Grauer's gorilla samples (600-

505 2100m; Table S1). We also observe compositional differences between high- and low- altitude

506 groups of Grauer's gorillas when considering oral and dietary taxa that differentiate mountain

507 and western lowland gorillas (Table S9). The high-altitude populations of Grauer's gorillas

508 appear indistinguishable from mountain gorillas, whereas low-altitude Grauer's gorillas show

509 no differences to western lowland gorillas, supporting the effect of ecological differences

510 connected to altitude that shape microbial and dietary composition. Generally, the dietary

511 profiles of Grauer's gorillas appear to contain dietary items present in both western lowland

512 and mountain gorillas (Figures 6 and S10). Our observation from dental calculus is in line with

513 previous studies (Table S11), which report considerable overlap in the diet of Grauer's gorillas

514 with western lowland (30 of the 41 plant families) and mountain (30 of the 45 plant families)

515 gorillas. In contrast, only 16 dietary families are consumed by both western lowland and

516 mountain gorillas.

517 It is also intriguing to speculate that the local environment itself could affect the composition

518 of the dental calculus community by serving as a source of at least some colonising taxa, for

519 example via the consumed phyllosphere and rhizosphere. A recent study of the human oral

520 microbiome found that specific taxa within the dental plaque community are evolutionarily

521 close to environmental bacteria, whereas those living on the tongue surface are more closely

522 related to other host-associated taxa (Shaiber et al., 2020). This finding suggested that dental

523 plaque microbial communities could be more affected by the environment than other host-

524 associated microbiomes. Within the gorilla oral microbiome, we detected microbial taxa that

28

525   likely have a dietary origin and differ in abundance among host subspecies, possibly reflecting

526   dietary differences. In particular, bacteria associated with the *Fabaceae* rhizosphere, including

527   *Agrobacterium deltaense, A. fabacearum* (Delamuta et al., 2020; Yan et al., 2017), and three

528   *Rhizobium* species (Poole et al., 2018) may be derived from the consumed dietary items or

529   accidentally ingested soil. These bacteria were significantly more abundant in the dental

530   calculus of western lowland and Grauer's gorillas than in mountain gorillas (Figure 3, Table

531   S5). Two lines of evidence provide a link between these microbial taxa and dietary differences

532   among the gorilla subspecies. First, our data suggest a higher abundance of the family

533   *Fabaceae* in western lowland and Grauer's gorillas (Figures 6 and S10, Table S10) compared

534   to mountain gorillas. Second, although no comparative data exists on the prevalence of roots

535   and rhizomes in the diet of different gorilla species, indirect evidence suggests that western

536   lowland gorillas consume more roots than mountain gorillas. Behavioural observations report

537   frequent root and rhizome consumption in western lowland gorillas of all ages (Fletcher and

538   Nowell, 2008) and comparative dentition analyses show increased tooth wear in western

539   lowland gorillas compared to mountain gorillas, in which tooth wear is well explained by the

540   time spent feeding on roots (Galbany et al., 2016).

541   Although dental calculus is a highly suitable material for the study of the oral microbiome,

542   several previous studies have identified caveats associated with dietary analyses from this

543   source material (Haas et al., 2022; Mann et al., 2020; Ottoni et al., 2019). Sample

544   contamination, the low proportion of eukaryotic reads, and the sparsity of eukaryotic reference

545   databases lead to biases and present a considerable challenge to the analysis of dental calculus.

546   In addition, reference genome contamination can result in false positive taxonomic assignments

547   (Mann et al., 2020). In the present study, the authenticity of putatively dietary taxa could not

548   be assessed using damage patterns due to low abundance of eukaryotic reads. Nevertheless, we

29

549 were able to detect dietary differences among gorilla subspecies, which are mostly in

550 agreement with previous knowledge of gorilla diets (Figure 5) (Michel et al., 2022; Remis et

551 al., 2001; Rogers et al., 2004; Rothman et al., 2014; Yamagiwa et al., 2005).

552

Metagenome-assembled genomes add to the understanding of wild microbiomes

554 The recovery of high quality MAGs from metagenomic data can provide an important

555 perspective on the evolutionary relationships among microbial taxa and between

556 microorganisms and their hosts. As reference databases are often biased towards organisms of

557 medical, agricultural, and industrial importance (Marcelino et al., 2020; Velsko et al., 2018),

558 MAGs recovered from novel environments, such as dental calculus oral microbiome of non-

559 human animals, may represent poorly described, understudied or completely unknown

560 microbial lineages (Jiao et al., 2021). Several of the high and medium quality MAGs

561 constructed in this study were recovered in higher abundance from environmental controls than

562 from any dental calculus samples and were considered to represent environmental

563 contaminants. However, other MAGs were highly abundant in the dental calculus samples and

564 belong to taxa that are members of the oral microbiome (Figures S5 & S6). Many of these

565 MAGs were distinct from the evolutionary closest sequenced genomes available in GTDB

566 (Figures S4, S7, S8, Figure 4), suggesting that they may represent undescribed bacterial

567 lineages, in line with discoveries of novel microbial diversity in unstudied host species (Levin

568 et al., 2021; Youngblut et al., 2020). The closest evolutionary relatives of several bacterial taxa

569 with high quality MAGs were found among isolates from primates, including humans (Figures

570 5a, S7, S8).

30

571   In this study, we recovered a near-complete MAG of *Limosilactobacillus gorillae*, which was

572   more closely related to a faecal isolate from a captive gorilla than to a faecal isolate from

573   another primate (Figure S7a). Usually associated with gorilla faeces (Tsuchida et al., 2014),

574   the presence of *L. gorillae* within the oral microbiome may be the result of coprophagy, a

575   common behaviour among wild gorillas (Graczyk and Cranfield, 2003). However, questions

576   remain regarding the persistence of faecal-associated bacteria within dental calculus and their

577   potential function within the plaque biofilm. Coprophagic behaviour is suggested to serve as a

578   route for vertical or horizontal transmission of gastrointestinal microbiota between individuals

579   (Abusleme et al., 2020) and may have a stabilising impact on host-associated microbial

580   communities overall (Bo et al., 2020).

581   We also reconstructed MAGs of *Neisseria*, *Rothia*, and *Veillonella* and confirmed the presence

582   of nitrate-reducing genes in these members of the gorilla oral microbiome (Figures 5a, S8).

583   The reduction of nitrate to nitrite from dietary sources by oral bacteria provides the precursor

584   to nitric oxide (NO), which is an important signalling and effector molecule (Jones et al., 2021).

585   Along with multiple benefits to circulatory health (Lundberg et al., 2018), nitric oxide is

586   increased in response to hypoxic stress (Feelisch, 2018; Levett et al., 2011), and results in

587   increased blood oxygen levels in the host (Beall et al., 2012). Human populations adapted to

588   high altitude exhale higher concentrations of nitric oxide compared to low-land populations

589   (Beall et al., 2001; Erzurum et al., 2007). For these reasons, regulation of nitric oxide

590   metabolism is thought to be beneficial to high-altitude adaptation (Beall et al., 2012).

591   Mountain gorillas live at elevation of up to 3,800 masl (Williamson et al., 2013) and hence are

592   exposed to hypoxic stress (Ruff et al., 2022). They also show higher abundance of *Neisseria*

593   and *Veillonella* compared to the other two gorilla subspecies, which live at considerably lower

31

594   altitudes (Figures 5b and S8d). Although it is not known if dietary nitrate amounts differ among

595   gorilla subspecies, the primarily herbivorous diet of mountain gorillas is naturally rich in

596   nitrate. Taken together, it is likely that high abundance of nitrate-reducing oral taxa may aid

597   mountain gorillas cope with the physiological demands of their high-altitude habitat. This

598   finding is particularly intriguing, as comparative genomic analyses failed to uncover host-

599   encoded genes related to high-altitude adaptation in mountain gorillas (Xue et al., 2015). Our

600   results suggest an important role for host-associated microbiomes in promoting host

601   adaptations.

602

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

604   The mitochondrial sequences recovered from the dental calculus metagenomes were used to

605   confirm gorilla subspecies identity. Six of the samples yielded mitochondrial genomes with

606   completeness >80% and coverage >3X (maximum 94% completeness and 77X coverage, Table

607   S3). However, in many cases even a few mitochondrial fragments were sufficient to allow

608   subspecies assignment. We could genetically distinguish gorilla species (eastern versus

609   western gorillas) with mitochondrial genome coverage as low as 2.6% and eastern gorilla

610   subspecies (mountain versus Grauer's) with coverage as low as 6.6%. Grauer's and mountain

611   gorilla mitochondrial genomes could be successfully distinguished despite differing by only

612   0.5% (72 positions, excluding the hypervariable D-loop region (van der Valk et al., 2018)).

613   This observation supports the usefulness of dental calculus as material for obtaining genetic

614   information about the host (Mann et al., 2018; Ozga et al., 2016; Warinner, 2016; Warinner et

615   al., 2015).

616

Conclusions

618 Our study is, to our knowledge, the first to investigate the evolution of the dental calculus oral

619 microbiome at the early stages of the speciation process of the host and adds to the new but

620 growing field of research on the microbiomes of wild animals. We find that in closely related

621 species, evolutionary relationships are less important than ecology in explaining taxonomic

622 composition and function of the oral microbiome. Host-associated microbial communities have

623 been proposed to contribute to host adaptation, partly because they can respond rapidly to

624 changing environmental conditions (Alberdi et al., 2016). We find that taxa enriched in the

625 mountain gorilla oral microbiome may facilitate their high-altitude lifestyle through increased

626 nitrate reduction potential and the associated physiological benefits. Our discovery of distinct

627 phylo- and rhizosphere taxa in the dental calculus of different gorilla subspecies suggests a

628 close connection between this host-associated microbiome and the local environment.

629 Colonisation of the host by environmental taxa with beneficial functions in local adaptation

630 and health may present an exciting evolutionary route to be explored in the future.

631

Methods

Sample collection

634 The study dataset consisted of dental calculus samples from 57 specimens of three gorilla

635 subspecies: 16 western lowlands gorillas, 22 Grauer's gorillas, and 19 mountain gorillas (Table

636 S1). Dental calculus samples of western lowland gorillas were collected at the Royal Museum

33

637   for Central Africa (RMCA, Tervuren, Belgium) and the Cleveland Museum of Natural History

638   (CMNH, USA). Samples of mountain gorillas came from the Swedish Museum of Natural

639   History (Naturhistoriska Riksmuseet - NRM, Stockholm, Sweden) and the RMCA. Samples of

640   Grauer's gorillas were collected primarily at RMCA, with a few samples from NRM and the

641   Royal Belgian Institute of Natural Sciences (RBINS, Brussels, Belgium). Our dataset consisted

642   of newly generated shotgun data from 26 specimens and published gorilla dental calculus

643   sequences: two samples from Brealey et al. (2020) and 29 samples from Fellows Yates et al.

644   (2021). We also included sequences from 35 published and newly generated extraction blanks,

645   twelve library preparation blanks, and four museum controls. One extraction and one library

646   blank had too few reads for taxonomic classification and were therefore excluded from

647   downstream analysis (Table S1). The museum controls consisted of two swabs taken at NRM

648   from a museum shelf surface holding reindeer (*Rangifer tarandus*) specimens and from the

649   surface of a brown bear (*Ursus arctos*) skull. In addition, we included data from two gorilla

650   specimens from previous studies (van der Valk et al., 2019, 2017): A petrous bone sample from

651   a specimen at NRM and a skin sample from RMCA (ENA accession numbers: ERR2503700

652   and ERR2868193, respectively). These sequences were generated to study the gorilla host,

653   however, we removed host reads and retained microbial reads for our analyses.

654   **Preparation of Genomic Libraries and Metagenomic Shotgun Sequencing**

655   All samples were processed in cleanroom facilities following appropriate methods for working

656   with ancient and historical DNA. DNA was extracted according to the protocol by Dabney et

657   al. (2013) with slight changes, as described in Brealey et al. (2020) and Fellows Yates et al.

658   (2021). The datasets differed in library preparation protocol, indexing strategy, sequencing

659   platform, read length, and sequencing depth (Table S1). Samples from four specimens were

660 processed in both laboratory facilities: Uppsala University, Uppsala, Sweden and the Max

661 Planck Institute for the Science of Human History, Jena, Germany (Fellows Yates et al., 2021).

662 We used these technical duplicates to assess putative batch effects (Fig. S3), but only retained

663 one sample per pair for biological analyses, selecting the one with the largest number of reads.

664 For the newly generated data, we followed the protocol as detailed in Brealey et al. (2021).

665 Briefly, dental calculus samples ranging in weight from < 5 mg and up to 20 mg were surface-

666 decontaminated using UV light (10 min at 254 nm) and washing in 500 µl of 0.5M

667 ethylenediaminetetraacetate (EDTA) for 1 min (Brealey et al., 2021, 2020; Ozga et al., 2016).

668 DNA was extracted using a silica-based method (Dabney et al., 2013) in batches of at most 16

669 samples with two negative controls. DNA was eluted in 45 µl of EB buffer (10 mM tris-

670 hydrochloride, pH 8.0; QIAGEN, Netherlands) supplemented with 0.05% (v/v) Tween-20.

671 Double-stranded genomic libraries for all newly generated and previously sequenced samples

672 were prepared following the double indexing protocol (Dabney and Meyer, 2012; Meyer and

673 Kircher, 2010). Newly generated libraries and those previously published by (Brealey et al.,

674 2020) included double in-line barcodes to guard against index hopping (van der Valk et al.,

675 2020). A detailed protocol for library preparation is provided in Brealey et al. (2020). Library

676 blanks were included for each batch of approximately 20 samples. Adapter-ligated libraries

677 were quantified by a quantitative PCR assay, allowing us to estimate an appropriate number

678 of index PCR cycles, which ranged from 8 to 20. Following indexing, individual libraries were

679 purified with Qiagen MinElute columns and quantified using a quantitative PCR assay. All

680 extraction and library blanks consistently showed lower DNA content than calculus samples

681 (Table S1). We pooled 1.5 µl of each indexed library (including blanks and museum controls)

682 and performed size selection with AMPure XP beads (Beckman Coulter, IN, USA) for

35

683   fragments of approximately 100-500 bp in length. The pooled library was sequenced by

684   SciLifeLab Uppsala on two Illumina NovaSeq S2 flowcells using paired-end 100 bp read

685   length and V1 sequencing chemistry.

686   **Preprocessing of sequencing data**

687   For each sample, the data generated from different sequencing runs was concatenated into a

688   single file (separately for reverse and forward reads) and poly-G tails, which are artefacts of

689   the two-colour sequencing chemistry of the NextSeq and NovaSeq Illumina platforms, were

690   removed using fastp (V0.20.0; Chen et al., 2018). We removed unpaired reads with BBTools

691   `repair.sh` (V38.61b; Bushnell, 2014). For newly generated data, we used AdapterRemoval

692   (V2.2.2; Schubert et al., 2016) to clip adapters, trim reads based on minimum phred quality

693   (>=30) and length (30bp), and merge forward and reverse reads. The unmerged reads were

694   excluded from downstream analysis (Table S1). In-line barcodes in newly generated data were

695   trimmed from both ends using a Python script (Brealey et al., 2020). Read quality filtering was

696   performed using PrinSeq-Lite (V0.20.4; Schmieder and Edwards, 2011) with a mean base

697   quality threshold of 30. PCR duplicates were removed using a Python script (Brealey et al.,

698   2020), which randomly kept one read among those with the same sequence. All commands and

699   scripts used in our bioinformatic analysis are available at 10.5281/zenodo.6861585.

700   Sequences of the *phi*X bacteriophage, which is used as an internal control for Illumina

701   sequencing (Mukherjee et al., 2015), were removed from the dataset by mapping against the

702   *phiX* genome (GenBank: GCA_000819615.1) using BWA-MEM (V0.7.17; Li and Durbin,

703   2009). Unmapped reads were retained using SAMtools (V1.12; Danecek et al., 2021) and

704   converted to FASTQ for downstream analysis, using the 'bamtofastq' function of BEDtools

705   (V2.29.2; Quinlan and Hall, 2010). To remove host reads and potential human contamination,

36

706   we mapped all the reads to a combined file containing western lowland gorilla and human

707   reference genomes (GenBank: GCF_000151905.2 and GCF_000001405.38, respectively),

708   repeating the steps above. Extraction blanks, library blanks, and museum controls were only

709   mapped to the human genome. The unmapped reads were used in oral microbiome and dietary

710   analyses, whereas the mapped reads were extracted, sorted, and indexed using SAMtools, and

711   used for the host genomic analyses.

712   **Host genome analysis**

713   *Mitochondrial genome analysis.* From the host mapped BAM files, we extracted all reads

714   mapping to the gorilla mitochondrial genome from position 1 to position 15,446, with a

715   mapping quality of 30 or higher, using SAMtools. Nucleotide positions 15,447-16,364, which

716   contain the D-loop, were excluded to prevent the hypervariable regions from introducing biases

717   in our analyses. We calculated sample coverage and the total number of reads mapping to the

718   mitochondrial genome with SAMtools and generated a consensus sequence in FASTA format

719   with ANGSD (V0.933; (Korneliussen et al., 2014).

720   *Subspecies assignment verification.* The subspecies identity of each sample was initially

721   assigned using museum records. We confirmed these assignments using diagnostic sites within

722   the host mitochondrial genome. Using 102 published mitochondrial genomes (Das et al., 2014;

723   Hallast et al., 2016; Hu and Gao, 2016; van der Valk et al., 2018; Xu and Arnason, 1996);

724   Table S12), we identified two sets of diagnostic sites that were fixed for different alleles in

725   each of the contrasted groups using a custom Python script. The first set was species-specific,

726   identifying membership in western versus eastern gorillas. The second set was subspecies-

727   specific, distinguishing between mountain and Grauer's gorillas.

728 *Molecular sexing of samples*. We used sexassign (Gower et al., 2019) to determine the sex of

729 each individual (Table S1). To this end, we mapped filtered reads, prior to removing host and

730 human reads, to the gorilla reference genome (GCF_000151905.2) and counted reads mapping

731 to the X chromosome and the autosomes. sexassign conducts a likelihood ratio test comparing

732 the observed X-to-autosome ratio to expected ratios for males (R_X ~0.5) and females (0.8 > R_X

733 < 1.2).

734 **Initial taxonomic classification and decontamination**

735 *Taxonomic classification with Kraken2/Bracken.* After removing host and human reads, the

736 retained unmapped reads were taxonomically classified using Kraken2 (V2.1.1; Wood et al.,

737 2019) with default parameters. We used the standard Kraken2 database (accessed by the

738 software on 1st September 2021), which includes all bacterial, archaeal and viral genomes from

739 NCBI. Abundances were re-estimated at the species level using Bracken (V2.6.2; Lu et al.,

740 2017), by redistributing reads from higher taxonomic levels to the species level, (setting read

741 length to 55 and the rest of parameters in default). The average read length was set to 55 bp,

742 which corresponded to the average of the mean read length of our data. The raw species table

743 and the corresponding metadata were combined into a phyloseq object (V1.34.0; (McMurdie

744 and Holmes, 2013) in R (V4.0.4; R Core Team, 2015). Taxonomic information was assigned

745 based on taxonomic IDs from the NCBI taxonomy database using the 'classification' function

746 of the taxize package (V0.9.99; (Chamberlain and Szöcs, 2013).

747 *Removal of low-quality samples and contaminant taxa.* Samples containing less than 300,000

748 processed reads were excluded from downstream analysis (for the distribution of read counts

749 per sample see Figure S11). We then evaluated the composition of the microbial communities

750 of the retained samples with FEAST (V0.1.0; Shenhav et al., 2019) using default parameters.

751 FEAST makes use of user-provided reference microbiomes to partition the composition of the

752 study samples and to estimate the proportional contribution of the reference "sources". We

753 provided 29 such sources sampled from six different environments that were chosen to reflect

754 both the oral cavity and potential contaminants. Specifically, we included human calculus (n=5;

755 Mann et al., 2018) and human plaque (n=5; Human Microbiome Project Consortium, 2012;

756 Lloyd-Price et al., 2017) as oral sources, and human gut (n=5; Human Microbiome Project

757 Consortium, 2012; Lloyd-Price et al., 2017), human skin (n=5; Oh et al., 2014), tundra soil

758 (n=5; Johnston et al., 2016) and laboratory contaminants excluding human sequences (n=4;

759 Salter et al., 2014) as potential contaminants (Table S13). Gorilla dental calculus samples for

760 which the proportion of the metagenomic community similar to the human oral microbiome

761 (human calculus and human plaque considered jointly) was lower than 3%, according to the

762 output of FEAST, were excluded from further analyses. The threshold was decided based on

763 the distribution of the oral proportions in the study samples (Figure S3b).

764 For the retained samples, we then carried out a multi-step approach to remove putative

765 contaminant microbial taxa. First, we used the R package decontam (V1.10.0; Davis et al.,

766 2018) with the 'isContaminant' function (method="combined", normalize = TRUE). The

767 package employs two main approaches: a prevalence-based approach, which identifies

768 contaminants based on their increased prevalence in blanks, and a frequency-based approach,

769 which relies on the observation of a reverse relationship between taxon abundance and DNA

770 quantity in the sample. We used a combination of both decontamination approaches, performed

771 separately for the subset of newly generated data, which also included two previously published

772 samples processed in the same facility (Uppsala University, Sweden; Brealey et al., 2020), and

773 data from gorilla dental calculus published by Fellows Yates et al. (2021) and processed at the

774 Max Planck Institute for the Science of Human History, Germany. We refer to these datasets

775   as 'newly generated' and Fellows Yates et al. (2021), respectively. Thresholds were allowed

776   to differ by dataset (0.2 for newly generated and 0.3 for Fellows Yates et al. (2021)) and

777   selected to minimise the number of oral taxa (taxa reported as oral in Chen et al. (2010) and

778   Fellows Yates et al. (2021)) among all taxa identified as contaminants (Figure S12). All taxa

779   identified as contaminants were then removed from the full dataset, regardless of which subset

780   they were identified in.

781   Second, we performed abundance filtering on the entire dataset, setting each taxon to 0 if it had

782   relative abundance <0.005% in a given sample. We chose this threshold, as it retained the

783   highest number and proportion of oral taxa, while excluding known contaminant taxa (Table

784   S2). Third, we considered the relative abundances of taxa in the museum controls to remove

785   likely environmental contaminants. A microbial taxon was removed as a likely contaminant if

786   it was found to have a higher relative abundance in at least one of the four museum controls

787   than in any of the samples. Finally, we considered bacterial genera identified as common

788   contaminants in typical molecular and specialised ancient DNA laboratories (Salter et al., 2014;

789   Weyrich et al., 2019). Some of the listed genera included taxa found in the Human Oral

790   Microbiome Database (HOMD; accessed July 14 2021; (Chen et al., 2010) and in the core

791   hominid oral microbiome (Fellows Yates et al., 2021). Therefore, we used a two-step approach

792   to remove these potential contaminant taxa. We directly removed genera that were listed only

793   as contaminants and did not appear in the oral databases. Taxa that were present in both

794   contaminant lists and the oral databases and had a sufficient number of reads in our dataset

795   were investigated further by running mapDamage2 (V2.0.9; Jónsson et al., 2013). To this end,

796   we identified the sample with the highest abundance of a given taxon (according to Kraken2

797   output, requiring at least 10,000 reads) and mapped the reads to the respective reference

798   genome. Taxa with too few reads to be tested were automatically retained. Since many taxa

799   needed to be investigated in this way, making a visual inspection impractical, we assessed the

800   presence of damage using a custom R script (Data and Materials). The presence of typical

801   post-mortem DNA damages was assumed if deamination (C-to-T in the 3' end and G-to-A in

802   the 5' end) was the most frequent change for at least two of the three terminal positions for

803   each fragment end and had the frequency above 0.02 for at least one position. This rule was

804   employed because in-line barcodes can create atypical DNA damage patterns in metagenomic

805   data, showing lower damage at 1st terminal position (Brealey et al., 2020). Bacterial taxa with

806   no damage were assumed to be modern contaminants and were therefore removed. The

807   resulting dataset was used for downstream taxonomic analyses.

808   *Sequence-level decontamination.* We applied additional filtering by removing sequencing

809   reads from taxa identified as contaminants (n=167) to allow for functional characterization of

810   microbial communities and reconstructing metagenome-assembled genomes (MAGs). To this

811   end, we used two approaches. First, we removed reads assigned to contaminants using the

812   'extract_kraken_reads.py' script from the KrakenTools suite (V1.2; Wood et al., 2019). Using

813   the read-level taxonomic assignments by Kraken2, we removed reads assigned to the 167

814   contaminant taxa identified above. This approach retained sequencing reads of all taxa not

815   identified as contaminants (i.e. those included in the final taxonomic dataset), but also a large

816   number of unclassified reads.

817   The reads which remained unclassified by Kraken2 could contain reads belonging to

818   contaminant taxa. Therefore, we performed a further decontamination step. We constructed a

819   dataset containing the genomes of the identified contaminant taxa and a selection of abundant

820   non-contaminant taxa from our dataset. To identify non-contaminant taxa, we selected all taxa

821   with more than 10,000 reads in at least one sample (as long as they were not identified as

41

822   contaminants), for a total of 72 unique taxa. The non-contaminant taxa were included to avoid

823   forcing the reads to map to the contaminants. For both contaminants and non-contaminants,

824   one genome per taxon was selected from the NCBI assembly summary files for bacteria,

825   archaea, and viruses (NCBI ftp accessed 14 April 2021, available at

826   https://github.com/markella-moraitou/Gorilla_dental_calculus/RD2_mapping/) with the

827   following order of priority: genome representation (full > partial), RefSeq (yes > no), assembly

828   level (full > chromosome > scaffold > contig), publication date (most recent). Twenty

829   contaminant, and 11 non-contaminant taxa did not have a reference genome available in NCBI

830   assembly. For them, we obtained a different genome from the same genus, unless that genus

831   was already represented, using the taxize R package, and selected one genome per missing

832   taxon as a replacement.

833   Finally, all references (consisting of 155 contaminant genomes and 61 non-contaminant

834   genomes) were concatenated into a combined database. We mapped the FASTQ files from the

835   previous decontamination step to this combined reference using BWA-MEM with default

836   settings and then used SAMtools to retain the unmapped reads and reads that mapped primarily

837   to the non-contaminant genomes.

838   **Taxonomic analyses**

839   All taxonomic analyses were performed at the species level. To account for the compositional

840   nature of the data, a centred-log ratio (CLR) normalisation was applied (Gloor et al., 2017)

841   using the 'transform' function of the microbiome R package (V1.12.0; Lahti and Shetty, 2018).

842   *Oral composition.* To assess differences in oral microbiome composition by subspecies and

843   sex, we used two distance measures: Jaccard distances (Jaccard, 1901), which only consider

42

844   presence-absence of taxa, and Aitchison's distances (Aitchison and Aitchison, 1986), which

845   also take into account relative abundances and are calculated as Euclidean distances on the

846   CLR-transformed data. We performed PERMANOVAs using these two distance metrics by

847   applying the 'adonis' function of the vegan package (V2.5-7; Oksanen et al., 2011) and setting

848   the parameter 'method' to 'jaccard' and 'euclidean' for the non-normalised and normalised

849   data, respectively. Because of the unequal distribution of host subspecies among the two

850   datasets (newly generated versus Fellows Yates et al. (2021); Figure S1), we included dataset

851   as a factor in each model and considered it prior to evaluating the contribution of biological

852   variables. We used the 'adonis.pair' function of the EcolUtils R package (V0.1; Salazar, 2019)

853   to perform pairwise PERMANOVAs for significant factors using Jaccard and Aitchison

854   distance matrices produced with the 'vegdist' function of the vegan package. Principal

855   Coordinate Analysis (PCoA) plots were generated using the 'plot_ordination' function of the

856   phyloseq package. Because sex assignment was missing for some samples (Table S1), we re-

857   ran PERMANOVAs for each distance measure for the subset of samples that could be sexed.

858   *Alpha diversity.* We used the 'estimate_richness' function of the phyloseq package to estimate

859   alpha diversity with two different metrics, applied to the non-normalized data: the Chao1

860   estimator (Chao et al., 2009), which evaluates species richness while accounting for the number

861   of species that are missing from the dataset, and the Shannon index (Shannon, 1948), which

862   reflects evenness. To assess the effect of the host subspecies on alpha diversity, we ran an

863   ANOVA model and confirmed the model met assumptions of normality, using a Shapiro-

864   Wilk's test on the Studentized residuals, and visually assessed the fit against the residuals

865   (Quinn and Keough, 2002). Both Chao1 and Shannon models were transformed to account for

866   skew ('sqrt(500-x)' and 'exp(x)', respectively). Finally, we used a Tukey post-hoc test to

867   compare the different levels of the factor(s) that were statistically significant.

868 *Identification of differentially abundant taxa.* We identified bacterial taxa that significantly

869 differed in abundance between host subspecies using ANCOM-II (Mandal et al., 2015), a

870 method appropriate for compositional data, which accounts for the sparsity of the taxon table

871 (Kaul et al., 2017). To account for the association between host subspecies and dataset (newly

872 generated versus Fellows Yates *et al.* (2021); Figure S1) and to reduce the possibility that we

873 identify dataset-specific taxa as differentially abundant, we ran the analysis twice. First, we

874 identified microbial taxa that differed in abundance between host subspecies, while accounting

875 for read depth. Second, we identified microbial taxa that differed in abundance between

876 datasets, while accounting for the host subspecies. Only genera that differed among subspecies

877 and did not differ by dataset were considered relevant for evaluating microbiome differences

878 among gorilla subspecies. Differentially abundant taxa were identified as those above the 0.9

879 quantile of the ANCOM W-statistic and Kruskal-Wallace post-hoc test (Bonferroni adjusted

880 p-values < 0.05), or those that were structurally absent from one or more subspecies. Only

881 microbial taxa identified at the genus or species level were considered.

882 Following this procedure, we also used ANCOM to identify taxa that differed in abundance

883 between mountain and western lowland gorillas, for both microbial and dietary taxonomic

884 datasets. These taxa were then used to explore the effects of altitude, separating Grauer's

885 gorillas into high- and low-altitude populations and performing comparisons using both

886 Jaccard and Aitchison distances in pairwise PERMANOVAs, in which differences in dataset

887 identity were included as factors.

888 **Functional analysis**

889 Decontaminated metagenomic reads were functionally characterised using an assembly-free

890 approach implemented in the HUMAnN2 pipeline (V0.11.2; Franzosa et al., 2018) using the

44

891   MetaPhlAn2, UniRef50, and ChocoPhlAn databases (Segata et al., 2012; Truong et al., 2015).

892   The output tables containing the abundances of gene families with specific function were

893   regrouped to show gene ontology (GO) terms (Ashburner et al., 2000; Gene Ontology

894   Consortium, 2021), normalised to copies per million reads and feature abundances across taxa

895   were aggregated in a single feature table, which was imported into R for further analysis. We

896   subset our dataset to GO terms referring to biological processes and used PERMANOVA on

897   their Euclidean distances to investigate functional differences of the oral microbiome by host

898   subspecies. We then compared the abundances of biological processes among host subspecies

899   using ANCOM, using the same approach as above.

900   **Metagenome-assembled genomes (MAGs)**

901   We used the metaWRAP pipeline (V1.3.2; Uritskiy et al., 2018) to construct MAGs from

902   decontaminated sequencing reads. To increase the amount of data available for building the

903   initial assemblies with MEGAHIT (Li et al., 2016), we concatenated the reads across all

904   samples. The initial assembly was split into genomic bins using three different metagenomic

905   binning tools (metabat2; V2.9.1 (Kang et al., 2019), maxbin2; V2 2.2.4 (Wu et al., 2016),

906   CONCOCT; V0.4.0 (Alneberg et al., 2013)) and we retained only those bins that were

907   consistently defined by all three tools. For each MAG, we assigned taxonomic identity with

908   GTDB-TK (Chaumeil et al., 2019) and constructed a phylogeny containing our MAGs and

909   closely-related reference genomes identified by GTDB-TK. The abundance of MAGs in each

910   sample was estimated using salmon (V0.13.1; Patro et al., 2017) and the CLR-transformed

911   abundances were visualised using pheatmap (Kolde and Others, 2012). For each assembled

912   and taxonomically identified MAG, we extracted and curated records of isolation sources

913   (sources listed in Table S8) following an approach described by Madin et al. (2020). This

45

914   information was used to categorise the isolation sources for taxa into one of the following

915   categories: "contaminant", "environmental", "host-associated", "oral", or "unknown".

916   To evaluate evolutionary relationships for select MAGs, reference genomes or gene sequences

917   for the identified species/genus were downloaded from GenBank and Genome Taxonomy

918   Database (GTDB; V202; (Parks et al., 2018). We created alignments using the amino acid

919   sequences of core genes with PhyloPhlAn (V.3.0.2; Segata et al., 2013) and constructed

920   maximum-likelihood phylogenies using FastTree (V2.1.10; Price et al., 2010), which we

921   visualised using ggtree (V3.1.0; Yu et al., 2017).

922   We used two different approaches to determine the presence of genes involved in the nitrate

923   reduction pathway in MAGs belonging to oral genera with nitrate reduction ability (see

924   references provided in Rosier et al. (2020)). First, we used the pangenomic tool Panaroo

925   (Tonkin-Hill et al., 2020) to determine presence of genes across a set of fragmented MAG

926   assemblies. Second, we performed a translated BLAST search (tblastn, V2.11.0; Camacho et

927   al., 2009) using protein sequences obtained from UniProt against MAG contigs. Genes in

928   MAGs and reference genomes were considered as present if detected with either method.

929   **Estimating ecological variation in oral microbiome composition**

930   Using altitude as a proxy for ecology and diet, we estimated the correlation between altitudinal,

931   geographical and microbiome distances. First, we approximated altitude and geographic

932   location (latitude, and longitude) for each specimen based on the collection locality as retrieved

933   from museum records. We constructed Euclidean distance matrices for altitudinal and

934   geographical distance using R package vegan. Using the ade4 R package (1.7-16; Dray et al.,

935   2007), we performed partial Mantel tests (10,000 permutations) between microbiome distance

936 matrices (Pearson correlation used for Jaccard distances and Spearman's rank coefficient

937 correlation for Aitchison taxonomic and Euclidean functional distances) and altitudinal

938 distances, while accounting for the effect of log-transformed geographical distances.

939 To further investigate the effects of altitude on Grauer's gorilla microbiomes, which is the

940 subspecies with the largest altitudinal range (Plumptre et al., 2016), we grouped the Grauer's

941 gorilla samples into low- (500-1000 m above sea level [masl]) and high-altitude (>1000 masl)

942 populations. We evaluated the contribution of altitude by performing a PERMANOVA, after

943 accounting for potential dataset and host subspecies effects. Post hoc tests were performed

944 using pairwise PERMANOVA, as described above.

945 **Dietary Analysis**

946 Eukaryotic reads in dental calculus were used to identify taxa potentially consumed by gorillas.

947 Before taxonomically classifying these reads, we removed reads assigned to prokaryotes and

948 viruses from the metagenomic dataset using the 'extract_kraken_reads.py' script from

949 KrakenTools. The resulting FASTQ files were used as input in a second run of Kraken2 using

950 the NCBI 'nt' database (accessed 27 September 2021). A joint feature table containing species

951 and genus-level assignments was produced using kraken-biom and imported into R for further

952 analysis. The genome coverage for these taxonomic classifications was insufficient for

953 verification using mapDamage2.

954 *Preprocessing of feature table.* The feature table produced by kraken-biom was used in a

955 phyloseq object along with a corresponding taxonomic matrix, created using the taxize R

956 package. Using this matrix, we aggregated all identified taxa to the genus level, and used the

957 decontam (V.1.10.0; Davis et al., 2018) R package to remove likely contaminants, as described

47

958   above for the oral microbiome analyses, including using different cut-offs for each dataset.

959   Then, all genera present in the museum controls or with fewer than 10 reads per sample were

960   removed, to reduce contaminants, spurious and unreliable classifications. Despite this filtering,

961   spurious classifications or contaminants were still present (species that are unlikely to be

962   consumed by gorillas, e.g. mollusks), so we restricted it to genera belonging to taxonomic

963   families that were previously reported to be consumed by gorillas (Michel et al., 2022; Remis

964   et al., 2001; Rogers et al., 2004; Rothman et al., 2014; Yamagiwa et al., 2005). Since only few

965   reference genomes of tropical plants are available and classically-used barcoding genes

966   represent only a small part of the genome, accurate taxonomic classification from a sparse

967   metagenomic sample is difficult and can likely only be achieved on genus or family level

968   (Mann et al., 2020). For this reason, we focused our analysis not only on genera known to be

969   consumed by gorillas, but also included all detected genera in the same family. The resulting

970   feature table was used to detect differentially abundant dietary components with ANCOM

971   using the same approach as described above, while accounting for differences in the number

972   of reads per sample.

973   *Collecting gorilla diet reference data.* We produced a reference dataset of gorilla foods (Table

974   S11) by searching the literature for plant species known to be consumed by the three gorilla

975   subspecies based either on direct observations, analyses of food remains and faeces or more

976   recently on molecular evidence (Michel et al., 2022; Remis et al., 2001; Rogers et al., 2004;

977   Rothman et al., 2014; Yamagiwa et al., 2005). We retrieved the corresponding taxonomic IDs

978   from NCBI using the taxize R package. For the taxa that had no match in the NCBI Taxonomy

979   database, we manually searched the database and corrected the names that were obviously

980   misspelt (difference of 1-2 characters from the NCBI entry). For the remaining taxa, we

981   manually searched the GBIF species database (Wheeler, 2004) and obtained GBIF taxonomic

48

982 IDs. However, several taxa (26 out of 314) could not be found in either database and therefore

983 were excluded from our analysis. The taxonomic IDs obtained from either NCBI or GBIF were

984 used with the 'classification' function of taxize to obtain the genus and family name for each

985 taxon.

Availability of Data and Materials

987 Sequencing data generated in this study have been deposited on ENA, under Project Accession

988 Number: PRJEB49638. Scripts for both preprocessing and downstream analysis are available

989 at 10.5281/zenodo.6861585.

990

Funding

998

Authors' contributions

1000 JCB, MM, AF extracted and prepared genomic libraries for the newly generated data. MM

1001 processed sequencing data, performed bioinformatic and statistical analyses (microbiome

49

1002   taxonomic and functional analyses, dietary and host genome analyses, altitudinal analyses) and

1003   wrote the manuscript draft. AF performed bioinformatic and statistical analyses (metagenome-

1004   assembled genomes, functional analyses, altitudinal analyses), wrote the respective parts of the

1005   manuscript and contributed to others. JCB performed initial bioinformatics analyses on a

1006   smaller dataset. JAFY and CW provided genomic data. KG conceived the research plan,

1007   interpreted the results, provided financial support and infrastructure, and contributed

1008   significantly to manuscript writing. All authors contributed to writing the final version of this

1009   manuscript.

1010

Acknowledgments

References

1025 Abusleme L, Dupuy AK, Dutzan N, Silva N, Burleson JA, Strausbaugh LD, Gamonal J, Diaz
1026   PI. 2013. The subgingival microbiome in health and periodontitis and its relationship
1027   with community biomass and inflammation. *ISME J* **7**:1016–1025.
1028   doi:10.1038/ismej.2012.174
1029 Abusleme L, O'Gorman H, Dutzan N, Greenwell-Wild T, Moutsopoulos NM. 2020.
1030   Establishment and Stability of the Murine Oral Microbiome. *J Dent Res* **99**:721–729.
1031   doi:10.1177/0022034520915485
1032 Adler CJ, Dobney K, Weyrich LS, Kaidonis J, Walker AW, Haak W, Bradshaw CJA,
1033   Townsend G, Sołtysiak A, Alt KW, Parkhill J, Cooper A. 2013. Sequencing ancient
1034   calcified dental plaque shows changes in oral microbiota with dietary shifts of the
1035   Neolithic and Industrial revolutions. *Nat Genet* **45**:450–5, 455e1. doi:10.1038/ng.2536
1036 Adler CJ, Malik R, Browne GV, Norris JM. 2016. Diet may influence the oral microbiome
1037   composition in cats. *Microbiome* **4**:23. doi:10.1186/s40168-016-0169-y
1038 Aitchison J, Aitchison JW. 1986. The Statistical Analysis of Compositional Data. Springer
1039   Netherlands.
1040 Alberdi A, Aizpurua O, Bohmann K, Zepeda-Mendoza ML, Gilbert MTP. 2016. Do
1041   Vertebrate Gut Metagenomes Confer Rapid Ecological Adaptation? *Trends Ecol Evol*
1042   **31**:689–699. doi:10.1016/j.tree.2016.06.008
1043 Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Loman NJ, Andersson
1044   AF, Quince C. 2013. CONCOCT: Clustering cONtigs on COverage and ComposiTion.
1045   *arXiv [q-bioGN]*.
1046 Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance.
1047   *Austral Ecol* **26**:32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x
1048 Arora T, Sharma R. 2011. Fermentation potential of the gut microbiome: implications for
1049   energy homeostasis and weight management. *Nutr Rev* **69**:99–106. doi:10.1111/j.1753-
1050   4887.2010.00365.x
1051 Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Michael Cherry J, Davis AP,
1052   Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A,
1053   Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene
1054   Ontology: tool for the unification of biology. *Nature Genetics*. doi:10.1038/75556
1055 Beall CM, Laskowski D, Erzurum SC. 2012. Nitric oxide in adaptation to altitude. *Free*
1056   *Radic Biol Med* **52**:1123–1134. doi:10.1016/j.freeradbiomed.2011.12.028
1057 Beall CM, Laskowski D, Strohl KP, Soria R, Villena M, Vargas E, Alarcon AM, Gonzales C,
1058   Erzurum SC. 2001. Pulmonary nitric oxide in mountain dwellers. *Nature* **414**:411–412.
1059   doi:10.1038/35106641
1060 Boehlke C, Rupf S, Tenniswood M, Chittur SV, Hannig C, Zierau O. 2020. Caries and
1061   periodontitis associated bacteria are more abundant in human saliva compared to other
1062   great apes. *Arch Oral Biol* **111**:104648. doi:10.1016/j.archoralbio.2020.104648
1063 Bo T-B, Zhang X-Y, Kohl KD, Wen J, Tian S-J, Wang D-H. 2020. Coprophagy prevention
1064   alters microbiome, metabolism, neurochemistry, and cognitive behavior in a small
1065   mammal. *The ISME Journal*. doi:10.1038/s41396-020-0711-6
1066 Brealey JC, Leitão HG, Hofstede T, Kalthoff DC, Guschanski K. 2021. The oral microbiota
1067   of wild bears in Sweden reflects the history of antibiotic use by humans. *Curr Biol*
1068   **31**:4650–4658.e6. doi:10.1016/j.cub.2021.08.010
1069 Brealey JC, Leitão HG, van der Valk T, Xu W, Bougiouri K, Dalén L, Guschanski K. 2020.
1070   Dental Calculus as a Tool to Study the Evolution of the Mammalian Oral Microbiome.

1071   *Mol Biol Evol* **37**:3003–3022. doi:10.1093/molbev/msaa135

1072   Briggs AW, Stenzel U, Johnson PLF, Green RE, Kelso J, Prüfer K, Meyer M, Krause J,
1073   Ronan MT, Lachmann M, Pääbo S. 2007. Patterns of damage in genomic DNA
1074   sequences from a Neandertal. *Proc Natl Acad Sci U S A* **104**:14616–14621.
1075   doi:10.1073/pnas.0704665104

1076   Bryan NS, Tribble G, Angelov N. 2017. Oral Microbiome and Nitric Oxide: the Missing Link
1077   in the Management of Blood Pressure. *Curr Hypertens Rep* **19**:33. doi:10.1007/s11906-
1078   017-0725-2

1079   Bushnell B. 2014. BBMap: A fast, accurate, splice-aware aligner (No. LBNL-7065E).
1080   Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).

1081   Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009.
1082   BLAST+: architecture and applications. *BMC Bioinformatics* **10**:421. doi:10.1186/1471-
1083   2105-10-421

1084   Cephas KD, Kim J, Mathai RA, Barry KA, Dowd SE, Meline BS, Swanson KS. 2011.
1085   Comparative Analysis of Salivary Bacterial Microbiome Diversity in Edentulous Infants
1086   and Their Mothers or Primary Care Givers Using Pyrosequencing. *PLoS ONE*.
1087   doi:10.1371/journal.pone.0023503

1088   Chamberlain SA, Szöcs E. 2013. taxize: taxonomic search and retrieval in R. *F1000Res*
1089   **2**:191. doi:10.12688/f1000research.2-191.v2

1090   Chao A, Colwell RK, Lin C-W, Gotelli NJ. 2009. Sufficient sampling for asymptotic
1091   minimum species richness estimators. *Ecology* **90**:1125–1133. doi:10.1890/07-2147.1

1092   Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify
1093   genomes with the Genome Taxonomy Database. *Bioinformatics*.
1094   doi:10.1093/bioinformatics/btz848

1095   Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.
1096   *Bioinformatics* **34**:i884–i890. doi:10.1093/bioinformatics/bty560

1097   Chen T, Yu W-H, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. 2010. The Human
1098   Oral Microbiome Database: a web accessible resource for investigating oral microbe
1099   taxonomic and genomic information. *Database* **2010**:baq013.
1100   doi:10.1093/database/baq013

1101   Cobb CM, Kelly PJ, Williams KB, Babbar S, Angolkar M, Derman RJ. 2017. The oral
1102   microbiome and adverse pregnancy outcomes. *Int J Womens Health* **9**:551–559.
1103   doi:10.2147/IJWH.S142730

1104   Corby PM, Bretz WA, Hart TC, Schork NJ, Wessel J, Lyons-Weiler J, Paster BJ. 2007.
1105   Heritability of oral microbial species in caries-active and caries-free twins. *Twin Res*
1106   *Hum Genet* **10**:821–828. doi:10.1375/twin.10.6.821

1107   Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, García
1108   N, Pääbo S, Arsuaga J-L, Meyer M. 2013. Complete mitochondrial genome sequence of
1109   a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl*
1110   *Acad Sci U S A* **110**:15758–15763. doi:10.1073/pnas.1314445110

1111   Dabney J, Meyer M. 2012. Length and GC-biases during sequencing library amplification: a
1112   comparison of various polymerase-buffer systems with ancient and modern DNA
1113   sequencing libraries. *Biotechniques* **52**:87–94. doi:10.2144/000113809

1114   Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,
1115   McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools.
1116   *Gigascience* **10**. doi:10.1093/gigascience/giab008

1117   Das R, Hergenrother SD, Soto-Calderón ID, Dew JL, Anthony NM, Jensen-Seaman MI.
1118   2014. Complete mitochondrial genome sequence of the Eastern gorilla (Gorilla beringei)

52

1119   and implications for african ape biogeography. *J Hered* **105**:752–761.
1120   doi:10.1093/jhered/esu056
1121   Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical
1122   identification and removal of contaminant sequences in marker-gene and metagenomics
1123   data. *Microbiome* **6**:226. doi:10.1186/s40168-018-0605-2
1124   Delamuta JRM, Scherer AJ, Ribeiro RA, Hungria M. 2020. Genetic diversity of
1125   Agrobacterium species isolated from nodules of common bean and soybean in Brazil,
1126   Mexico, Ecuador and Mozambique, and description of the new species Agrobacterium
1127   fabacearum sp. nov. *Int J Syst Evol Microbiol* **70**:4233–4244.
1128   doi:10.1099/ijsem.0.004278
1129   Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R.
1130   2010. Delivery mode shapes the acquisition and structure of the initial microbiota across
1131   multiple body habitats in newborns. *Proc Natl Acad Sci U S A* **107**:11971–11975.
1132   doi:10.1073/pnas.1002601107
1133   Donati C, Zolfo M, Albanese D, Tin Truong D, Asnicar F, Iebba V, Cavalieri D, Jousson O,
1134   De Filippo C, Huttenhower C, Segata N. 2016. Uncovering oral Neisseria tropism and
1135   persistence using metagenomic sequencing. *Nat Microbiol* **1**:16070.
1136   doi:10.1038/nmicrobiol.2016.70
1137   Doran DM, McNeilage A, Greer D, Bocian C, Mehlman P, Shah N. 2002. Western lowland
1138   gorilla diet and resource availability: new evidence, cross-site comparisons, and
1139   reflections on indirect sampling methods. *Am J Primatol* **58**:91–116.
1140   doi:10.1002/ajp.10053
1141   Dray S, Dufour A-B, Others. 2007. The ade4 package: implementing the duality diagram for
1142   ecologists. *J Stat Softw* **22**:1–20.
1143   Eriksson H, Bagge E, Båverud V, Fellström C, Jansson DS. 2014. Erysipelothrix
1144   rhusiopathiae contamination in the poultry house environment during erysipelas
1145   outbreaks in organic laying hen flocks. *Avian Pathol* **43**:231–237.
1146   doi:10.1080/03079457.2014.907485
1147   Erzurum SC, Ghosh S, Janocha AJ, Xu W, Bauer S, Bryan NS, Tejero J, Hemann C, Hille R,
1148   Stuehr DJ, Feelisch M, Beall CM. 2007. Higher blood flow and circulating NO products
1149   offset high-altitude hypoxia among Tibetans. *Proc Natl Acad Sci U S A* **104**:17593–
1150   17598. doi:10.1073/pnas.0707462104
1151   Feelisch M. 2018. Enhanced nitric oxide production is a universal response to hypoxic stress.
1152   *National Science Review*. doi:10.1093/nsr/nwy041
1153   Fellows Yates JA, Velsko IM, Aron F, Posth C, Hofman CA, Austin RM, Parker CE, Mann
1154   AE, Nägele K, Arthur KW, Arthur JW, Bauer CC, Crevecoeur I, Cupillard C, Curtis
1155   MC, Dalén L, Díaz-Zorita Bonilla M, Díez Fernández-Lomana JC, Drucker DG,
1156   Escribano Escrivá E, Francken M, Gibbon VE, González Morales MR, Grande Mateu
1157   A, Harvati K, Henry AG, Humphrey L, Menéndez M, Mihailović D, Peresani M,
1158   Rodríguez Moroder S, Roksandic M, Rougier H, Sázelová S, Stock JT, Straus LG,
1159   Svoboda J, Teßmann B, Walker MJ, Power RC, Lewis CM, Sankaranarayanan K,
1160   Guschanski K, Wrangham RW, Dewhirst FE, Salazar-García DC, Krause J, Herbig A,
1161   Warinner C. 2021. The evolution and changing ecology of the African hominid oral
1162   microbiome. *Proc Natl Acad Sci U S A* **118**. doi:10.1073/pnas.2021655118
1163   Fletcher A, Nowell A. 2008. The development of feeding behaviour in wild western lowland
1164   gorillas (Gorilla gorilla gorilla). *Behaviour*. doi:10.1163/156853907783244747
1165   Flynn KJ, Baxter NT, Schloss PD. 2016. Metabolic and Community Synergy of Oral Bacteria
1166   in Colorectal Cancer. *mSphere* **1**. doi:10.1128/mSphere.00102-16

53

1167   Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, Lipson KS,
1168   Knight R, Caporaso JG, Segata N, Huttenhower C. 2018. Species-level functional
1169   profiling of metagenomes and metatranscriptomes. *Nat Methods* **15**:962–968.
1170   doi:10.1038/s41592-018-0176-y

1171   Galbany J, Imanizabayo O, Romero A, Vecellio V, Glowacka H, Cranfield MR, Bromage
1172   TG, Mudakikwa A, Stoinski TS, McFarlin SC. 2016. Tooth wear and feeding ecology in
1173   mountain gorillas from Volcanoes National Park, Rwanda. *Am J Phys Anthropol*
1174   **159**:457–465. doi:10.1002/ajpa.22897

1175   Ganas J, Robbins MM, Nkurunungi JB, Kaplin BA, McNeilage A. 2004. Dietary variability
1176   of mountain gorillas in bwindi impenetrable National Park, Uganda. *Int J Primatol*
1177   **25**:1043–1072. doi:10.1023/b:ijop.0000043351.20129.44

1178   Gao S, Li S, Ma Z, Liang S, Shan T, Zhang M, Zhu X, Zhang P, Liu G, Zhou F, Yuan X, Jia
1179   R, Potempa J, Scott DA, Lamont RJ, Wang H, Feng X. 2016. Presence of
1180   Porphyromonas gingivalis in esophagus and its association with the clinicopathological
1181   characteristics and survival in patients with esophageal cancer. *Infectious Agents and*
1182   *Cancer*. doi:10.1186/s13027-016-0049-x

1183   Gene Ontology Consortium. 2021. The Gene Ontology resource: enriching a GOld mine.
1184   *Nucleic Acids Res* **49**:D325–D334. doi:10.1093/nar/gkaa1113

1185   Gerner-Smidt P, Keiser-Nielsen H, Dorsch M, Stackebrandt E, Ursing J, Blom J, Christensen
1186   AC, Christensen JJ, Frederiksen W, Hoffmann S. 1994. Lautropia mirabilis gen. nov.,
1187   sp. nov., a gram-negative motile coccus with unusual morphology isolated from the
1188   human mouth. *Microbiology* **140 (Pt 7)**:1787–1797. doi:10.1099/13500872-140-7-1787

1189   Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. 2017. Microbiome Datasets Are
1190   Compositional: And This Is Not Optional. *Front Microbiol* **8**:2224.
1191   doi:10.3389/fmicb.2017.02224

1192   Gower G, Fenderson LE, Salis AT, Helgen KM, van Loenen AL, Heiniger H, Hofman-
1193   Kamińska E, Kowalczyk R, Mitchell KJ, Llamas B, Cooper A. 2019. Widespread male
1194   sex bias in mammal fossil and museum collections. *Proc Natl Acad Sci U S A*
1195   **116**:19019–19024. doi:10.1073/pnas.1903275116

1196   Graczyk TK, Cranfield MR. 2003. Coprophagy and intestinal parasites: implications to
1197   human-habituated mountain gorillas (Gorilla gorilla beringei) of the Virunga Mountains
1198   and Bwindi Impenetrable Forest. *Primate Conserv* **9**:58–64.

1199   Haas FB, Eisenhofer R, Schreiber M, Weyrich LS, Rensing SA. 2022. Neanderthals did not
1200   likely consume Physcomitrium patens – a model moss species.
1201   doi:10.1101/2022.01.26.472964

1202   Hallast P, Maisano Delser P, Batini C, Zadik D, Rocchi M, Schempp W, Tyler-Smith C,
1203   Jobling MA. 2016. Great ape Y Chromosome and mitochondrial DNA phylogenies
1204   reflect subspecies structure and patterns of mating and dispersal. *Genome Res* **26**:427–
1205   439. doi:10.1101/gr.198754.115

1206   Hezel MP, Weitzberg E. 2015. The oral microbiome and nitric oxide homoeostasis. *Oral Dis*
1207   **21**:7–16. doi:10.1111/odi.12157

1208   Human Microbiome Project Consortium. 2012. Structure, function and diversity of the
1209   healthy human microbiome. *Nature* **486**:207–214. doi:10.1038/nature11234

1210   Hu X-D, Gao L-Z. 2016. The complete mitochondrial genome of eastern lowland gorilla,
1211   Gorilla beringei graueri, and comparative mitochondrial genomics of Gorilla species.
1212   *Mitochondrial DNA A DNA Mapp Seq Anal* **27**:1484–1485.
1213   doi:10.3109/19401736.2014.953103

1214   Hyde ER, Andrade F, Vaksman Z, Parthasarathy K, Jiang H, Parthasarathy DK, Torregrossa

1215 AC, Tribble G, Kaplan HB, Petrosino JF, Bryan NS. 2014a. Metagenomic analysis of
1216   nitrate-reducing bacteria in the oral cavity: implications for nitric oxide homeostasis.
1217   *PLoS One* **9**:e88645. doi:10.1371/journal.pone.0088645
1218 Hyde ER, Luk B, Cron S, Kusic L, McCue T, Bauch T, Kaplan H, Tribble G, Petrosino JF,
1219   Bryan NS. 2014b. Characterization of the rat oral microbiome and the effects of dietary
1220   nitrate. *Free Radic Biol Med* **77**:249–257. doi:10.1016/j.freeradbiomed.2014.09.017
1221 Jaccard P. 1901. Étude comparative de la distribution florale dans une portion des Alpes et
1222   des Jura. *Bull Soc Vaud sci nat* **37**:547–579.
1223 Janiak MC, Montague MJ, Villamil CI, Stock MK, Trujillo AE, DePasquale AN, Orkin JD,
1224   Bauman Surratt SE, Gonzalez O, Platt ML, Martínez MI, Antón SC, Dominguez-Bello
1225   MG, Melin AD, Higham JP. 2021. Age and sex-associated variation in the multi-site
1226   microbiome of an entire social group of free-ranging rhesus macaques. *Microbiome*
1227   **9**:68. doi:10.1186/s40168-021-01009-w
1228 Jiao J-Y, Liu L, Hua Z-S, Fang B-Z, Zhou E-M, Salam N, Hedlund BP, Li W-J. 2021.
1229   Microbial dark matter coming to light: challenges and opportunities. *Natl Sci Rev*
1230   **8**:nwaa280. doi:10.1093/nsr/nwaa280
1231 Johnston ER, Rodriguez-R LM, Luo C, Yuan MM, Wu L, He Z, Schuur EAG, Luo Y, Tiedje
1232   JM, Zhou J, Konstantinidis KT. 2016. Metagenomics Reveals Pervasive Bacterial
1233   Populations and Reduced Community Diversity across the Alaska Tundra Ecosystem.
1234   *Front Microbiol* **7**:579. doi:10.3389/fmicb.2016.00579
1235 Jones AM, Vanhatalo A, Seals DR, Rossman MJ, Piknova B, Jonvik KL. 2021. Dietary
1236   Nitrate and Nitric Oxide Metabolism: Mouth, Circulation, Skeletal Muscle, and Exercise
1237   Performance. *Med Sci Sports Exerc* **53**:280–294. doi:10.1249/MSS.0000000000002470
1238 Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast
1239   approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*
1240   **29**:1682–1684. doi:10.1093/bioinformatics/btt193
1241 Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive
1242   binning algorithm for robust and efficient genome reconstruction from metagenome
1243   assemblies. *PeerJ* **7**:e7359. doi:10.7717/peerj.7359
1244 Kartzinel TR, Hsing JC, Musili PM, Brown BRP, Pringle RM. 2019. Covariation of diet and
1245   gut microbiome in African megafauna. *Proc Natl Acad Sci U S A* **116**:23588–23593.
1246   doi:10.1073/pnas.1905666116
1247 Kaul A, Mandal S, Davidov O, Peddada SD. 2017. Analysis of Microbiome Data in the
1248   Presence of Excess Zeros. *Front Microbiol* **8**:2114. doi:10.3389/fmicb.2017.02114
1249 Kolde R, Others. 2012. Pheatmap: pretty heatmaps. *R package version* **1**:726.
1250 Kolenbrander PE, Palmer RJ Jr, Periasamy S, Jakubovics NS. 2010. Oral multispecies
1251   biofilm development and the key role of cell-cell distance. *Nat Rev Microbiol* **8**:471–
1252   480. doi:10.1038/nrmicro2381
1253 Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation
1254   Sequencing Data. *BMC Bioinformatics* **15**:356. doi:10.1186/s12859-014-0356-4
1255 Lahti L, Shetty S. 2018. Introduction to the microbiome R package.
1256   http://bioconductor.statistik.tu-
1257   dortmund.de/packages/3.6/bioc/vignettes/microbiome/inst/doc/vignette.html
1258 Levett DZ, Fernandez BO, Riley HL, Martin DS, Mitchell K, Leckstrom CA, Ince C, Whipp
1259   BJ, Mythen MG, Montgomery HE, Grocott MP, Feelisch M, Caudwell Extreme Everest
1260   Research Group. 2011. The role of nitrogen oxides in human adaptation to hypoxia. *Sci*
1261   *Rep* **1**:109. doi:10.1038/srep00109
1262 Levin D, Raab N, Pinto Y, Rothschild D, Zanir G, Godneva A, Mellul N, Futorian D, Gal D,

1263   Leviatan S, Zeevi D, Bachelet I, Segal E. 2021. Diversity and functional landscapes in
1264   the microbiota of animals in the wild. *Science* **372**. doi:10.1126/science.abb5352
1265   Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, Yamashita H, Lam T-W. 2016.
1266   MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced
1267   methodologies and community practices. *Methods* **102**:3–11.
1268   doi:10.1016/j.ymeth.2016.02.020
1269   Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler
1270   transform. *Bioinformatics* **25**:1754–1760. doi:10.1093/bioinformatics/btp324
1271   Li J, Nasidze I, Quinque D, Li M, Horz H-P, André C, Garriga RM, Halbwax M, Fischer A,
1272   Stoneking M. 2013. The saliva microbiome of Pan and Homo. *BMC Microbiol* **13**:204.
1273   doi:10.1186/1471-2180-13-204
1274   Li Y, Ismail AI, Ge Y, Tellez M, Sohn W. 2007. Similarity of Bacterial Populations in Saliva
1275   from African-American Mother-Child Dyads. *Journal of Clinical Microbiology*.
1276   doi:10.1128/jcm.00771-07
1277   Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, Brady A, Creasy HH,
1278   McCracken C, Giglio MG, McDonald D, Franzosa EA, Knight R, White O,
1279   Huttenhower C. 2017. Strains, functions and dynamics in the expanded Human
1280   Microbiome Project. *Nature* **550**:61–66. doi:10.1038/nature23889
1281   Lu J, Breitwieser FP, Thielen P, Salzberg SL. 2017. Bracken: estimating species abundance
1282   in metagenomics data. *PeerJ Comput Sci* **3**:e104. doi:10.7717/peerj-cs.104
1283   Lundberg JO, Carlström M, Weitzberg E. 2018. Metabolic Effects of Dietary Nitrate in
1284   Health and Disease. *Cell Metab* **28**:9–22. doi:10.1016/j.cmet.2018.06.007
1285   Madin JS, Nielsen DA, Brbic M, Corkrey R, Danko D, Edwards K, Engqvist MKM, Fierer
1286   N, Geoghegan JL, Gillings M, Kyrpides NC, Litchman E, Mason CE, Moore L, Nielsen
1287   SL, Paulsen IT, Price ND, Reddy TBK, Richards MA, Rocha EPC, Schmidt TM,
1288   Shaaban H, Shukla M, Supek F, Tetu SG, Vieira-Silva S, Wattam AR, Westfall DA,
1289   Westoby M. 2020. A synthesis of bacterial and archaeal phenotypic trait data. *Sci Data*
1290   **7**:170. doi:10.1038/s41597-020-0497-4
1291   Maisels F, Bergl RA, Williamson EA. 2018. Gorilla gorilla (amended version of 2016
1292   assessment). *The IUCN Red List of Threatened Species* 2018–2012.
1293   Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R, Peddada SD. 2015. Analysis of
1294   composition of microbiomes: a novel method for studying microbial composition.
1295   *Microbial Ecology in Health & Disease*. doi:10.3402/mehd.v26.27663
1296   Mann AE, Fellows Yates JA, Fagernäs Z, Austin RM, Nelson EA, Hofman CA. 2020. Do I
1297   have something in my teeth? The trouble with genetic analyses of diet from
1298   archaeological dental calculus. *Quat Int*. doi:10.1016/j.quaint.2020.11.019
1299   Mann AE, Sabin S, Ziesemer K, Vågene ÅJ, Schroeder H, Ozga AT, Sankaranarayanan K,
1300   Hofman CA, Fellows Yates JA, Salazar-García DC, Frohlich B, Aldenderfer M,
1301   Hoogland M, Read C, Milner GR, Stone AC, Lewis CM Jr, Krause J, Hofman C, Bos
1302   KI, Warinner C. 2018. Differential preservation of endogenous human and microbial
1303   DNA in dental calculus and dentin. *Sci Rep* **8**:9822. doi:10.1038/s41598-018-28091-9
1304   Marcelino, R, Vanessa, Holmes EC, Sorrell TC. 2020. The use of taxon-specific reference
1305   databases compromises metagenomic classification. *BMC Genomics* **21**:184.
1306   doi:10.1186/s12864-020-6592-2
1307   Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. 2016. Biogeography of a
1308   human oral microbiome at the micron scale. *Proceedings of the National Academy of*
1309   *Sciences* **113**:E791–E800.
1310   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

Nearing JT, DeClercq V, Van Limbergen J, Langille MGI. 2020. Assessing the Variation within the Oral Microbiome of Healthy Adults. *mSphere* **5**. doi:10.1128/mSphere.00451-20

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

57

1359 Ozga AT, Nieves-Colón MA, Honap TP, Sankaranarayanan K, Hofman CA, Milner GR,
1360   Lewis CM Jr, Stone AC, Warinner C. 2016. Successful enrichment and recovery of
1361   whole mitochondrial genomes from ancient human dental calculus. *Am J Phys*
1362   *Anthropol* **160**:220–228. doi:10.1002/ajpa.22960
1363 Ozga AT, Ottoni C. 2021. Dental calculus as a proxy for animal microbiomes. *Quat Int*.
1364   doi:10.1016/j.quaint.2021.06.012
1365 Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz
1366   P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially
1367   revises the tree of life. *Nat Biotechnol* **36**:996–1004. doi:10.1038/nbt.4229
1368 Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-
1369   aware quantification of transcript expression. *Nat Methods* **14**:417–419.
1370   doi:10.1038/nmeth.4197
1371 Plumptre A, Nixon S, Caillaud D, Hall JS, Hart JA, Nishuli R, Williamson EA. 2016. Gorilla
1372   beringei ssp. graueri. The IUCN Red List of Threatened Species 2016: e.
1373   T39995A102328430 DOI 10.2305/IUCN. UK. 2016-2. *RLTS T39995A17989838 en*.
1374 Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: from saprophytes to endosymbionts.
1375   *Nat Rev Microbiol* **16**:291–303. doi:10.1038/nrmicro.2017.171
1376 Price MN, Dehal PS, Arkin AP. 2010. FastTree 2--approximately maximum-likelihood trees
1377   for large alignments. *PLoS One* **5**:e9490. doi:10.1371/journal.pone.0009490
1378 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic
1379   features. *Bioinformatics* **26**:841–842. doi:10.1093/bioinformatics/btq033
1380 Quinn GP, Keough MJ. 2002. Experimental Design and Data Analysis for Biologists.
1381   doi:10.1017/cbo9780511806384
1382 Radini A, Nikita E, Buckley S, Copeland L, Hardy K. 2017. Beyond food: The multiple
1383   pathways for inclusion of materials into ancient dental calculus. *Am J Phys Anthropol*
1384   **162 Suppl 63**:71–83. doi:10.1002/ajpa.23147
1385 R Core Team. 2015. R: A Language and Environment for Statistical Computing. Vienna,
1386   Austria: R Foundation for Statistical Computing.
1387 Remis MJ, Dierenfeld ES, Mowry CB, Carroll RW. 2001. Nutritional aspects of western
1388   lowland gorilla (Gorilla gorilla gorilla) diet during seasons of fruit scarcity at Bai
1389   Hokou, Central African Republic. *Int J Primatol* **22**:807–836.
1390 Rogers ME, Abernethy K, Bermejo M, Cipolletta C, Doran D, McFarland K, Nishihara T,
1391   Remis M, Tutin CEG. 2004. Western gorilla diet: a synthesis from six sites. *Am J*
1392   *Primatol* **64**:173–192. doi:10.1002/ajp.20071
1393 Rosier BT, Moya-Gonzalvez EM, Corell-Escuin P, Mira A. 2020. Isolation and
1394   Characterization of Nitrate-Reducing Bacteria as Potential Probiotics for Oral and
1395   Systemic Health. *Front Microbiol* **11**:555465. doi:10.3389/fmicb.2020.555465
1396 Rosier BT, Takahashi N, Zaura E, Krom BP, MartÍnez-Espinosa RM, van Breda SGJ, Marsh
1397   PD, Mira A. 2022. The Importance of Nitrate Reduction for Oral Health. *J Dent Res*
1398   220345221080982. doi:10.1177/00220345221080982
1399 Rothman JM, Nkurunungi JB, Shannon BF, Bryer MAH. 2014. High Altitude Diets:
1400   Implications for the Feeding and Nutritional Ecology of Mountain Gorillas In: Grow
1401   NB, Gursky-Doyen S, Krzton A, editors. High Altitude Primates. New York, NY:
1402   Springer New York. pp. 247–264. doi:10.1007/978-1-4614-8175-1_14
1403 Roy J, Arandjelovic M, Bradley BJ, Guschanski K, Stephens CR, Bucknell D, Cirhuza H,
1404   Kusamba C, Kyungu JC, Smith V, Robbins MM, Vigilant L. 2014. Recent divergences
1405   and size decreases of eastern gorilla populations. *Biol Lett* **10**:20140811.
1406   doi:10.1098/rsbl.2014.0811

58

1407 Ruff CB, Junno J-A, Burgess ML, Canington SL, Harper C, Mudakikwa A, McFarlin SC.
1408   2022. Body proportions and environmental adaptation in gorillas. *American Journal of*
1409   *Biological Anthropology* **177**:501–529. doi:10.1002/ajpa.24443
1410 Salazar G. 2019. EcolUtils: Utilities for community ecology analysis. *R Package Version 0 1*
1411   *Available online: https://github com/GuillemSalazar/EcolUtils (accessed on 22 October*
1412   *2019)*.
1413 Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J,
1414   Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically
1415   impact sequence-based microbiome analyses. *BMC Biol* **12**:87. doi:10.1186/s12915-
1416   014-0087-z
1417 Sawafuji R, Saso A, Suda W, Hattori M, Ueda S. 2020. Ancient DNA analysis of food
1418   remains in human dental calculus from the Edo period, Japan. *PLoS One* **15**:e0226654.
1419   doi:10.1371/journal.pone.0226654
1420 Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets.
1421   *Bioinformatics* **27**:863–864. doi:10.1093/bioinformatics/btr026
1422 Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming,
1423   identification, and read merging. *BMC Res Notes* **9**:88. doi:10.1186/s13104-016-1900-2
1424 Segata N, Börnigen D, Morgan XC, Huttenhower C. 2013. PhyloPhlAn is a new method for
1425   improved phylogenetic and taxonomic placement of microbes. *Nat Commun* **4**:2304.
1426   doi:10.1038/ncomms3304
1427 Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. 2012.
1428   Metagenomic microbial community profiling using unique clade-specific marker genes.
1429   *Nat Methods* **9**:811–814. doi:10.1038/nmeth.2066
1430 Shaiber A, Willis AD, Delmont TO, Roux S, Chen L-X, Schmid AC, Yousef M, Watson AR,
1431   Lolans K, Esen ÖC, Lee STM, Downey N, Morrison HG, Dewhirst FE, Mark Welch JL,
1432   Eren AM. 2020. Functional and genetic markers of niche partitioning among enigmatic
1433   members of the human oral microbiome. *Genome Biol* **21**:292. doi:10.1186/s13059-020-
1434   02195-w
1435 Shannon CE. 1948. A mathematical theory of communication. *The Bell System Technical*
1436   *Journal* **27**:379–423. doi:10.1002/j.1538-7305.1948.tb01338.x
1437 Shenhav L, Thompson M, Joseph TA, Briscoe L, Furman O, Bogumil D, Mizrahi I, Pe'er I,
1438   Halperin E. 2019. FEAST: fast expectation-maximization for microbial source tracking.
1439   *Nat Methods* **16**:627–632. doi:10.1038/s41592-019-0431-x
1440 Smith SN, Colston TJ, Siler CD. 2021. Venomous Snakes Reveal Ecological and
1441   Phylogenetic Factors Influencing Variation in Gut and Oral Microbiomes. *Front*
1442   *Microbiol* **12**:657754. doi:10.3389/fmicb.2021.657754
1443 Soares-Castro P, Araújo-Rodrigues H, Godoy-Vitorino F, Ferreira M, Covelo P, López A,
1444   Vingada J, Eira C, Santos PM. 2019. Microbiota fingerprints within the oral cavity of
1445   cetaceans as indicators for population biomonitoring. *Sci Rep* **9**:13679.
1446   doi:10.1038/s41598-019-50139-7
1447 Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL. 1998. Microbial complexes in
1448   subgingival plaque. *Journal of Clinical Periodontology*. doi:10.1111/j.1600-
1449   051x.1998.tb02419.x
1450 Suzuki TA. 2017. Links between Natural Variation in the Microbiome and Host Fitness in
1451   Wild Mammals. *Integr Comp Biol* **57**:756–769. doi:10.1093/icb/icx104
1452 Takenoshita Y, Yamagiwa J. 2008. Estimating gorilla abundance by dung count in the
1453   northern part of Moukalaba-Doudou National Park, Gabon. *African Study Monographs*
1454   *Supplementary Issue* **39**:41–54.

1455 Teles R, Wang C-Y. 2011. Mechanisms involved in the association between peridontal
1456   diseases and cardiovascular disease. *Oral Diseases*. doi:10.1111/j.1601-
1457   0825.2010.01784.x
1458 Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, Gladstone RA, Lo
1459   S, Beaudoin C, Floto RA, Frost SDW, Corander J, Bentley SD, Parkhill J. 2020.
1460   Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol*
1461   **21**:180. doi:10.1186/s13059-020-02090-4
1462 Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower
1463   C, Segata N. 2015. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat*
1464   *Methods* **12**:902–903. doi:10.1038/nmeth.3589
1465 Tsuchida S, Kakooza S, Mbehang Nguema PP, Wampande EM, Ushida K. 2018.
1466   Characteristics of Gorilla-Specific Lactobacillus Isolated from Captive and Wild
1467   Gorillas. *Microorganisms* **6**. doi:10.3390/microorganisms6030086
1468 Tsuchida S, Kitahara M, Nguema PPM, Norimitsu S, Fujita S, Yamagiwa J, Ngomanda A,
1469   Ohkuma M, Ushida K. 2014. Lactobacillus gorillae sp. nov., isolated from the faeces of
1470   captive and wild western lowland gorillas (Gorilla gorilla gorilla). *Int J Syst Evol*
1471   *Microbiol* **64**:4001–4006. doi:10.1099/ijs.0.068429-0
1472 Tsuchida S, Nezuo M, Tsukahara M, Ogura Y, Hayashi T, Ushida K. 2015. Draft Genome
1473   Sequence of Lactobacillus gorillae Strain KZ01T, Isolated from a Western Lowland
1474   Gorilla. *Genome Announc* **3**. doi:10.1128/genomeA.01196-15
1475 Tsuzukibashi O, Uchibori S, Kobayashi T, Umezawa K, Mashimo C, Nambu T, Saito M,
1476   Hashizume-Takizawa T, Ochiai T. 2017. Isolation and identification methods of Rothia
1477   species in oral cavities. *J Microbiol Methods* **134**:21–26.
1478   doi:10.1016/j.mimet.2017.01.005
1479 Turner LA, Bucking C. 2019. The role of intestinal bacteria in the ammonia detoxification
1480   ability of teleost fish. *J Exp Biol* **222**. doi:10.1242/jeb.209882
1481 Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP—a flexible pipeline for genome-
1482   resolved metagenomic data analysis. *Microbiome* **6**:158. doi:10.1186/s40168-018-0541-
1483   1
1484 van der Hoek Y, Pazo WD, Binyinyi E, Ngobobo U, Stoinski TS, Caillaud D. 2021. Diet of
1485   Grauer's Gorillas (Gorilla beringei graueri) in a Low-Elevation Forest. *Folia Primatol*
1486   **92**:126–138. doi:10.1159/000515377
1487 van der Valk T, Díez-Del-Molino D, Marques-Bonet T, Guschanski K, Dalén L. 2019.
1488   Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in
1489   Eastern Gorillas. *Curr Biol* **29**:165–170.e6. doi:10.1016/j.cub.2018.11.055
1490 van der Valk T, Lona Durazo F, Dalén L, Guschanski K. 2017. Whole mitochondrial genome
1491   capture from faecal samples and museum-preserved specimens. *Mol Ecol Resour*
1492   **17**:e111–e121. doi:10.1111/1755-0998.12699
1493 van der Valk T, Sandoval-Castellanos E, Caillaud D, Ngobobo U, Binyinyi E, Nishuli R,
1494   Stoinski T, Gilissen E, Sonet G, Semal P, Kalthoff DC, Dalén L, Guschanski K. 2018.
1495   Significant loss of mitochondrial diversity within the last century due to extinction of
1496   peripheral populations in eastern gorillas. *Sci Rep* **8**:6551. doi:10.1038/s41598-018-
1497   24497-7
1498 van der Valk T, Vezzi F, Ormestad M, Dalén L, Guschanski K. 2020. Index hopping on the
1499   Illumina HiseqX platform and its consequences for ancient DNA studies. *Mol Ecol*
1500   *Resour* **20**:1171–1181. doi:10.1111/1755-0998.13009
1501 Velsko IM, Fellows Yates JA, Aron F, Hagan RW, Frantz LAF, Loe L, Martinez JBR,
1502   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

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,

61

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