1   **Hunter-gatherer oral microbiomes are shaped by contact network structure**

2   Federico Musciotto[1,2+], Begoña Dobon[2,3+], Michael Greenacre[4,5], Alex Mira[6,7], Abigail E. Page[8],

3   Mark Dyble[9], Sylvain Viguier[9], Daniel Smith[10], Nikhil Chaudhary[11], Gul Deniz Salali[9], Pascale

4   Gerbault[12], Rodolph Schlaepfer[2], Leonora H. Astete[13], Marilyn Ngales[13], Jesus Gomez-

5   Gardenes[14,15], Vito Latora[16,17,18], Federico Battiston[2,19], Jaume Bertranpetit[3#], Lucio Vinicius[2,9#],

6   Andrea Bamberg Migliano[2,9#]

7   [1]Dipartimento di Fisica e Chimica, Università di Palermo; Palermo, Italy.

8   [2]Department of Anthropology, University of Zurich; Zurich, Switzerland.

9   [3]Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra); Barcelona, Spain.

10   [4]Department of Economics and Business, Universitat Pompeu Fabra & Barcelona Graduate
11   School of Economics; Barcelona, Spain.

12   [5]Faculty of Biosciences, Fisheries and Economics, University of Tromsø; Norway

13   [6]Department of Health and Genomics, Center for Advanced Research in Public Health, FISABIO
14   Foundation; Valencia, Spain.

15   [7]CIBER Center for Epidemiology and Public Health; Madrid, Spain

16   [8]Department of Population Health, London School of Hygiene and Tropical Medicine; London,
17   United Kingdom.

18   [9]Department of Anthropology, University College London; London, United Kingdom.

19   [10]Bristol Medical School, University of Bristol; Bristol, United Kingdom.

20   [11]Department of Archaeology, University of Cambridge; Cambridge, United Kingdom.

21   [12]Department of Genetics and Evolution, University of Geneva, Geneva, Switzerland.

22   [13]Lyceum of the Philippines University, Intramuros, Manila, Philippines.

23   [14]GOTHAM Lab, Institute for Biocomputation and Physics of Complex Systems, and
24   Department of Condensed Matter Physics, University of Zaragoza; Zaragoza, Spain.

25   [15]Center for Computational Social Science (CCSS), Kobe University; Kobe, Japan.

26 [16]School of Mathematical Sciences, Queen Mary University of London; London, United
27 Kingdom.

28 [17]Dipartimento di Fisica ed Astronomia, Università di Catania and INFN; Catania, Italy.

29 [18]Complexity Science Hub Vienna (CSHV); Vienna, Austria.

30 [19]Department of Network and Data Science, Central European University; Vienna, Austria.

31

32 [#]equal contribution

33 +equal contribution

34 Corresponding author: andrea.migliano@uzh.ch jaume.bertranpetit@upf.edu

35

36

37 **Abstract**

38 **Ancestral humans evolved a complex social structure still observed in extant hunter-**

39 **gatherers. Here we investigate the effects of extensive sociality and mobility on the oral**

40 **microbiome of 138 Agta hunter-gatherers from the Philippines. Comparisons of**

41 **microbiome composition showed that the Agta are more similar to Central African Bayaka**

42 **hunter-gatherers than to neighboring farmers. We also defined the Agta social microbiome**

43 **as a set of 137 oral bacteria (only 7% of 1980 amplicon sequence variants) significantly**

44 **influenced by social contact (quantified through wireless sensors of short-range**

45 **interactions). We show that interaction networks covering large areas, and their strong**

46 **links between close kin, spouses, and even unrelated friends, can significantly predict**

47 **bacterial transmission networks across Agta camps. Finally, more central individuals to**

48 **social networks are also bacterial supersharers. We conclude that hunter-gatherer social**

49   **microbiomes, which are predominantly pathogenic, were shaped by evolutionary tradeoffs**

50   **between extensive sociality and disease spread.**

51

Introduction

53   Hominins have significantly diverged from other African apes regarding social behaviour and

54   structure[1]. Compared to polygynous mating and male philopatric residence patterns typically

55   found in chimpanzees, bonobos and gorillas, archaeological and ethnographic evidence point to a

56   stepwise emergence of features such as pair bonding, multilocal residence, high mobility

57   between residential camps and increased co-residence with unrelated individuals[2,3]. Such traits

58   were the foundations of multilevel social structuring appearing in ancestral *Homo sapiens* and

59   possibly earlier hominins. The niche of extant hunter-gatherers may offer a window into past

60   human adaptations as it still exhibits features prevalent before the advent of agriculture, such as a

61   high-quality diet including meat and tubers, and multilevel sociality. Multilevel organization

62   results in interconnected social networks covering large areas and multiple residential camps[4],

63   and in frequent interactions between individuals differing by sex, age and relatedness level.

64   Interconnected networks may have accelerated the evolution of cultural innovations in humans

65   compared to other apes[5,6]. However, efficient networks may also facilitate the spread of

66   infectious diseases[7], potentially affecting the structure and composition of hunter-gatherer

67   microbiomes. Previous studies have investigated the role of diet, ecology and environment in

68   hunter-gatherer oral, gut and milk microbiomes[8-16] and revealed higher oral microbiome

69   diversity in hunter-gatherers than in farming populations[17]. However, they have not been able to

70   isolate the contribution of high sociality and individual mobility to microbial transmission from

3

71   other factors such as shared environments or diet. Although the more fluid and complex sociality

72   of hunter-gatherers results in high levels of camp coresidence[2], cooperation and social

73   interactions among unrelated individuals[18], its potential effects on microbiome transmission have

74   been mostly neglected. We conducted a comprehensive investigation of the oral microbiome of

75   Agta hunter-gatherers to analyse the specific effect of sociality and social network structure on

76   the composition of the Agta oral microbiome, with a companion article examining the separate

77   role of environmental (diet) and biological (age, sex, host genotype) factors[19].

78   We obtained both oral microbiome sequences and high-resolution social network data

79   from the same 138 Agta hunter-gatherers from the Philippines. We also collected oral

80   microbiome data for 21 Bayaka hunter-gatherers from the Congo, and 14 Palanan farmers

81   neighboring the Agta territory. We sequenced the 16S rRNA region and identified 6409

82   amplicon sequence variants (ASVs)[20], later reduced to 1980 ASVs (with at least 10 counts and

83   present in at least two individuals), to detect fine-scale variation between individuals. We also

84   collected data on proximity interactions and social networks using radio sensor technology

85   recording close-range dyadic interactions every two minutes for 5-7 days[6,18] from four Agta

86   camps, and from two longer multi-camp experiments (interactions recorded every hour for one

87   month). Proximity data were supplemented with information on household composition, kinship

88   and affinal relationships from all Agta individuals.

89   Our extensive dataset on oral microbiome composition and social interactions from the

90   same individuals allowed us to investigate in more depth the possible effects of sociality on oral

91   microbiome transmission and composition in Agta hunter-gatherers. Our aims were to

92   investigate the roles of hunter-gatherer niche and geography on oral microbiome diversity in

93   hunter-gatherers from two continents and a neighboring farming population from the Philippines;

94   to determine which fraction of the Agta oral microbiome specifically responds to levels of social

95   interaction; to identify levels of pathogenicity of the oral microbiome transmitted through social

96   contact; to investigate any potential tradeoffs between increased sociality and the spread of

97   infectious disease; and to verify potential tradeoffs at individual level by testing whether 'hyper

98   social' individuals also shared more bacteria. In the following, we provide evidence that the oral

99   microbiome of extant hunter-gatherers was partially shaped by tradeoffs between extensive

100 sociality and the spread of infectious disease.

101

102 **Results**

103 **Hunter gatherer niches shape the oral microbiome.** To investigate the contributions of

104 lifestyle versus environment to the hunter-gatherer oral microbiome, we compared the Agta

105 (n=138) to smaller samples of Bayaka hunter-gatherers from the Congo (n=21) and neighboring

106 Palanan farmers from the Philippines (n=14) (see Methods). Both Agta (mean of 252.1±90 ASVs

107 per individual) and Bayaka (280.1±83) exhibited significantly more ASVs than Palanan farmers

108 (163.4±34) (P<0.0001; Figure 1a), and higher levels of ASV diversity as measured by Faith's

109 Phylogenetic Diversity index (Figure 1b). Comparisons based on the total set of ASVs in each

110 population (controlling for differences in sample size through subsampling) revealed that the

111 Agta shared more bacteria with African Bayaka (471.2±33.9) than with neighboring Palanan

112 farmers (423.4±23.8) (Figure 1c). Finally, Agta and Bayaka resampled groups showed

113 respectively 651.5±93.7 and 688.1±61.9 exclusive ASVs, against only 285.1±16.9 in Palanan

114 farmers (Figure 1d). In summary, the two hunter-gatherer populations show higher microbiome

115 diversity and uniqueness than Palanan farmers, consistent with findings that farming

116   significantly reduced gut microbiome diversity[21]. Results therefore demonstrate the precedence

117   of niche over geography in shaping hunter-gatherer oral microbiomes.
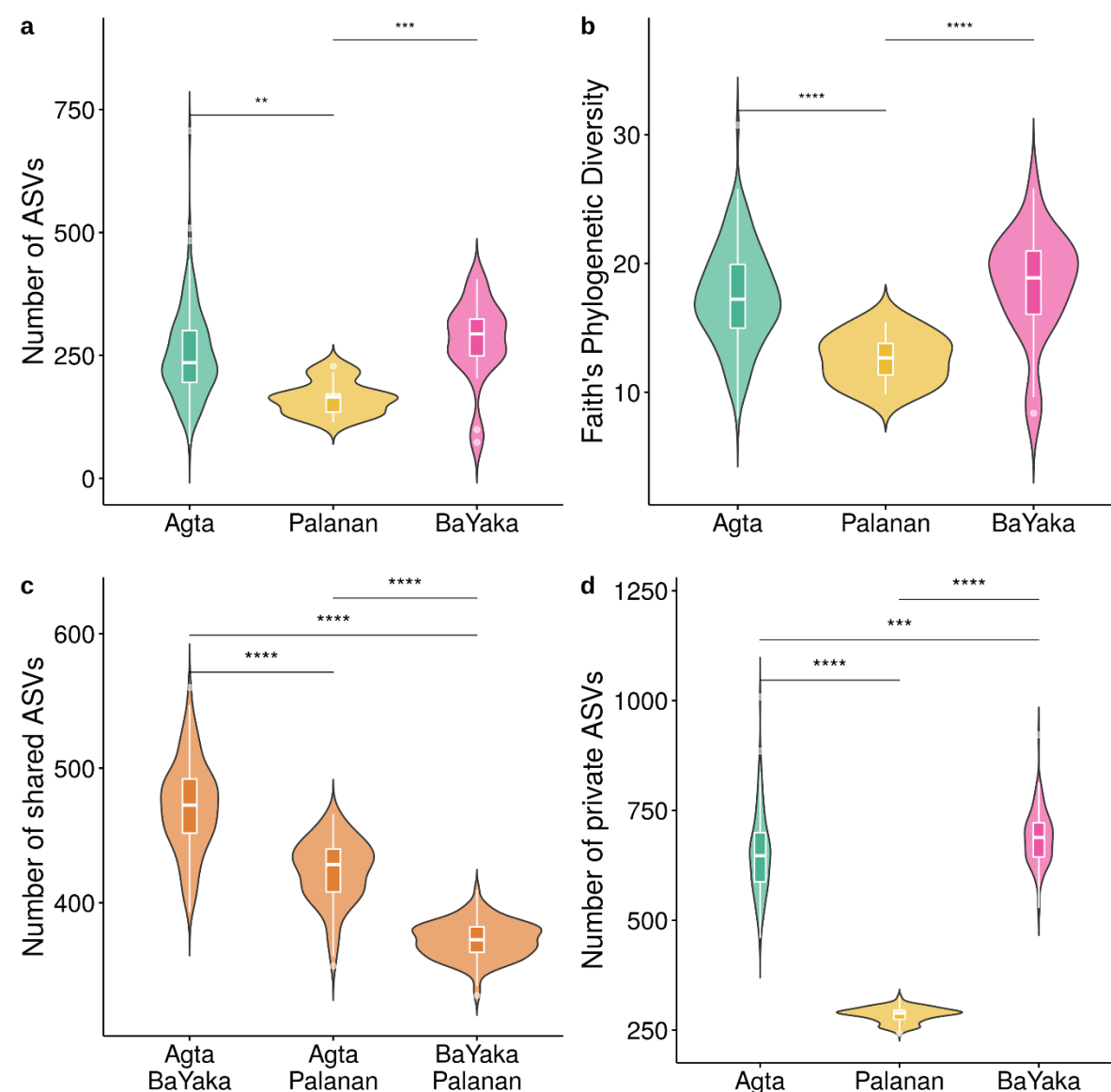
118

119

120   **Figure 1. Oral microbiome diversity in Agta hunter-gatherers, neighbouring Palanan**

121   **farmers, and Bayaka hunter-gatherers.** a) Number of ASVs in the Agta (n=138), Bayaka

122   (n=21) and Palanan (n=14). b) Oral microbiome diversity assessed by Faith's Phylogenetic

6

123   Diversity index accounting for ASV phylogenetic distances (Agta=17.52±3.83;

124   Bayaka=18.45±4.10; Palanan farmers=12.62±1.86); c) Shared ASVs between populations,

125   estimated by randomly sampling 10 individuals from each population (averaged over 100

126   permutations); d) Exclusive ASVs per individual, estimated by randomly sampling 10

127   individuals from each population (100 permutations). Boxplot midlines represent medians, and

128   box limits represent first and third quartiles (****: FDR-adjusted P<0.0001; ***: P<0.001; **:

129   P<0.01).

130   **The social microbiome is a socially transmitted fraction of the oral microbiome**. Primate

131   social 'pan-microbiomes' were recently defined as the totality of microorganisms present in a

132   host population or species[22], but this definition also includes microorganisms acquired due to

133   common diet or environment. Here, we define the 'social microbiome' as the oral microbiome

134   specifically transmitted through social interactions. To identify the socially transmitted fraction

135   of the Agta oral microbiome, we used the contact network recorded through radio sensor devices

136   and split all Agta dyadic social interactions into a strong (top 25% from the distribution of dyadic

137   link weights) and a weak set (the remaining 75% links; see Methods). We then tested for

138   differences in the proportion of each of the 1980 ASVs between the strong and weak sets. We

139   identified 137 ASVs (7% of the Agta oral microbiome; see Supplementary Figure 1 and

140   Supplementary Table 1) whose presence was significantly higher in the strong set, and therefore

141   statistically associated with higher frequencies of social interactions. In the following we

142   investigate the transmission patterns and composition of the hunter-gatherer social microbiome.

143

144   **The hunter-gatherer social microbiome is predominantly pathogenic.** Human sociality is

145   associated with multiple fitness benefits, including increased reproductive success[7], reputation[23],

146   food sharing[24], cooperation[3,25] and cultural transmission[26], but may also facilitate pathogen

147   transmission[7,18]. In our dataset, from the 18 ASVs that could be classified at species level, 14 are

148   socially transmitted, 9 of which (64.3%) are typically pathogenic, and 10 (71.4%) are typically

149   oral. By contrast, all four non-socially transmitted species were non-pathogenic and typically

150   oral.

151   We were able to classify 1886 of the 1980 ASVs at genus level, resulting in 36 socially

152   (those representing ASVs included in the social microbiome) and 62 non-socially transmitted

153   genera (the remaining ones). Among the social genera, 61.8% were classified as typically or

154   exclusively pathogenic (21 out of 34; two genera could not be classified), against only 16.4%

155   among non-socially transmitted genera (10 out of 61; one genus could not be classified). We

156   identified many socially transmitted genera either typically (*Aggregatibacter*, *Capnocytophaga*)

157   or uniquely (*Corynebacterium*) associated with dental plaque formation, gingivitis and calculus,

158   the full red complex of periodontal disease (*Porphyromonas*, *Treponema* and *Tannerella*), and

159   other potential periodontal pathogens (*Prevotella*, *Desulfobulbus*, *Fusobacterium*)[27–30]. The

160   classification of bacterial genera as pathogenic is not unequivocal for those cases where some

161   species within the genus can be pathogenic and others non-pathogenic. Thus, practical criteria

162   were applied in these cases for assigning a genus to the pathogenic group (see Methods).

163   Following these criteria, the social microbiome clearly has a higher proportion of pathogenic

164   organisms than the non-socially transmitted portion.

165   We also found pathogenic bacteria typical of the gut (*Rickenellaceae*), non-human

166   environments (*Tetragenococcus*, *Comamonas*), respiratory tract (*Staphylococcus*, *Moraxella*,

167   *Streptococcus pneumoniae*), and both urogenital and respiratory tracts (*Mycoplasma*), suggesting

168   that their spread may be facilitated by oral transmission[31–33]. In summary, the predominantly

169   pathogenic nature of the social microbiome suggests a trade-off between benefits of hunter-

170   gatherer sociality and costs associated with disease transmission.

171   **Hunter-gatherer multilevel social structure shapes social microbiome sharing.** Hunter-

172   gatherer sociality is characterised by specific interaction channels not found in non-human apes,

173   such as long-term pair bonding and households, extended families, friendships among unrelated

174   individuals, and frequent between-camp relocation. We estimated the effect of relatedness level,

175   residence camp and friendships on the probability of sharing socially transmitted bacteria. First,

176   we built a bacterial sharing network, where the weight of each Agta dyadic link is given by how

177   many of the 137 social bacteria are shared by the two Agta individuals (rather than by the

178   strength of its social bond, as in the social network). Next, we classified all dyadic links in this

179   network into: i) levels of kinship (mother-offspring, father-offspring, siblings: r=0.5, other kin:

180   r=0.25 or r=0.125, non-kin: r=0.0625 or lower, spouses, friends: defined as non-kin at the top

181   25% distribution of social dyadic weights, and other non-kin) and ii) residence (same or different

182   camp, same or different household). Finally, we compared the mean weight of each type of

183   dyadic link in our bacterial sharing network to its mean weight in a sample of 1000 networks of

184   the same size and topology, but where the dyadic classification was randomised (Figure 2 and

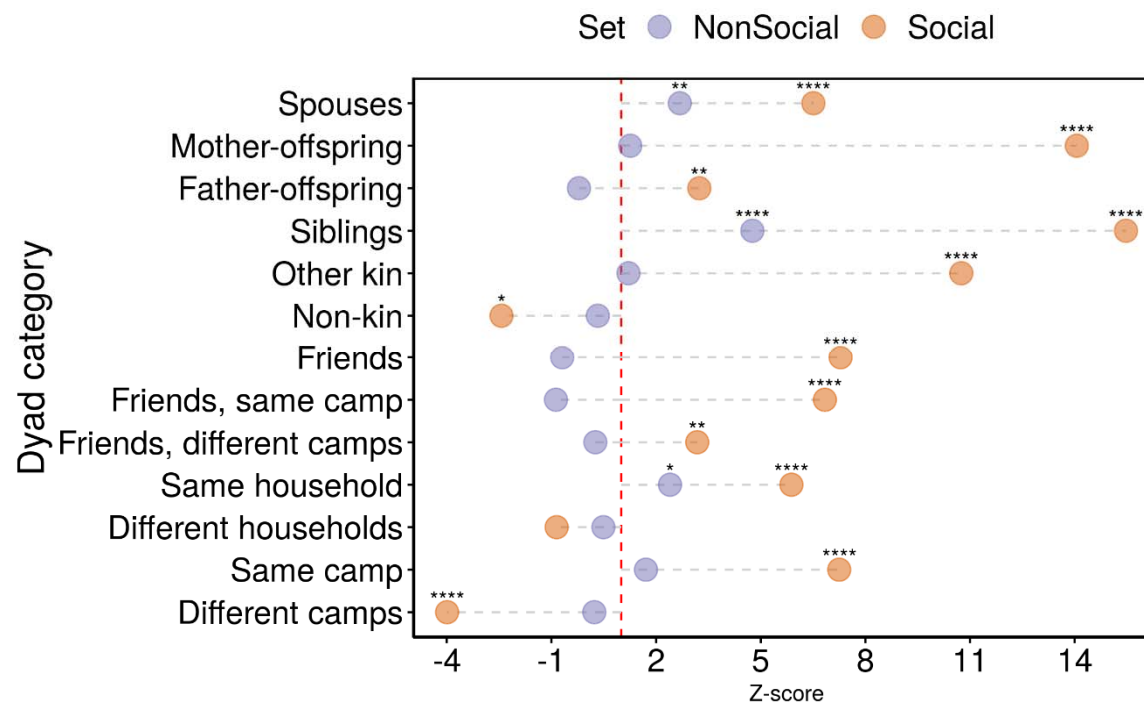185   Supplementary Tables 2 and 3).

Figure 2. Effect of kinship, friendship and residence on dyadic bacterial sharing. Dyads were classified into kinship levels; same or different households; same or different camps; and between friends in the same or different camps. Dots show the z-score, or the standardised ratio of the mean link weight in real to randomised networks, in either social (orange) or non-socially transmitted bacteria (purple). Vertical red dashed line indicates a ratio of 1, or no difference between the number of shared bacteria in real and randomised networks. For socially transmitted bacteria, kinship, friendship and residence in the same household or camp are associated with significantly higher bacterial sharing than predicted from randomised networks of the same size and structure. By contrast, dyads from different camps or non-kin share significantly fewer bacteria than expected by chance; bacterial sharing in dyads from different households do not differ from randomised networks. For non-socially transmitted bacteria, the only dyadic categories significantly increasing bacterial sharing were siblings, spouses, and dyads from the same household (all of which share the same close environment). See Supplementary Tables 2 and 3 for values on mean weights for real and randomised networks. (****: FDR-adjusted $P<0.0001$; ***: $P<0.001$; **: $P<0.01$).

Results showed that some dyadic categories share significantly more socially transmitted bacteria than expected by chance. First, we observed higher bacterial sharing within the same household and camp than between different households and camps, an expected consequence of

10

206   the Agta multilevel social structure. Kinship effects were also clear, with the highest levels of

207   social microbiome sharing found in mother-offspring pairs, followed by siblings known interact

208   every day in households and playgroups. High sharing between spouses also confirmed the

209   importance of human pair bonding in microbial transmission. In addition, strong friendship links

210   were also associated with increased bacterial sharing. Social bacterial sharing between friends in

211   the same camp is as high as between close kin or within households. Friends in different camps

212   also share a higher proportion of social bacteria than expected by chance, which is possibly a

213   consequence of high between-camp mobility. By contrast, non-kin or individuals from different

214   camps share fewer social bacteria than expected, further demonstrating the role of Agta

215   friendships in the transmission of social bacteria across households and whole camps.

216   The same analysis performed instead on non-socially transmitted ASVs did not reveal

217   significant effects on bacterial sharing from most dyadic categories, except for three types:

218   spouses, siblings, and same household. A possible explanation is that some non-socially

219   transmitted bacteria may be shared due to a common environment and diet in the same

220   household. For example, we have shown in our parallel study[19] that the proportion of meat

221   versus rice in individual diets affects the composition of the oral microbiome. Therefore, similar

222   diets may explain the presence of the same ASVs within the individuals of a household

223   irrespective of social interaction levels. However, the effects of sociality and shared diets seem to

224   be independent. This is shown by the fact that socially transmitted bacteria are equally likely to

225   be related or not to diet: 13 socially transmitted genera were found also associated to diet (41

226   genera), whereas 23 were not (57 genera) (proportion test: chi-squared=0.44, P=0.51; See

227   companion paper[19] for further data and analysis). There we also show that host genotype

228   correlates with the presence of certain ASVs19. While high genetic relatedness may play a role

229   in bacteria sharing within households, none of the ASVs associated with the host genotype were

230   present in the social microbiome. Therefore, our analyses seem to distinguish between the effects

231   of social contact from shared environment or genes within dyadic types. Overall, the results

232   show the roles of mobility and the multiple interaction channels created by multilevel sociality in

233   social microbiome sharing, similarly to what is also observed in cultural transmission[6,18,26,34].

234

235   **Frequency of social contact predicts social microbiome sharing**. Although previous studies

236   have investigated patterns of bacterial sharing in human groups, they have often been unable to

237   comprehensively characterize transmission patterns due to limited information on social

238   contact[35]. In order to obtain a full picture of individual contact and exposure levels, we built

239   social networks based on proximity data from four camps and two multi-camp locations (Figure

240   3a-b). Overall, Agta social networks reveal a multilevel structure of households (mostly

241   consisting of strong kin links) connected by a few strong links (mostly among unrelated friends)

242   in each camp, and in the case of multi-camp groups, camps interconnected mostly due to visits

243   among friends. We also observed equality of interactions within and between sexes and across

244   age groups[6,18]. This pattern creates multiple channels for social transmission of bacteria both

245   within and between camps, between close kin and unrelated individuals, and finally across whole
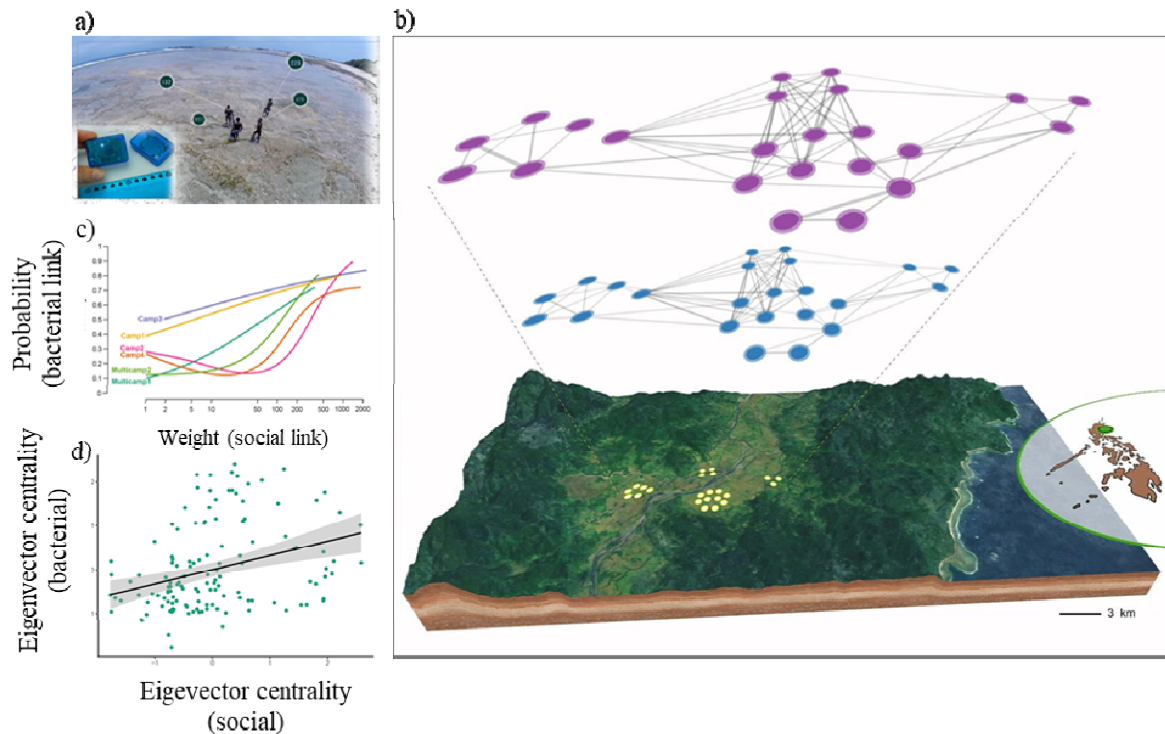
246   multi-camp structures.

Figure 3. Characterisation of the social microbiome. a) Recording networks of social interactions using radio sensor technology. b) Reinforcement analysis estimates the probability of a link occurring in the bacterial sharing network (top layer, purple) based on the weight of the same link in the social contact network (bottom layer, blue). Network nodes (circles) represent the same Agta individuals in the bacterial sharing and social contact network. Panel displays networks from multi-camp 1 (23 individuals). Map shows geographical location of four camps interconnected by frequent migration. c) Probabilities of links in the Agta bacterial sharing network increase with their weights in the social contact network. Curves estimated by generalised additive modelling (binomial option). Data from four Agta camps and two multi-camps. d) Eigenvector centralities in bacterial sharing and social contact networks. Linear regression based on pooled data from four Agta camps and two multi-camp structures. Virtually similar results were obtained by including camp either as a fixed factor in a multiple regression (with or without interactions), or as a random factor (on intercept and slope) in a mixed effects linear regression.

To further assess whether increasing levels of social contact predict higher levels of sharing of socially transmitted bacteria, we applied reinforcement analysis[36] (Figure 3b) to assess

13

265   whether the social network predicts (or reinforces) the bacterial sharing network in each Agta

266   camp. We calculated the conditional probability of each link between two individuals A and B in

267   the bacterial network provided the same weighted link is present in the social network

268   (Supplementary Table 4 and 5). For all four camps and two multi-camp structures, results

269   showed that the weight of a dyadic link in the social network significantly predicts the

270   probability of the same link occurring in the bacterial sharing network (Figure 3c and

271   Supplementary Figure 2). Specifically, a larger dyadic weight in the social network implies a

272   higher probability that the same individuals also share at least one socially transmitted ASV. For

273   example, for multi-camp 1, while weak social network links (with weights under 10 recorded

274   social interactions) show a probability below 20% of sharing any socially transmitted bacteria,

275   strong links (over 200 recorded contacts) are associated with a probability above 70%. Overall,

276   the results confirm that the Agta social microbiome is shaped by their social interactions.

277

278   **Hypersocial individuals are supersharers**. We also investigated whether more socially

279   interactive individuals exhibited higher social microbiome diversity. We calculated eigenvector

280   centralities for all individuals in the bacterial sharing network, resulting in a significant and

281   positive slope in a regression on eigenvector centralities from the same individuals in the social

282   network (b=0.32, P=0.0001, $R^2 = 0.1$, n=138, Figure 3d). We also identified 16 individuals

283   ranked at the top quartile of eigenvector centralities in both networks as potential microbial

284   'superspreaders' or 'superacquirers', that is, "supersharers". They do not stem from a specific

285   age (7 to 68 years) or sex (six males, ten females), which is compatible with the egalitarian social

286   structure of hunter-gatherers allowing individuals from any age or sex to be potentially central to

287   social networks.

Discussion

We have identified and characterized a socially transmitted fraction of the Agta hunter-gatherer oral microbiome. This fraction (7%) is surprisingly small, since in principle all 1980 identified ASVs could be orally transmitted between closely interacting people. Nonetheless, our results demonstrate a significant and independent role of social interactions on the transmission of oral microbiome, in addition to other factors such as shared environment (household and diet) and host characteristics (age, sex and genes) previously investigated in other hunter-gatherer populations[8–16] and in the same Agta population[19]. The transmission of the 137 bacteria classified into the social microbiome seem to be facilitated by the extended sociality of hunter-gatherers and its various transmission channels, ranging from spouses to unrelated friends often residing in different camps. Together, reinforcement analysis, multiple channels of social interaction, and supersharers show that social microbiome sharing is strongly shaped by hunter-gatherer multilevel sociality. From an evolutionary perspective, sociality has considerably changed from our closest ape relatives to ancestral humans, when adopting a hunter-gathering lifestyle meant exhibited higher frequency of social contact with unrelated individuals, larger networks of extended kin across large geographical regions, and more egalitarian interactions between and within sexes and across ages. Such changes may have affected patterns of pathogen transmission and affected the human microbiome as observed in current hunter-gatherers. As with our study of the Agta, future research should collect data on both social networks and social microbiomes from the same populations of non-human apes; such a dataset would provide a comparative basis for analysing of the role of social evolution on the human social microbiome. Hunter-gatherer social networks are efficient systems of cultural transmission, and its specific channels organised around kinship, friendship and camp interconnectivity are central for the organisation of

15

311   between-household cooperation, food sharing and social learning[24,26]. Agta mothers with higher

312   social network centrality enjoy increased access to help and reproductive success[7], but our results

313   have shown that efficient networks may also facilitate the spread of infectious diseases, and

314   hence significantly affect the structure and composition of the Agta microbiome. Crucially the

315   frequency of pathogenic bacteria is much higher in the socially than in the non-socially

316   transmitted fraction of the Agta oral microbiome. Together with the association between

317   hypersocial individuals and increased bacterial sharing, this suggests a tradeoff between potential

318   fitness benefits and costs of increased pathogen transmission. We conclude that the

319   predominantly pathogenic oral social microbiome we identified in hunter-gatherers may be at

320   least partially the outcome of a tradeoff between the advantages of multilevel sociality and the

321   cost of infectious disease.

322

323

324 **Methods**

325 **Ethnographic data collection**

326 *Agta demography*. Ethnographic data collection took place over two seasons in April-June 2013

327 and February-October 2014. We censused 915 Agta individuals (54.7% male) across 20 camps.

328 For the current study we selected four camps and two multi-camp structures where we collected

329 data both on proximity networks and saliva samples. Accurate ages were estimated following

330 relative aging protocols[37]. Relatedness (biological and affinal) was based on household

331 genealogies. To resolve inconsistencies, we took either the genealogy from the most

332 knowledgeable individual (i.e. mother over aunt) or the genealogy that reduced other

333 inconsistencies (i.e. discarding six-month interbirth intervals). Genealogies contained 2953 living

334 and dead Agta. We used the *R* packages *pedigree*, *kinship2*, and *igraph* to measure

335 consanguineous relatedness (r)[34,38]. For comparative purposes, we obtained 14 saliva samples

336 from neighbouring Palanan farmers, making sure individuals were unrelated by directly asking.

337 *Bayaka demography*. Ethnographic data collection took place over two seasons in April-June

338 2013 and February-October 2014. We collected saliva samples from 21 individuals for

339 microbiome analyses.

340 *Ethics.* This study was approved by UCL Ethics Committee (UCL Ethics code 3086/003) and

341 carried out with permission from local government and community members. Informed consent

342 was obtained from all participants, after group and individual explanation of research objectives

343 in the indigenous language. A small compensation (usually a thermal bottle or cooking utensils)

344 was given to each participant. The National Commission for Indigenous Peoples (NCIP), advised

345 us that the process of Free Prior Informed Consent with the tribal leaders, youth and elders

346 would be necessary to validate our data collection under their supervision. This was done in2017

17

347   with the presence of all tribal leathers, elders and youth representatives at the NCIP regional

348   office, with the mediation of the regional officer and the NCIP Attorney. The validation process

349   was approved unanimously by the tribal leaders, and the NCIP, and validated the full 5 years of

350   data collection.

351

352   **Oral microbiome analysis**

353   *Microbial DNA extraction and 16S rRNA gene sequencing*. A total of 190 saliva samples were

354   selected from Agta hunter-gatherers (n = 155) and Palanan farmers in the Philippines (n = 14),

355   and Bayaka hunter-gatherers from the Congo (n = 21). Microbial DNA was extracted following

356   the protocol for manual purification of DNA for Oragene·DNA/saliva samples. The 16S rRNA

357   gene V3-V4 region was amplified by PCR with primers containing Illumina adapter overhang

358   nucleotide sequences. All PCR products were validated through an agarose gel and purified with

359   magnetic beads. Index PCR was then performed to create the final library also validated through

360   an agarose gel. All samples were pooled together at equimolar proportions and the final pool was

361   qPCR-quantified before MiSeq loading. Raw Illumina pair-end sequence data were

362   demultiplexed and quality-filtered with *QIIME* 2 2019.1[39] and *DADA2*[40], which generates single

363   nucleotide exact amplicon sequence variants (ASV or ESV). ASVs are biologically meaningful

364   as they identify a specific sequence and allow for higher resolution than operational taxonomic

365   units (OTUs)[41] or clusters of sequences above a similarity threshold, and thus an ASV is

366   equivalent to a 100% similar OTU. Taxonomic information was assigned to ASVs using a naïve

367   Bayes taxonomy classifier against the *SILVA* database release 132 with a 99% identity

368   sequence[42].

369   Reads outside the kingdom Bacteria or assigned to mitochondria or chloroplasts were

370   removed. Phylogenetic analyses aligned sequences with *MAFFT*[43] and generated a rooted

371   phylogenetic tree with *FastTree2*[44] using default settings via *QIIME* 2. We generated an Alpha

372   rarefaction curve with R package *vegan* to confirm that sample richness had been fully observed

373   (Supplementary Figure 3).

374   Samples with extremely low number of reads (8000) were removed. This resulted in

375   6409 ASVs (later reduced to 1980 ASVs present in at least two individuals and abundance of at

376   least 10 counts per individual) and 173 individuals: 138 Agta, 21 Bayaka, and 14 Palanan

377   farmers.

378   *Identification of the Agta social microbiome*. In our Agta sample, we first selected a set of strong

379   social links (top 25% of the weight distribution from each camp and multi-camp). For each ASV,

380   we calculated the proportion of strong links (f_A^s) where a given ASV A was present. Next, we

381   calculated the same proportion in the complementary set of 75% weak social links (f_A^w). We then

382   computed for each ASV A the score $s_A^s = \frac{f_A^s - f_A^w}{f_A^s + f_A^w}$, or normalised difference between the two

383   proportions. This score can be paired with the z-score $\frac{f_A^s - f_A^w}{\sqrt{f_A(1-f_A)(\frac{1}{n^s}+\frac{1}{n^w})}}$ which quantifies the

384   deviation from the null hypothesis that the two proportions are equal, with n^s and n^w as

385   respectively the numbers of strong and weak links and f_A as the proportion of total links that

386   share ASV A. We then selected as affected by social interaction 137 ASVs with s >0.5 and P<

387   0.05. P-values were adjusted by False Discovery Rate (FDR).

388   We performed a sensitivity analysis to investigate the consequences of varying the

389   threshold defining strong versus weak social links in the Agta social network. Instead of 137

390   ASVs resulting from selecting dyads at top 25% of the weight distribution, we obtained:

391   Top 45%: 166 ASVs

392   Top 35%: 156 ASVs

393   Top 25%: 137 ASVs

394   Top 15%: 138 ASVs

395   Top 5%: 99 ASVs

396   The list shows that whether the strong set consists of a very reduced number dyads with

397   very strong weights (top 5%), or instead includes nearly half the dyads (top 45%), the number of

398   ASVs significantly responsive to social contact varies from 99 (5%) to 166 (8%), representing a

399   small fraction of the total of 1980 ASVs found in the Agta. Therefore, setting the threshold at the

400   top 25% did not affect our results and conclusions.

401

402   **Agta social network data, construction and analysis**

403   *Mote devices.* Motes are wireless sensing devices storing all between-device communications

404   within a specified distance and have been described in detail elsewhere[6,7,18,45]. We used the

405   UCMote Mini (with a TinyOS operating system) sealed into wristbands or belts, labelled with a

406   unique number, and identified with coloured string to avoid accidental swaps. Motes require no

407   grounded infrastructure and collect interactions even when individuals are away from camps.

408   Individuals arriving at a camp after the start of data collection were given a mote and entry time

409   was recorded, while those leaving a camp before the end of data collection had their exit time

410   recorded. To prevent swaps individuals were checked twice daily, and mote numbers were

411   checked upon return. Any swaps were later corrected by reassigning data to the correct

412   individuals.

413   Data were later downloaded via a PC side application in *Java*. Data were limited to 5am-

414   8pm. We ran raw data through a stringent data-processing system in *Python* to prevent data

415   corruption. Data were matched to ID numbers and start-stop times of each mote. The result was a

416   matrix with the number of recorded beacons for all possible dyads and their weights.

417   For the camp-level experiment, all individuals from four camps wore motes from five to

418   seven days. Each device sent a message every two minutes that contained its unique ID, a time

419   stamp and the signal strength. Messages are stored by any other mote within a three-meter

420   radius, a frequently used threshold[46,47]. For the multi-camp experiment, adult individuals from

421   two areas (consisting of seven and three camps respectively) wore motes for one month.

422   *Effect of dyad category on bacterial sharing.* The bacterial sharing network was constructed by

423   defining link weights as the number of social bacteria shared by two individuals. Dyads in the

424   network were classified into: i) levels of kinship (mother-offspring, father-offspring, siblings:

425   r=0.5, other kin: r=0.25 or r=0.125, non-kin: r=0.0625 or lower, spouses, friends: defined as

426   non-kin at the top 25% distribution of social dyadic weights, and other non-kin) and ii) residence

427   (same or different camp, same or different household). Mean weights were calculated for each

428   dyadic category (Supplementary Table 2). Then, we produced 1000 network randomisations

429   based on a single-step ID swap between nodes. For example, if dyad 1 consisted of two spouses

430   in the real network, randomisation preserved dyad 1 and its weight, but randomly replaced the

431   two nodes (potentially changing the dyadic classification to siblings, friends, etc.). We calculated

432   the mean weights for each dyadic category in the 1000 randomised networks, and then calculated

433   one-sample t-tests with the mean weight in the real network as the test value. We repeated the

434   analysis for non-social bacteria (Supplementary Table 3).

435   *Reinforcement analysis.* In a multilayer network, reinforcement analysis measures the overlap in

436   links between different layers to quantify the probability of finding a link on a layer conditioned

437   on the weight of the same link on another[48]. Reinforcement between two network layers α and

438   α', or $P(\alpha'|\alpha)$, is defined as

439    $P(\alpha'|\alpha) = \dfrac{\sum_{ij} a_{ij}^{[\alpha']} w_{ij}^{[\alpha]}}{\sum_{ij} w_{ij}^{[\alpha]}}$ (1)

440   where $a_{ij}^{[\alpha']}$ is the adjacency matrix of conditioned layer α', and $w_{ij}^{[\alpha]}$ is the adjacency matrix of

441   the conditioning layer α. We split the weight of social links into three tertiles, and computed

442   equation (1) for each. We obtained increasing values of reinforcement from the lower to the

443   higher tertile, providing evidence of an effect of social contact on bacterial sharing.

444

445   **ASV classification**

446   *ASV diversity metrics*. To distinguish between the effects of lifestyle and shared ecology on the

447   microbiome, we compared the diversity of the oral microbiome of Agta hunter-gatherers with

448   neighbouring Palanan farmers in the Philippines and with Bayaka hunter-gatherers in Congo.

449   Using the 6409 ASVs dataset we calculated the number of observed ASVs in each population

450   with *R* package *Phyloseq* (version 1.30.0) and the Faith's Phylogenetic Diversity index was

451   calculated with *R* package *picante* using the generated rooted phylogenetic tree. To estimate the

452   number of shared ASVs in the Agta, Bayaka and Palanan farmers, we sampled a random subset

453   of 10 samples for each population without replacement, calculated the shared ASVs between the

454   populations and repeated this procedure 100 times. Global differences between groups and

455   pairwise comparisons were assessed by Kruskal-Wallis and Wilcoxon Rank Sum tests

456   respectively and plotted by the *R* package *ggpubr*. Pairwise p-values were adjusted by False

457   Discovery Rate (FDR).

458   *Classification of oral bacteria as pathogens.* ASVs were classified as oral pathogens if they have

459   been reported as etiological agents of periodontitis[49,50] or dental caries[51,52]. For gum disease

460   pathogens, these included the classical "red" and "orange" complex of periodontal pathogens and

461   the recent update by Pérez-Chaparro and cols[50] based on systematic review and metaanalysis.

462   For caries pathogens, the list of active microorganisms detected through metatranscriptomics of

463   cavities was used. Common oral commensals potentially causing endocarditis or systemic

464   infections in immunocompromised patients only were not considered pathogens. Bacteria

465   reported as etiological agents of lower respiratory infections (e.g. pneumonia, whooping cough,

466   bronchitis, or sinusitis) and biofilm-mediated infections (e.g. lactational mastitis, medical

467   implant biofilm infections, chronic lung infections, osteomyelitis or chronic wounds) were also

468   considered pathogens, including organisms present in healthy carriers[53,54]. Bacteria causing

469   urinary tract infections or sexually transmitted diseases transiently found in the oral cavity were

470   also considered pathogens[55]. Bacteria were classified as "oral" if detected in more than 10% of

471   the population in oral samples according to the Human Oral Microbiome database. If a bacterial

472   species or genus had been isolated from the oral cavity of an animal, it was also classified as

473   oral.

474   For assignment of bacteria to pathogenic or non-pathogenic, we used species-level ASVs,

475   given that there are multiple cases where different species from the same genus had a different

476   assignment. If taxonomic classification of the ASV was only possible at the genus level, it was

477   considered a pathogen if: i) >90% of named species within the genus were pathogenic, or ii) the

478   genus included a major pathogenic species but the remaining species within the genus were not

479   classified as oral by the Human Oral Microbiome Database[56]. ASV with a top hit to a sequence

480   classified as "Oral taxa" in databases but without a species assignment were not considered

481   named species and were discarded from the analysis. Cases where taxonomic classification of the

482   ASV was only possible at the family level or higher were also discarded.

23

483

484

Acknowledgements

Data availability

16S amplicon data (EGAS00001005317) are deposited at the European Genome-phenome Archive (EGA), which is hosted at the EBI and the CRG. Data at the individual level on age, kinship relationships, household composition, camp assignation and social contacts that support the findings of this study are available on request from the corresponding author (A.B.M.). The individual data are not publicly available due to information that could compromise research participant privacy.

Code availability

Source code and data for visualization are available at https://doi.org/10.5281/zenodo.6338840

References

1. Migliano, A. B., Vinicius, L. & Lahr, M. M. Life history trade-offs explain the evolution of human pygmies. *Proc. Natl. Acad. Sci. U. S. A.* **104**, (2007).

2. Hill, K. R. *et al.* Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science (80-.).* **331**, 1286–1289 (2011).

3. Apicella, C. L., Marlowe, F. W., Fowler, J. H. & Christakis, N. A. Social networks and cooperation in hunter-gatherers. *Nature* **481**, 497–501 (2012).

4. Smith, K. M., Larroucau, T., Mabulla, I. A. & Apicella, C. L. Hunter-gatherers maintain assortativity in cooperation despite high levels of residential change and mixing. *Curr. Biol.* **28**, 3152–3157 (2018).

5. Derex, M. & Boyd, R. Partial connectivity increases cultural accumulation within groups. *Proc. Natl. Acad. Sci.* **113**, 2982–2987 (2016).

6. Migliano, A. B. *et al.* Hunter-gatherer multilevel sociality accelerates cumulative cultural evolution. *Sci. Adv.* **6**, (2020).

7. Page, A. E. *et al.* Hunter-Gatherer Social Networks and Reproductive Success. *Sci. Rep.* **7**, (2017).

8. Nasidze, I. *et al.* High Diversity of the Saliva Microbiome in Batwa Pygmies. *PLoS One* **6**, e23352 (2011).

9. Schnorr, S. L. *et al.* Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* **5**, 3654 (2014).

10. Rampelli, S. *et al.* Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Curr. Biol.* **25**, 1682–1693 (2015).

11. Smits, S. A. *et al.* Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* **357**, 802–806 (2017).

12. Fragiadakis, G. K. *et al.* Links between environment, diet, and the hunter-gatherer microbiome. *Gut Microbes* **10**, 216–227 (2019).

13. Schnorr, S. L. The Diverse Microbiome of the Hunter-Gatherer. *Nature* **518**, S14–S15 (2015).

14. Gomez, A. *et al.* Gut Microbiome of Coexisting BaAka Pygmies and Bantu Reflects Gradients of Traditional Subsistence Patterns. *Cell Rep.* **14**, 2142–2153 (2016).

15. Moeller, A. H. The shrinking human gut microbiome. *Curr. Opin. Microbiol.* **381. A. H.,** 30–35 (2017).

16. Meehan, C. L. *et al.* Social networks, cooperative breeding, and the human milk microbiome. *Am. J. Hum. Biol.* **30**, e23131 (2018).

17. Lassalle, F. *et al.* Oral microbiomes from hunter☐gatherers and traditional farmers reveal shifts in commensal balance and pathogen load linked to diet. *Mol. Ecol.* **27**, 182–195 (2018).

18. Migliano, A. B. *et al.* Characterization of hunter-gatherer networks and implications for cumulative culture. *Nat. Hum. Behav.* **1**, 1–6 (2017).

19. Dobon, B. *et al.* The making of the oral microbiome in Agta hunter-gatherers. (2021).

20. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).

21. Schnorr, S. L. *et al.* Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* **5**, 3654 (2014).

22. Sarkar, A. *et al.* Microbial transmission in animal social networks and the social microbiome. *Nat. Ecol. Evol.* **4**, 1020–1035 (2020).

23. Chaudhary, N. *et al.* Polygyny without wealth: popularity in gift games predicts polygyny in Bayaka Pygmies. *R. Soc. open Sci.* **2**, 150054 (2015).

24. Dyble, M. *et al.* Networks of Food Sharing Reveal the Functional Significance of Multilevel Sociality in Two Hunter-Gatherer Groups. *Curr. Biol.* **26**, (2016).

25. Smith, D. *et al.* Camp stability predicts patterns of hunter-gatherer cooperation. *R. Soc. Open Sci.* **3**, (2016).

26. Salali, G. D. *et al.* Knowledge-sharing networks in hunter-gatherers and the evolution of cumulative culture. *Curr. Biol.* **26**, 2516–2521 (2016).

27. Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C. & Kent Jr., R. L. Microbial complexes in subgingival plaque. *J. Clin. Periodontol.* **25**, 134–144 (1998).

28. Pérez-Chaparro, P. J. *et al.* Newly Identified Pathogens Associated with Periodontitis: A Systematic Review. *J. Dent. Res.* **93**, 846–858 (2014).

29. Simón-Soro, A., Guillen-Navarro, M. & Mira, A. Metatranscriptomics reveals overall active bacterial composition in caries lesions. *J. Oral Microbiol.* **6**, 25443 (2014).

30. Simón-Soro, A. & Mira, A. Solving the etiology of dental caries. *Trends Microbiol.* **23**, 76–82 (2015).

31. Jung, H.-S., Ehlers, M. M., Lombaard, H., Redelinghuys, M. J. & Kock, M. M. Etiology of bacterial vaginosis and polymicrobial biofilm formation. *Crit. Rev. Microbiol.* **43**, 651–667 (2017).

32. Natsis, N. E. & Cohen, P. R. Coagulase-Negative Staphylococcus Skin and Soft Tissue Infections. *Am. J. Clin. Dermatol.* **19**, 671–677 (2018).

33. Haq, I. J., Battersby, A. C., Eastham, K. & McKean, M. Community acquired pneumonia in children. *BMJ* **356**, j686 (2017).

34. Dyble, M. *et al.* Sex equality can explain the unique social structure of hunter-gatherer bands. *Science (80-.).* **348**, 796–798 (2015).

35. Brito, I. L. *et al.* Transmission of human-associated microbiota along family and social networks. *Nat. Microbiol.* **4**, 964–971 (2019).

36. Battiston, F., Nicosia, V. & Latora, V. Structural measures for multiplex networks. *Phys. Rev. E* **89**, 32804 (2014).

37. Diekmann, Y. *et al.* Accurate age estimation in small-scale societies. *Proc. Natl. Acad. Sci. U. S. A.* **114**, (2017).

38. Page, A. E., Myers, S., Dyble, M. & Migliano, A. B. Why so many Agta boys? Explaining 'extreme' sex ratios in Philippine foragers. *Evol. Hum. Sci.* **1**, (2019).

39. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

40. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

41. Callahan, B. J., McMurdie, P. J. & Holmes, S. P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **11**, 2639–2643 (2017).

42. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).

43. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

44. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490–e9490 (2010).

597   45.   Page, A. E. *et al.* Testing adaptive hypotheses of alloparenting in Agta foragers. *Nat.*
598   *Hum. Behav.* **3**, 1154–1163 (2019).
599   46.   Brent, L. J. N., Lehmann, J. & Ramos-Fernández, G. Social network analysis in the study
600   of nonhuman primates: A historical perspective. *Am. J. Primatol.* **73**, 720–730 (2011).
601   47.   Isella, L. *et al.* What's in a crowd? Analysis of face-to-face behavioral networks. *J. Theor.*
602   *Biol.* **271**, 166–180 (2011).
603   48.   Battiston, F., Nicosia, V. & Latora, V. Structural measures for multiplex networks. *Phys.*
604   *Rev. E* **89**, 32804 (2014).
605   49.   Socransky, S. S., Haffajee, A. D., Cugini, M. A., Smith, C. & Kent Jr., R. L. Microbial
606   complexes in subgingival plaque. *J. Clin. Periodontol.* **25**, 134–144 (1998).
607   50.   Pérez-Chaparro, P. J. *et al.* Newly Identified Pathogens Associated with Periodontitis: A
608   Systematic Review. *J. Dent. Res.* **93**, 846–858 (2014).
609   51.   Simón-Soro, A., Guillen-Navarro, M. & Mira, A. Metatranscriptomics reveals overall
610   active bacterial composition in caries lesions. *J. Oral Microbiol.* **6**, 25443 (2014).
611   52.   Simón-Soro, A. & Mira, A. Solving the etiology of dental caries. *Trends Microbiol.* **23**,
612   76–82 (2015).
613   53.   Leung*, A. K. C. & Hon, A. H. C. W. and K. L. Community-Acquired Pneumonia in
614   Children. *Recent Patents on Inflammation & Allergy Drug Discovery* vol. 12 136–144
615   (2018).
616   54.   Natsis, N. E. & Cohen, P. R. Coagulase-Negative Staphylococcus Skin and Soft Tissue
617   Infections. *Am. J. Clin. Dermatol.* **19**, 671–677 (2018).
618   55.   Jung, H.-S., Ehlers, M. M., Lombaard, H., Redelinghuys, M. J. & Kock, M. M. Etiology
619   of bacterial vaginosis and polymicrobial biofilm formation. *Crit. Rev. Microbiol.* **43**, 651–
620   667 (2017).
621   56.   Chen, T. *et al.* The Human Oral Microbiome Database: a web accessible resource for
622   investigating oral microbe taxonomic and genomic information. *Database (Oxford).* **2010**,
623   baq013–baq013 (2010).
624
625
626