PNAS
www.pnas.org

Supplementary Information for

The evolution and changing ecology of the African hominid oral microbiome

James A. Fellows Yates, Irina M. Velsko, Franziska Aron, Cosimo Posth, Courtney A. Hofman, Rita M. Austin, Cody E. Parker, Allison E. Mann, Kathrin Nägele, Kathryn Weedman Arthur, John W. Arthur, Catherina C. Bauer, Isabelle Crevecoeur, Christophe Cupillard, Matthew C. Curtis, Love Dalén, Marta Díaz-Zorita Bonilla, J. Carlos Díez Fernández-Lomana, Dorothée G. Drucker, Elena Escribano Escrivá, Michael Francken, Victoria E. Gibbon, Manuel Gonzalez Morales, Ana Grande Mateu, Katerina Harvati, Amanda G. Henry, Louise Humphrey, Mario Menéndez, Dušan Mihailović, Marco Peresani, Sofía Rodríguez Moroder, Mirjana Roksandic, Hélène Rougier, Sandra Sázelová, Jay T. Stock, Lawrence Guy Straus, Jiří Svoboda, Barbara Teßmann, Michael J. Walker, Robert C. Power, Cecil M. Lewis, Krithivasan Sankaranarayanan, Katerina Guschanski, Richard Wrangham, Floyd E. Dewhirst, Domingo C. Salazar-Garcia, Johannes Krause, Alexander Herbig, Christina Warinner.

James A. Fellows Yates, Christina Warinner
Email: fellows@shh.mpg.de, warinner@shh.mpg.de

This PDF file includes:

Supplementary text
Figures S1 to S14
Table S1
Legends for Datasets S1 to S3
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S3

Table of Contents

Supplementary Information Text
S1. Primate host species and subspecies included in this study 4
S1.1 Alouatta (Outgroup) 4
S1.1.1 *Alouatta palliata* (Mantled howler monkey) 4
S1.2 Gorilla 5
S1.2.1 *Gorilla beringei* (Eastern gorilla) 5
S1.2.2 *Gorilla gorilla* (Western lowland gorilla) 6
S1.3 Pan 6
S1.3.1 *Pan troglodytes* (Common chimpanzee) 7
S1.4 Homo 8
S1.4.1 *Homo neanderthalensis* (Neanderthal) 8
S1.4.2. *Homo sapiens* (Modern Human) 9
S2. Laboratory procedure 11
S2.1 Dental calculus collection 12
S2.2 DNA extraction 12
S2.3 Library preparation and sequencing 13
S2.3.1 Shallow Sequenced Dataset 13
S2.3.2 Deep Sequenced Dataset 14
S2.4 Sequencing 15
S2.4.1 Dental calculus 15
S2.4.2 Controls 15
S2.4.3 Sequence data 15
S3. Data processing and quality filtering 15
S3.1 Publicly available data 15
S3.1.1 Dental calculus 15
S3.1.2 Modern human microbiome and environment 16
S3.2 Sequencing Quality Control and Human DNA Removal 16
S3.2.1 Shallow sequenced dataset 17
S3.2.2 Deep Sequenced dataset 17
S3.3 Taxonomic binning and classification 17
S3.4 Preservation assessment and removal of low quality samples 19
S3.4.1 Cumulative percent decay 20
S3.4.2 SourceTracker 22
S3.4.3 Comparison of methods 24
S3.4.4 Abundance of eukaryotic content as a preservation indicator? 24
S3.5 Ancient DNA authentication 24
S3.6 Contamination assessment and removal 25
S4. Microbial compositional analysis 26
S4.1 Principal coordinates analysis (PCoA) 26
S4.2 PERMANOVA 27
S4.3 Hierarchical clustering 28
S4.4 Indicator analysis 30
S4.5 Are dietary practices associated with oral microbiome composition? 31
S5. Core microbiome 32
S5.1 Biofilm spatial organization and development: current state of knowledge 32
S5.1.1 Periodontal disease and associated microbial taxa 34
S5.2 Parameter selection of core microbiome analysis 36
S5.3 Comparison among African hominid core microbiomes 39
S5.4 Evolutionary and ecological relationships among core oral microbes of African hominids 40
S5.5 Evolution of oral disease virulence factors in late-colonisers 42
S5.6 Host-associated differences in early-coloniser Streptococcus species 43
S5.6.1 Streptococcus α-amylase-binding proteins as an indicator of host-microbe coevolution 44
S5.7 Detection of amylase-binding protein genes abpA and abpB 46
S5.8 Salivary amylase, cooked starches, and the Expensive-Tissue Hypothesis 47

S5.9 Dating expansion of starch-associated streptococci using amylase binding proteins 51
S5.9.1 *abpA* 51
S5.9.2 *abpB* 52
S6. Microbial phylogenetics 52
S6.1 Super-reference construction 53
S6.2 Alignment and species selection 54
S6.3 Performance of super-reference vs. single reference mapping 55
S6.4 Single reference genome mapping statistics 56
S6.5 Variant calling and single-allelic position assessment 56
S6.6 Phylogenetic trees 57
S7. Functional and metabolic pathway analysis 60
S7.1 HUMAnN2 60
S7.1.1 Pathway abundance 61
S7.1.2 KEGG ortholog distribution 61
S7.1.3 Species contributions to KEGG orthologs 64
S7.1.4 Metabolic category PCAs 65
S7.2 AADDER 66
S7.2.1 SEED profile 66
S7.2.2 Species contributions to SEED proteins 69
S7.2.3 Metabolic category PCAs 70

Supplementary Figures

Fig. S1. Primate species analysed in this study and their dental calculus 72
Fig. S2. Flowchart of study processing, authentication and analysis procedures 73
Fig. S3. Data quality and authentication filters applied to dental calculus samples in this study 74
Fig. S4. Representative ancient authenticity metrics for Neanderthals from Pesturina (PES001) and de Nadale (GDN001) caves 75
Fig. S5. Principal Coordinates Analysis (PCoA) of dental calculus and sources 76
Fig. S6. Hierarchical clustering of different Homo calculus microbiomes, comparing different lifestyles and regions 77
Fig. S7. Alluvial diagram showing effects of increasing the minimum abundance threshold to the MALT OTU table-based core microbiome calculations 78
Fig. S8. Alluvial diagram showing effects of dropping and retaining a single-individual Gorilla population in core microbiome calculations at genus and species taxonomic levels 79
Fig. S9. Core microbiome strategy and virulence genes 80
Fig. S10. Comparison of the number of multi-allelic single nucleotide polymorphisms (SNPs) identified when using a single representative genome mapping strategy versus a multi-reference genome (super-reference) mapping approach 81
Fig. S11. Neighbour joining trees of eight well-supported core calculus taxa from single representative mappings of the deep sequenced dataset 82
Fig. S12. Replication of production dataset phylogenies with low-coverage and damage-containing screening dataset with additional European individuals 83
Fig. S13. Principal Components Analysis (PCA) biplots of functional annotations with the top 10 positive and negative loadings separating Homo from the other hosts 84
Fig. S14. Principal Components Analysis (PCA) using KEGG orthologs or SEED-classified proteins belonging to specific metabolic pathway categories 85

Supplementary Tables

Table S1. Summary of host species and dental calculus samples included in study 86

Legends for Datasets S1 to S3

Data S1. Laboratory metadata: detailed sample, laboratory and analytical metadata (.xlsx)
Data S2. Metagenomic taxa tables: OTU tables from MALT taxonomic assignment (.xlsx)
Data S3. Core microbiome taxa (.xlsx)

References 88

S1. Primate host species and subspecies included in this study

To better understand oral microbiome evolution and ecology in primates, dental calculus from six host species spanning three hominid genera and one New World howler monkey outgroup were collected and analysed in this study (Table S1). These taxa were selected to focus on African members of the family Hominidae (great apes including 'archaic' and modern humans) and using the New World howler monkey species *Alouatta palliata* as an outgroup. A cladogram illustrating the relationships between these host species is provided in Fig. S1A.

Dental calculus samples in this study were obtained from primate research stations, museums, dental clinics, and archaeological collections originating from 19 countries in Africa, Europe, and Central America (Fig. 1a; Data S1). For present-day modern humans, dental calculus was obtained under informed consent during routine dental cleanings, and research protocols for analyzing de-identified samples were approved by the Institutional Review Board for Human Research Participant Protection at the University of Oklahoma (IRB#4543). Archaeological dental calculus within the genus *Homo* was obtained from three groups chosen to represent distinctive phases in human history: 1) pre-antibiotic era modern humans, 2) pre-agricultural modern humans, and 3) extinct *Homo*. All non-human primate dental calculus was obtained postmortem from wild populations. To avoid potential group- or collection-specific batch effects, dental calculus samples for each host species were drawn from at least two different populations, with the exception of the howler monkey outgroup, which were all obtained from a single island in Nicaragua.

A total of 115 new dental calculus samples from 109 individuals were initially screened for this study, of which 75 individuals passed genetic quality filters. In addition, we also analysed previously published calculus data from four Neanderthals, one chimpanzee (1), and ten present-day modern humans (2), of which all but one Neanderthal sample passed our genetic quality filters. Our combined total sample size was thus 124 individuals, of which 89 passed genetic quality filters. An overview of each host species, contributing institution, and collection details are provided below and are summarised in Data S1. Below we describe each species, their dietary habits, and known factors affecting their oral environment (e.g., saliva production, salivary pH, salivary protein composition, and antifeedant intake), which may influence their oral microbiome. Please note that in the following sections we use the following temporal terminology: 'Ma' for million years ago; 'ka' for thousand years ago.

S1.1 Alouatta (Outgroup)

Alouatta is a genus of New World monkeys in the family Atelidae native to tropical forests in Central and South America. New World monkeys (Platyrrhini) are estimated to have diverged from Old World monkeys and apes (Catarrhini) approximately 40 Ma (3, 4).

S1.1.1 Alouatta palliata (Mantled howler monkey)

Mantled howler monkeys are among the largest Central American monkeys. They consume a seasonal diet mostly of young leaves and ripe fruits, focusing primarily on trees in the family Moraceae, especially *Ficus* spp., *Brosimum alicastrum*, and *Poulsenia armata*, but also the trees *Cecropia obtusifolia*, *Spondias radlkoferi*, and *Dipholis minutiflora (5)*. Microwear analysis of *A. palliata* dentition shows a high degree of anisotropy (6) reflecting the relative toughness of their diet and repetitive masticatory abrasion by plant phytoliths. Leaves in howler monkey diets are rich in plant defensive compounds, especially tannins (7, 8), which are antifeedants produced by plants that bind dietary proteins and interfere with nutrient digestion (9). In response, *A. palliata* has large salivary glands to aid in food lubrication and the neutralization of dietary tannins (9). *A. palliata* saliva is slightly alkaline, which reduces complexing of tannins with dietary proteins, and it contains high amounts of proline-rich proteins (PRPs), which have a high affinity for binding and neutralizing tannins and other polyphenolic compounds (7). Consequently, salivary expression of tannin-binding salivary proteins, such as PRPs, is thought to be a physiological adaptation to a tannin-rich diet (7). Although PRP expression can be dietarily induced in other taxa, such

as rodents (9), PRP expression in howler monkeys is thought to be continuous (8). Like all New World monkeys, *A. palliata* do not express α-amylase in saliva. In primates, α-amylase expression in saliva emerged in Catarrhini (Old World monkeys, apes, and humans) through the insertion of a γ-actin pseudogene after the Platyrrhini-Catarrhini lineages split approximately 40 Ma, followed by a retroviral insertion in the hominoid lineage (superfamily Hominoidea) after the split with Old World monkeys approximately 20 Ma (10). Heavy dental calculus deposits are regularly observed in *A. palliata* skeletal remains (11).

The historic *A. palliata* dental calculus included in this study (n=5) was obtained from the Ometepe Biological Field Station operated by the Maderas Rainforest Conservancy (MRC) on Ometepe Island, Nicaragua. Dental calculus (Fig. S1B) was collected postmortem from a skeletal anatomy collection prepared from animals that had died of natural causes between 2000 and 2013. The skeletons were collected from the forest floor, and, if not already decomposed, bodies underwent a brief period of burial prior to integration into the skeletal anatomy collection. Permissions for dental calculus sampling and destructive analysis were provided by the director of the Maderas Rainforest Conservancy at the Ometepe Biological Field Station, Nicaragua. Further details about the individuals sampled are included in Data S1.

S1.2 Gorilla

Gorilla is an Old World hominid genus in the family Hominidae (great apes) and subfamily Gorillinae that includes two extant species: *G. beringei* (eastern gorillas) and *G. gorilla* (western gorillas). Gorillas are estimated to have diverged from other African great apes 5.6-11.2 Ma, and eastern and western lineages are estimated to have split approximately 150-300 ka (12).

S1.2.1 Gorilla beringei (Eastern gorilla)

Eastern gorilla is a critically endangered species of gorilla native to parts of central and eastern Africa (13). Individuals belonging to its two subspecies, mountain gorilla (*Gorilla beringei beringei*) and Grauer's gorilla (*Gorilla beringei graueri*, formerly known as eastern lowland gorilla), were sampled in this study; these subspecies are estimated to have diverged during or just before the Younger Dryas approximately 10 ka (14). Mountain gorillas primarily inhabit medium and high altitude Afromontane forests located along the volcanic slopes of the Virunga mountains and within the Bwindi Impenetrable National Park in East Africa (15). The population size of the Virunga mountain gorillas dropped to below 300 individuals in the 1980s, but subsequent conservation efforts have resulted in population increases to over 600 individuals in the wild today (16–18). The current estimate of total mountain gorillas today (Virunga and Bwindi populations) is approximately 1,000 (19). Mountain gorillas subsist almost entirely on an herbivorous diet focused on hard and tough plant structural material, such as leaves, stems, roots, wood, bark, and pith, as well as flowers and fruit (20). Mountain gorilla diet shows relatively little seasonal variation, with the exception of the seasonal availability of bamboo shoots (21). As with *A. palliata*, the dentition of the mountain gorilla exhibits anisotropic microwear consistent with repetitive masticatory abrasion by plant phytoliths. Mountain gorillas have been previously noted to consistently harbour 'black' calculus deposits (22, 23), which were also observed across all gorillas sampled in this study (Fig. S1c,d), and to have high levels of dental disease in general (24). Grauer's gorillas are the largest gorilla subspecies, and they inhabit tropical lowland forests in eastern central Africa. Their current population is estimated at under 4,000 individuals (18). They consume a more seasonal diet than mountain gorillas, and their diets are characterised by a greater proportion of fruit (21). The salivary composition of eastern gorillas is not well-studied. However, like all great apes, eastern gorillas express salivary amylase (25).

The historic eastern gorilla dental calculus samples in this study (n=21) were obtained from the Natural History Museum of Stockholm (NRM; mountain gorilla, n=15; Grauer's gorilla, n=1), the Royal Belgian Institute of Natural Sciences (Grauer's gorilla, n=1), and the Royal Museum of Central Africa (Grauer's gorilla, n=4) (18, 26). All eastern gorillas in this study were wild individuals from the early 20th century. Permissions for dental calculus sampling and destructive analysis were provided by the curators of the Vertebrate (Royal Museum of Central Africa) and Mammalogy (Royal Belgian Institute of Natural

Sciences; Naturhistoriska Riksmuseet, Stockholm) collections, respectively. Further details about the individuals sampled are included in Data S1.

S1.2.2 Gorilla gorilla (Western lowland gorilla)

Western lowland gorillas (*Gorilla gorilla gorilla*) are a critically endangered species of gorilla native to parts of central and western Africa, where they inhabit lowland rainforests (27). Western gorillas consume a seasonally varied diet consisting of fruits, leaves, herbs, roots, bark, pith, and termites (28). During the fruiting season, a majority of feeding time is spent consuming ripe tree fruits, while leaves make up the bulk of the diet during the rest of the year. Overall, western gorillas are more frugivorous than their eastern counterparts, and they selectively consume food parts or plants with higher nutrient (sugar, protein) and lower antifeedant (tannin) content than Old World monkeys (29). Western gorillas exhibit a high amount of dietary overlap with sympatric chimpanzees, differing only during periods of fruit scarcity; during these periods, western gorillas are more likely to consume fibrous foods containing tannins and phenols than chimpanzees (30). Compared with chimpanzees, gorillas have a larger body size, larger colons, and slower digesta passage rates, which enables digestion of more fibrous foods. With respect to their dentition, western gorilla molars are less highly crested than their eastern counterparts, which may be related to the higher amount of soft ripe fruit in their diet (30). However, despite differences in their diets, western and eastern gorillas do not exhibit significant differences in dental microwear (6). The salivary composition of western gorillas was recently studied and found to have an overall high protein concentration, and major salivary proteins were shared across gorillas, chimpanzees, and modern humans, although at different abundances and with different glycosylation patterns (31). Gorillas also express latherin (LATH/BASE), a surfactant and wetting agent found in the saliva and sweat of a range of mammals (32), but not in modern humans due to pseudogenization (31, 32). Copy number variation of the *AMY1* gene in gorillas is uncertain, but previously published quantitative PCR data have confirmed that modern humans have a higher *AMY1* copy number than either chimpanzees or gorillas (33). Basal salivary α-amylase activity in western gorillas is lower than that of modern humans but significantly higher than that of chimpanzees (31, 34), and it increases with age (35). The basis of this elevated activity is not known. Stress is known to stimulate salivary α-amylase expression; however, gorillas have relatively low cortisol levels compared to chimpanzees (35). Tannin-rich diets have also been shown to induce salivary α-amylase expression in mice (36), and gorillas consume more tannin-rich diets and starch (in the form of roots) than chimpanzees (35); however, it is not known if dietary tannins influence salivary α-amylase expression in apes. Tannins inhibit salivary α-amylase activity and bind salivary α-amylase with only slightly lower affinity than PRPs, and it has been proposed that tannin neutralization may be a secondary function of salivary α-amylase, after its primary function in starch digestion (36).

The historic western gorilla dental calculus samples in this study (n=8) were obtained from the Hamann-Todd Collection at the Cleveland Museum of Natural History (CMNH; Cleveland, USA; n=7) and the Natural History Museum of Stockholm (NRM, n=1). All CMNH individuals were a part of the collection of C. Hamann and T. Todd who obtained wild-shot primates from Africa in the early 20th century and subsequently donated them to the CMNH. Permissions for dental calculus sampling and destructive analysis were provided by the collections manager of the CMNH Physical Anthropology department. The NRM individual was originally classified as *Gorilla beringei beringei,* but was later confirmed using genetic testing by T. van der Valk and K. Guschanski to be *Gorilla gorilla gorilla*. Permission for destructive sampling of dental calculus was provided by the NRM mammal curator. Further details about the individuals sampled are included in Data S1.

S1.3 Pan

Pan is a genus in the subfamily Homininae and tribe Panini, and it includes two extant species: *P. troglodytes* (common chimpanzee) and *P. paniscus* (bonobo). Chimpanzees are estimated to have diverged from Hominini 3.7-7.5 Ma (12). Only *Pan troglodytes* were analysed in this study.

S1.3.1 Pan troglodytes (Common chimpanzee)

Common chimpanzees are an endangered species of chimpanzee native to parts of west, central, and east Africa. Individuals belonging to three subspecies, eastern chimpanzee (*Pan troglodytes schweinfurthii)*, Nigeria-Cameroon chimpanzee (*Pan troglodytes ellioti*) and western chimpanzee (*Pan troglodytes verus*), were sampled in this study. Eastern chimpanzees are estimated to have diverged from the other subspecies 540-630 ka, and Nigeria-Cameroon and western chimpanzees are estimated to have split around 250 ka (37). Chimpanzees live in diverse habitats, including savannas and forests (both deciduous and evergreen), and they are among the most consistently frugivorous primates, maintaining frugivory even during times of fruit scarcity (38). As such, they can be considered ripe fruit specialists (39). Chimpanzees are highly selective feeders, and their diet quality is similar even across diverse landscapes. As a result, chimpanzee habitat quality is a poor predictor of diet quality (38). Fruits from trees and climbers make up the majority of the chimpanzee diet, but herbs and insects are also eaten, and vertebrate animal prey is occasionally consumed. Consistent with their primarily frugivorous diet, chimpanzee molars exhibit high microwear complexity and low anisotropy (6). Ripe fruits are a strongly preferred food, and the main fallback foods are piths and leaves, which are preferred over unripe fruits (39, 40). Food selection appears to be driven by sugar content rather than antifeedant avoidance (41), and many preferred chimpanzee fruits are rich in indigestible components. Chimpanzees in part reduce their intake of lignin and other toxic or digestibility-reducing compounds (such as tannins) through "wadging", or compacting unpalatable plant material (such as seeds or fibrous material in leaves or fruit skins) for expectoration. Wadging is not known to be performed by other non-human primates, such as gorillas, but it is similar to quid production by modern humans (42, 43). Despite a preference for sugary food content, dental disease in wild individuals, including calculus formation (Fig. S1 E-F), generally remains low compared to gorillas (24), although it can appear more extensive when associated with the loss of molar attached gingival tissue (44). In captive individuals, calculus occurs at lower amounts in younger individuals primarily on labial surfaces of anterior teeth, whereas in adult individuals it is prevalent across all teeth (45). The salivary composition of chimpanzees has been studied and was found to have a lower protein concentration than gorillas, but higher than modern humans (31). Major salivary proteins identified within chimpanzee saliva were also shared with gorillas and modern humans, but differed in concentration and glycosylation patterns (31). Salivary immunoglobulins were found at relatively high levels in chimpanzees and modern humans, but not gorillas (31), suggesting greater immunological activity in the chimpanzee and modern human oral cavity. Chimpanzees have two diploid copies of the AMY1 gene (46), and copy number variation has not been observed. Salivary amylase expression in chimpanzees is significantly lower than in modern humans (46, 47), gorillas (35), and starch-utilizing Old World monkeys, such as hamadryas and gelada baboons (47).

A total of 21 historic wild chimpanzee dental calculus samples are included in this study. Dental calculus from eastern chimpanzees (n=7) was obtained from Kibale National Park histological collection (KNP; Fort Portal, Uganda). Dental calculus samples were collected from wild individuals who had died between 1992 and 2007 from illness or injury. All samples were collected postmortem after a brief period of burial, followed by storage in the KNP histological collection. Permissions for dental calculus sampling and destructive analysis were provided by R.W. of the Kibale Chimpanzee Project. Dental calculus from Nigerian-Cameroon chimpanzees (n=6) was obtained from the Hamann-Todd Collection at the Cleveland Museum of Natural History (CMNH; Cleveland, USA). All skeletons were obtained from wild-shot individuals from Africa in the early 20th century (see *Gorilla* section above for more details about the collection). Permissions for dental calculus sampling and destructive analysis were provided by the collections manager of the Physical Anthropology department. Dental calculus from western chimpanzees (n=7) was obtained from the chimpanzee skull collection housed in the Department of Palaeoanthropology at the Senckenberg Research Institute and Natural History Museum Frankfurt in Germany. Permission for dental calculus sampling and destructive analysis were provided by the curator. The chimpanzee skeletal material was obtained from local villages in Liberia where the chimpanzees had been hunted, killed, and cooked for food (48). Data for one additional western chimpanzee sample from Gola Forest, Sierra Leone was obtained from a public data repository (1). Further details about all chimpanzees included in this study are provided in Data S1.

S1.4 Homo

Homo is a genus in the subfamily Homininae and the tribe Hominini (hominins), and it includes one extant species: *H. sapiens* (modern human), and multiple extinct ('archaic') species, of which Neanderthals are best understood. We use the term 'modern' here as it is a widely-known term, although we acknowledge the ongoing and active debate surrounding the definitions and descriptions of *Homo* groups (49) Modern humans and Neanderthals are estimated to have diverged approximately 400-740 ka (50, 51).

S1.4.1 Neanderthal

Neanderthals are a group of *Homo* that inhabited parts of the Middle East, Europe, and Asia during the Middle and Late Pleistocene, and they went extinct during the Middle to Upper Palaeolithic transition approximately 40 ka (52). A total of 17 Neanderthal dental calculus samples are included in this study (Data S1). The age of the samples range from early Neanderthals dating to approximately 100 ka at Pešturina (53) to Neanderthals dating to approximately 70 ka at De Nadale Cave (54) and Banyoles (55), to late Neanderthals dating between 55 and 47 ka such as at Fumane Cave (56–58), the Troisième caverne of Goyet (59), and Zafarraya (60) (chronological difficulties described in 61, 62–66). Analysis of Neanderthal genomes has revealed that they were highly inbred (67), suggesting a long history of population bottlenecking and small population sizes. Neanderthals admixed with both Denisovans (68, 69) (another 'archaic' *Homo* species) and Upper Palaeolithic modern humans (70–72). With respect to Neanderthal diet, bone collagen stable isotope analysis and faunal assemblages have been conventionally interpreted to indicate a near complete reliance on large game hunting (73, 74), making them top level predators. Although some researchers have described them as carnivores (74), other research has complicated this view (75). A number of Neanderthal sites have yielded evidence for a broader resource exploitation of birds (i.e., at Fumane (76)), small prey (77), and even fish, shellfish, and marine mammals (78, 79). Dental enamel micro- and mesowear suggests that Neanderthals living in southern wooded environments consumed a broader range of dietary items than those living in more northern or open environments (80–82). Charred seeds and plant microfossils of starch-rich foods have been recovered at many Neanderthal sites (83, 84), and starch granules and phytoliths have been identified within Neanderthal dental calculus, providing a glimpse of the diverse range of starchy foods in the diet, including grass seeds, legumes, fruit, and underground storage organs (USOs), that otherwise leave few traces in the archaeological record (85–88). Elevated nitrogen stable isotope values that have been observed in Neanderthal bone collagen have also been reevaluated in light of broader patterns of Pleistocene isotopic ecology (89, 90), and recent compound-specific isotopic analysis of glutamic acid and phenylalanine indicates that plants may have accounted for as much of 20% of their dietary proteins (90–92). An attempt has been made to characterise Neanderthal diet using eukaryotic DNA sequences recovered from calculus (1), but improbable findings (93) and unresolved questions regarding sediment leaching, aDNA characteristics authentication, and sample treatment history have raised doubts about the study methodology, so this will not be further considered here. At present, a broad body of current evidence suggests that Neanderthals both hunted and gathered and that wild starchy foods were likely consumed when present in the environment.

To date, two high coverage Neanderthal genomes (Vindija and Altai) have been published (67, 69), but with the exception of *MUC7* and *AMY1*, the genes of salivary proteins in Neanderthals have not been studied. However, it is known that Neanderthals have two diploid copies of the *AMY1* gene, and copy number variation has not yet been observed. Modeling of the evolutionary history of the *AMY1* gene in modern humans suggests that the high copy number observed in modern humans postdates the modern human-Neanderthal split (94). The gene MUC7 encodes an abundant salivary protein involved in microbial aggregation, and variants in this gene in modern humans are associated with microbiome structure at oral and pharyngeal mucosal surfaces, as well as supragingival plaque (95).

A total of 18 Neanderthal dental calculus samples were included in this study. These samples originate from six sites in Spain (Banyoles (55), La Güelga (96), Cueva de Valdegoba (97), El Sidrón (1), Cueva del Boquete de Zafarraya (60, 98), Sima de las Palomas del Cabezo Gordo (99–101), one site in Serbia (Pešturina Cave (53, 102)), two sites in Italy (Fumane (103), de Nadale (54)), and two sites in Belgium (Troisième caverne of Goyet (59, 104), Spy (1)). Permission for sampling and destructive analysis for

plant microfossils of the dental calculus sample from Banyoles (BAN, n=1; ca. 66 ka) was obtained from the private owners of the mandible by D.C.S.G and R.C.P, and later extended to genetic analysis of a subsample from the original sampling. Permissions for sampling and destructive analysis of the dental calculus samples from La Güelga (GUE, n=1; >45 ka), Cueva de Valdegoba (VLD, n=1; >45 ka), and Sima de las Palomas del Cabezo Gordo (SPM, n=2; ca. 63-37 ka) were provided by the respective excavation leaders and head (palaeo)anthropologists (M.J.W. for SPM) to D.C.S.G (and R.C.P for VLD). Samples from La Güelga were obtained from the Universidad Nacional de Educación a Distancia, the sample from Cueva de Valdegoba was obtained from Universidad de Burgos, and calculus from Sima de las Palomas Cabezo Gordo was taken as a subsample from previous analyses (88) and the remains now curated at the Murcian Association for the Study of Palaeoanthropology and the Quaternary (MUPANTQUAT). The dental calculus sample from Cueva del Boquete de Zafarraya (ZAF, n=1; ca. 45 ka) was obtained and sampled at the Museo de Málaga (Spain), with permissions for sampling and destructive analysis provided to D.C.S.G and R.C.P. by the Junta de Andalucía. The dental calculus sample from Pešturina (PES, n=1; ca. 102 ka) was obtained from the University of Belgrade and permissions for sampling and destructive analysis were given by the director of the Center for the Palaeolithic and Mesolithic Studies, Faculty of Philosophy, University of Belgrade (D.M.). Dental calculus samples from Fumane cave (FUM, n=3; ca. 54-48 ka) and de Nadale cave (GDN, n=2; ca. 70 ka) were obtained from the University of Ferrara and permissions for sampling and destructive analysis were provided by the University of Ferrara (M.P.) and Soprintendenza archeologia, belle arti e paesaggio per le province di Verona, Rovigo e Vicenza (SAPAB – VR). Dental calculus samples from the Troisième caverne of Goyet (GOY, n=2; 45-40 ka) were collected from remains housed in the Anthropology and Prehistory collections of the Scientific Heritage Service of the Royal Belgian Institute of Natural Sciences. Permission for dental calculus sampling and destructive analysis were provided by the curator of the Anthropological and Prehistory Collections, and by the Troisième caverne of Goyet project coordinator (H.R). Finally, data from two additional Neanderthal sites at El Sidrón (n=2; ca. 49 ka) and Spy (n=2; 40 ka) were obtained from a public data repository (1). Further details about the individuals sampled are included in Data S1.

S1.4.2. Homo sapiens (Modern Human)

Modern humans (*Homo sapiens*) evolved in Africa approximately 300 ka (105, 106) and subsequently radiated throughout the Old World ca. 220-40 ka (107–111), admixing with other members of *Homo*, at least with Neanderthals and Denisovans, in the Near East/Europe and Asia, respectively (70–72). By the terminal Pleistocene (ca. 15 ka), modern human populations had expanded to the Americas and now inhabited every continent except Antarctica (107). Pleistocene modern human populations were diverse, practicing a wide variety of gathering, hunting, and fishing food acquisition strategies that were adapted to local environments. Early gathered foods included starchy geophytes (corms, bulbs, tubers, and rhizomes), which were likely cooked, and charred rhizomes and parenchyma fragments are attested in South Africa as early as 170 ka at Border Cave (112) and at 120 ka at the Klasies River site (113). The Klasies River site has also yielded modern human remains with dental calculus containing starch granules from grass seeds (Triticeae) and probable USOs, and similar grass and USO starch granules have also been reported in modern human dental calculus at Skhul, Israel (ca. 120 ka) (85, 114). The continued targeting of geophytes through the end of the Pleistocene in Africa, Eurasia, and Australia is supported by a wide body of evidence, including charred geophytes dating to 55 ka in Australia (115), starch granules and charred tubers and corms dating to 45 ka in Malaysia, and starch granules and charred plant tissues dating to ca. 30 ka in Europe and central Africa (85, 116–118). At the same time, diverse modern human populations in western Eurasia also began experimenting with collecting and grinding wild grains (117, 119, 120). Following the Younger Dryas and the onset of the Holocene ca. 11.7 ka (121, 122), agriculture began to develop in different locations, with early centers of domestication forming in the Near East and China (123, 124), and later in Mexico, Northern and Andean South America, Eastern North America, Central and West Africa, South Asia, Japan, and New Guinea (125).

In this study, we focus on prehistoric, and historic modern human populations from Africa and Europe, and contemporary modern human populations from Europe. Although the chronological systems of these two continents differ, the African Later Stone Age (beginning ca. 50-20 ka (126, 127) roughly corresponds in time with the European Late Middle and Upper Palaeolithic (ca. 50-11 ka). The earliest modern

humans in our study date to the Upper Paleolithic ca. 30 ka, approximately 15 ky after modern humans first arrived in Europe and 10 ky after the extinction of Neanderthals. Although these humans did not directly interact with Neanderthals, earlier populations did (128, 129). Subsequent post-glacial warming ca. 14 ka during the Bølling-Allerød interstadial is associated with major modern human population turnovers in Europe, and in particular an influx of ancestry from the Near East (130, 131). Following the Mesolithic period (which has a large range of start and end dates across Europe (132–138), Europe underwent a Neolithic transition beginning ca. 8.5-5 ka (139–141) as farming and animal husbandry largely replaced foraging in most regions. This Neolithic transition was initiated by a migration of Near Eastern farmers who replaced and admixed with local populations (142, 143). In Africa, the transition from foraging to food production was slightly later and more variable (144, 145). Ruminant pastoralism began ca. 7 ka in the Nile Valley and Sudan, and subsequently spread to the south-central Sahara by 6 ka (146). With the desertification of the Sahara, pastoralism continued to move southwards, spreading into West Africa and the East African savannas by 4 ka, and into Southern Africa within the last few millennia (146). Crop cultivation of Southwest Asian plant domesticates, such as wheat and barley, appeared by 5 ka in the Egyptian and Nubian Nile Valley and by 2.8 ka in the northern Horn of Africa (147, 148). Indigenous African plant domesticates appear as early as 4 ka in the West African Sahel, such as pearl millet, followed by the emergence of a wide range of African crops across much of the continent during the subsequent three millennia.

Modern human oral biology and microbiology is well-studied, and more than 600 prevalent microbial species are characteristic of the modern human oral microbiome (149). Dental calculus formation is nearly ubiquitous among modern human populations, but varies in abundance and prevalence within populations (150). Modern human saliva is well-studied, and to date more than 5,000 host proteins have been identified in saliva, of which salivary α-amylase is the most abundant, accounting for 4.8% of the total salivary proteome (151), and proteins in the salivary PRP family collectively account for 20-30% of all salivary proteins (w/w) (152). Although most salivary proteins are shared among the great apes, humans lack latherin, a surfactant protein likely involved in food lubrication, due to a frameshift mutation in the gene (31). Modern human saliva is very watery, having the lowest protein concentration of the great apes, a property that may be related to vocalization or food lubrication in modern humans (31). Proteomic studies have also identified 130 proteins in the modern human acquired enamel pellicle (AEP), of which salivary α-amylase is also the most abundant, after excluding keratin contaminants (153). Alpha-PRPs (aPRPs), histatins, and statherins are also present in modern human AEP, but they are less detectable using current methods (153), making them difficult to quantify, and evidence suggests aPRPs rapidly degrade in the AEP after initial formation (154). In contrast to non-human primates and ‘archaic’ humans, modern humans exhibit high salivary α-amylase (*AMY1*) copy number variation, with a mean of ~6-8 diploid copies in most populations, and a reported range 1-30 diploid copies overall (155, 156). The copy number expansion observed in modern humans is argued to relate to dietary shifts in early human evolutionary history, and specifically to an increased reliance on starch-rich geophytes, such as rhizomes, bulbs, corms, and tubers (46), which are important edible resources in grasslands and arid environments (157). Further correlations between copy number and dietary starch reliance in living populations today are less clear (155), but recent work on archaeological populations has found *AMY1* diploid copy numbers of 5, 6, and 13 among European Mesolithic hunter gatherers (158), confirming that *AMY1* copy number expansion occurred prior to agriculture (155), and a copy number of 6 has been imputed for a 45 ka genome from Ust’Ishim in Siberia (94). Further genetic modeling indicates AMY1 copy number expansion was likely already established in modern human populations prior to their expansion out of Africa, and possibly as early as the Middle Pleistocene, shortly after the split with Neanderthals (94).

To account for biogeographic variability as well as changes that may have occurred in the modern human oral microbiome across major lifestyle and dietary transitions, we analysed 39 ancient and historic modern human dental calculus samples across 34 individuals from Africa and Europe dating prior to the adoption of food production (n=25) and after the adoption of food production but before use of therapeutic antibiotics (n=14), as well as 18 present-day calculus samples from two populations in Europe (see Table S1 for individual site details and citations). We analysed thirteen samples dating to the African Later Stone Age prior to food production from the sites of Oakhurst, South Africa (n=5; ca. 7-5 ka), Taforalt, Morocco (n=7; ca. 15-14 ka), and Mota, Ethiopia (n=1; ca. 4.5 ka). We analysed thirteen samples dating to the European Upper Palaeolithic or Mesolithic prior to food production from the sites of Pavlov, Czech

Republic (n=1; ca. 30 ka), Dolní Věstonice, Czech Republic (n=2; ca. 31 ka), El Mirón, Spain (n=1; ca. 18.7 ka), Rigney, France (n=1; ca. 15.5 ka), El Collado, Spain (n=6; ca 9.6-8.4 ka), and Große Ofnethöhle, Germany (n=1; ca 8.3 ka; Fig. S1G). We analysed ten samples from African food producing but pre-antibiotic sites at Essasoira (Mogador), Morocco (n=5, ca. 280 ya) and Polyoak, South Africa (n=5, ca. 200 ya). We analysed four samples from European food producing but pre-antibiotic sites from Lisbon, Portugal (n=3, ca. 280 ya) and Basque region, Spain (n=1, ca. 1.9 ka). Present-day dental calculus was obtained from dental clinics in Jaen (n=10) and Valencia (n=8), Spain, and represents a sampling of the general population, including both periodontally healthy and diseased individuals (2).

Permission for destructive DNA analysis for samples for Oakhurst (OAK) and Polyoak (PYK) (South Africa) was provided by the UCT Human Skeletal Repository Repository Committee. Ethics approval was provided by the UCT Human Research Ethics Committee (Division of Clinical Anatomy and Biological Anthropology, University of Cape Town) under approval #715/2017, which was obtained in collaboration with V.E.G. Additional permission was granted from the South African Heritage Resources Agency and Heritage Western Cape for this study; part of the process included community consultation. Permission for collection and destructive analysis of calculus samples from individuals from Taforalt (Morocco; TAF) for archaeobotanical and archaeogenetic analysis were provided by the director of the National Institute of Archaeology and Heritage (INSAP), with export permission was granted to A.G.H. The Ethiopian Authority for Research and Conservation of Cultural Heritage and the National Museum of Ethiopia granted permission for destructive analysis of calculus taken from Bayira who was discovered in Mota Cave, Ethiopia. Samples from the Czech sites of Pavlov (PLV) and Dolní Věstonice (DLV) were taken as subsamples from previously published archaeobotanical analysis (159). Permission for additional destructive sampling for DNA analysis on these subsets was obtained from the collections curator (J.Sv.) at The Czech Academy of Sciences, Institute of Archeology (Brno), Centre for Palaeolithic and Palaeoanthropology in Dolní Věstonice. The sample from El Mirón (EMN) was obtained by subsampling previously collected calculus for archaeobotanical analysis (118, 160). Permission for destructive analysis of this sample was obtained by the excavation leaders (L.G.S., M.M.G) and obtained from the Universidad de Burgos. Dental calculus from the Rigney I (RIG) specimen was housed in the collections of the Regional Archaeological Service of Bourgogne-Franche-Comté (Besançon, France), who also provided the permission for sampling and destructive analysis (C.C.). Permission for sampling and destructive analysis for biomolecular methods of the samples from El Collado (ECO) was received from the Museo de Prehistòria de València to D.C.S.G. Dental calculus from the modern human individual from Große Ofnethöhle (Germany, Collection ID: OSUT 4043; OFN) was obtained from the Osteological Collection housed at the Institute for Archaeological Sciences, Section Palaeoanthropology at the University of Tübingen in Germany. Permission for dental calculus sampling and destructive analysis were provided by the director of the collection (K.H.). Permission for destructive DNA analysis for samples from Essaouira (Morocco; ESA), Lisbon (Portugal; LIS), and a sample from the modern day Basque Pyrenees region (Spain; BSQ) was provided by the curator of the Rudolf-Virchow-Sammlung in Berlin, Germany (B.T.). The data from present-day samples from Jaen (Spain; JAE) have been previously published in Velsko et al. (2). Permission for collection of present-day samples from VLC (Valencia, Spain) for analysis of oral microbiome and food debris was given under informed consent by anonymous patients undergoing routine dental healthcare cleaning by odontologists. To protect their privacy, human DNA has been removed from the sequencing data by discarding reads that map to the HG19 human reference genome. Further details about the individuals sampled are included in Tables S1 and S2.

We also generated 8 new accelerator mass spectrometry dates for modern humans, supplementing 36 previously published radiometric and trapped charge dates, for a total of 44 directly or indirectly dated ancient individuals in this study (Data S1).

S2. Laboratory procedure

For the purposes of this study, all calculus samples other than the present-day modern human samples (JAE and VLC) were treated as ancient during laboratory procedures and data analysis. This includes 20th and 21st century primate samples because, despite being relatively recent, they had undergone maceration and/or burial processes prior to analysis.

All ancient and historic calculus samples except from the OME and KNP individuals were extracted in a dedicated ancient DNA cleanroom facility at the Max Planck Institute for the Science of Human History (MPI-SHH). OME and KNP samples were extracted in a dedicated ancient DNA cleanroom facility at the Laboratories of Molecular Anthropology and Microbiome Research (LMAMR; University of Oklahoma). DNA was extracted using standard ancient DNA protocols following Dabney et al. (161), with minor adaptations for each laboratory. Present-day modern human calculus samples (JAE and VLC) were extracted in a modern DNA laboratory at the MPI-SHH using a DNeasy PowerSoil DNA extraction kit. The PowerSoil kit is widely used in both medical and ecological studies and has become a field-wide standard for modern microbiome research (162–164). Prior research has shown that both protocols yield equivalent microbial profiles when applied to ancient microbiome samples (165), but the Dabney protocol allows greater ancient DNA recovery. For this reason, we chose to extract the present-day calculus samples using the DNeasy PowerSoil kit and the ancient microbiome samples using the Dabney protocol.

All library constructions for shallow sequencing analysis (non-UDG treated libraries) and deep sequencing analysis (with full-UDG treatment) libraries were performed at the MPI-SHH. All library constructions for Illumina sequencing followed the double stranded, dual-indexed protocols of Meyer and Kircher (166, 167), with minor modifications. Negative controls were included in each extraction and library batch. Details for DNA extraction and library construction are provided below. Sample codes followed the internal conventions of the internal MPI-SHH LIMS system. See the External Data Repository Section R6.2.3 for details.

S2.1 Dental calculus collection

Unless otherwise stated, ancient and historic dental calculus samples were collected at their respective institutions or in the sampling room at the dedicated ancient DNA laboratory at the MPI-SHH. Prior to sampling, surfaces were cleaned with NaOCl, NaOH, or covered in aluminium foil to create a clean working environment. Wearing latex or nitrile gloves and a facemask, the tooth, cranium or mandible was held over either a constructed foil catchment bowl with a microcentrifuge tube perforated through the bottom (for loose teeth) (168), or a plastic weigh boat lined with a folded wax-weigh paper or aluminum foil square in the middle (for intact crania or mandibles). Calculus deposits were then gently removed using the broad edge of a sterile dental scaler or scalpel into the corresponding container and transferred to a microcentrifuge tube. Gloves and foil were replaced after each sampling and all utensils were thoroughly wiped down with bleach or alcohol wipes prior to reuse. Photos before and after sampling were taken where possible. Calculus was pooled from multiple teeth depending on the amount of available deposits, and further details regarding the specific teeth sampled are provided in Data S1. Full Protocols for both methods can be seen in https://dx.doi.org/10.17504/protocols.io.7hphj5n and https://dx.doi.org/10.17504/protocols.io.bqecmtaw).

Calculus from El Mirón (EMN), Dolní Věstonice (DLV), and Pavlov (PLV) were collected prior to this study, and the collection methods are described in (118, 159). Present-day modern human dental calculus for Jaén and Valencia was collected as described in (2) and stored frozen until analysis.

S2.2 DNA extraction

All ancient, historic, and non-present day modern human dental calculus samples were processed as follows, unless otherwise noted below. For pretreatment, if samples had evidence of consolidants, large deposits underwent a 15 min 0.5 ml acetone (VWR BDH Chemicals) soak without rotation, were spun down (1 min, 15,800 RCF), and acetone removed via pipette and evaporation. For surface DNA decontamination, samples of 2-10 mg were placed in a weigh-boat and UV irradiated for 1 min each side followed by a 15 min wash of 1 ml pH 8 0.5 M EDTA (Life Technologies) (169) with 15-20 min rotation at room temperature (RT). For very small or powdery calculus, UV irradiation was performed on open tubes for 1 min and the EDTA wash was performed with a short vortexing instead of the 15 min incubation. Samples were centrifuged (2 min, 8,000 RCF) and the wash supernatant removed. Due to concerns for sample loss because of very powdery and small sample size, EDTA surface decontamination was not performed for samples from MTM, MTS, MTK, LOB, CDC, GOY, and MOA. Following decontamination, calculus chunks were crushed using a pestle and the resulting fine powder was decalcified by incubation

in a solution of 1 ml 0.5 M EDTA for 24 h at RT or 37˚C under gentle rotation, and then additionally digested for 48-72 h at RT following the addition of 50 µl 10 mg/ml Proteinase K (Sigma-Aldrich). The suspension was then centrifuged and supernatant transferred into a binding buffer, following Dabney et al.(161), but substituting the High Pure Viral Nucleic Acid Large Volume Kit (Roche) in place of the MinElute PCR Purification Kit (Qiagen), as described in (170). In brief, DNA was bound onto silica membrane filter columns and purified following the High Pure Viral Nucleic Acid Kit workflow by twice washing the membrane with the manufacturer's wash buffer. Before elution, a dry spin of the silica membrane was performed (30 sec at 8000 RCF, rotate tube 180º, spin again for 30 sec at 8,000 RCF, at room temperature) to remove excess ethanol. DNA was eluted in 100 µl EB-Buffer (Qiagen) with added 0.05% Tween 20 across two eluting steps (1 min each, 8,000 RCF). DNA concentrations were measured with a Qubit 3.0 (Life Technologies) High Sensitivity Kit for double-stranded DNA (Life Technologies).

Calculus from the site KNP was extracted in the ancient DNA facility at LMAMR, as above but with the following laboratory-specific adaptations. Following the initial 24 h decalcification, 250 µl of mixed pellet/EDTA was removed for other analyses. This volume was replaced with 250 µl of fresh 0.5 M EDTA, and 100 µl of proteinase K (Qiagen) was added and the solution. Digestion proceeded for 96 h at RT under gentle rotation. Following digestion, samples were extracted with the Dabney et al. protocol (161) using the MinElute PCR Purification Kit (Qiagen), as described in Ozga et al. (171). For historic calculus from individuals from the site OME, DNA was extracted also at LMAMR but following in Ozga et al. (171).

Present-day calculus extraction and library preparation of the VLC samples were performed using a DNeasy PowerSoil DNA extraction kit (Qiagen) as described for JAE as described in (2). Per-sample details of the decontamination extraction procedure can be seen in Data S1.

Bench protocols for both ancient and modern dental calculus DNA extraction procedures are accessible at: https://dx.doi.org/10.17504/protocols.io.bqbmmsk6 and https://dx.doi.org/10.17504/protocols.io.7p8hmrw respectively.

S2.3 Library preparation and sequencing

We produced two sets of libraries for this project. We first constructed libraries without uracil-DNA glycosylase (UDG) treatment to retain characteristic deaminated cytosines (resulting in uracils, interpreted as Ts by DNA polymerases) indicative of aDNA (designated the 'shallow sequencing' or 'screening dataset'). We then used a subset of the extracts to construct UDG-treated (172) libraries, which removes aDNA associated damage, which we then deeply sequenced ('deep sequencing dataset' or 'production' dataset). For a summary, see Data S1 and External Data Repository File R1. Bench protocols are available as follows: non-UDG library preparation for ancient DNA from dental calculus (https://dx.doi.org/10.17504/protocols.io.bqcsmswe); full-UDG library preparation for dental calculus ancient DNA (https://dx.doi.org/10.17504/protocols.io.bmysk7we) and modern dental calculus DNA (https://dx.doi.org/10.17504/protocols.io.bqd7ms9n); indexing and amplification/pooling (https://dx.doi.org/10.17504/protocols.io.bakticwn and https://dx.doi.org/10.17504/protocols.io.beqkjduw).

S2.3.1 Shallow Sequenced Dataset

Libraries were generated for all samples at the MPI-SHH following previously described methods (166 with full modified protocol at: https://dx.doi.org/10.17504/protocols.io.bqcsmswe), unless otherwise noted. Prior to library construction, DNA from present-day clinical calculus (JAE and VLC) was sheared to a target length of 200 bp using a Covaris M220 Focused-Ultrasonicator. For all samples, an input up to 20 µl of DNA extract (to a maximum of 100 ng DNA) was used to create each library. Blunt ends of the DNA were repaired with 0.4 U T4 DNA Polymerase and 0.024 U T4 Polynucleotide Kinase (New England Biolabs), including NEB Buffer 2 (New England Biolabs), 1 mM ATP (New England Biolabs), 0.8 mg/ml BSA (New England Biolabs) 0.25 mM dNTPs for present-day samples/0.1 mM dNTPs for ancient samples (Thermo Scientific), and balanced with water. Reactions were incubated for 15 min at 15°C followed by 15 min for 25°C. Afterwards, extracts were purified by MinElute Purification (Qiagen) following the manufacturer's protocol, and eluted in 20 µl EB-Buffer (Qiagen) containing 0.05 % Tween 20 (Sigma-Aldrich). Sequencing adapters were ligated with the Quick Ligation Kit (New England Biolabs) using 1x

Quick Ligase Buffer, 18 µl of the blunt end repaired DNA and 0.25 uM of the Adapter Mix, and incubated for 20 min at 22°C. DNA again was purified with the MinElute Purification Kit (Qiagen), but eluting in 22 µl of EB/Tween buffer. Adapter fill in was performed with 0.4 U Bst 2.0 DNA Polymerase (New England Biolabs) using 1x Isothermal Buffer (New England Biolabs), 0.25 nM dNTPs for present-day samples/0.125 mM dNTPs for ancient samples (ThermoScientific), and balanced with water and 20 µl of the adapter ligated DNA. Incubation was for 30 min at 37°C followed by 10 min for 80°C. Quantification of libraries was carried out with IS7 and IS8 primers in a quantitative PCR (qPCR, DyNamo SYBR Green qPCR Kit; Thermo Fisher Scientific) on a LightCycler 48 (Roche), the libraries were split into 1*10^10 Copies DNA per reaction for the following indexing step. PfuTurbo DNA Polymerase (Agilent) was used to dual-index libraries (167) with 0.25nM dNTPs for present-day samples/0.38 mM dNTPs for ancient samples, 0.3mg/mL BSA, and 0.2uM each of P5 and P7, and amplified in a 100 µl reaction, after which a MinElute-Purification (Qiagen) was performed and the samples eluted in 50 µl EBT. The post-indexing libraries were then quantified with IS5 and IS6 Primers with the qPCR quantification above. Afterwards they were then quantified with an Agilent 4200 TapeStation NucleicAcid System with the D1000-Kit (Agilent). Every library that showed a concentration under 10 nM underwent a second amplification using Herculase II Fusion DNA Polymerase (Agilent) including 2-4 µl of the indexed library in a 100 µl reaction with 0.25mM dNTPs and 0.4uM of each of the IS5/IS6 Primers. After MinElute-Purification, the libraries were eluted in 20 µl EBT and quantified again on the Agilent 4200 (see above). An equimolar pool of each batch of non-UDG treated libraries (negative controls separately) were then prepared for sequencing.

S2.3.2 Deep Sequenced Dataset

In order to perform more detailed analyses of specific genomes and genes, we selected a subset of individuals for UDG treatment and deeper sequencing: *Alouatta*, OME002, OME003, OME005; *Gorilla*, DJA002, MTM009; *Pan*, EBO003, KNP001, KNP004, KNP005, KNP009; *Homo* (Neanderthal), FUM002, GOY005, PES001; *Homo* (pre-agricultural), ECO002, ECO004, EMN001, OAK002, OAK005, TAF008 (tab labelled ‘Deep Sequenced Dataset’ in Data S1, and see section S6 for selection procedure). UDG treatment was performed following the protocol of Briggs et al. (172) to remove damage (cytosine to uracil deamination) for more accurate SNP calling. In brief, up 30 µl of extract (to a maximum of 100 ng DNA) was UDG treated by incubating the extract for 3 h at 37°C with 0.06 U of USER Enzyme (New England Biolabs), 0.4 U of T4 PNK (New England Biolabs) NEB Buffer 2 (New England Biolabs), 0.1 mg/ml BSA (New England Biolabs), and 0.3 mM dNTPs (New England Biolabs), balanced with water, up to a total volume of 75 µl. After the incubation, 0.115 U of T4 DNA Polymerase (New England Biolabs) was added to continue with a blunt end-repair. Following an incubation at 25°C for 20 min and 12°C for 10 min, the DNA was purified with the MinElute PCR Purification Kit (Qiagen) following the manufacturer's protocol, but eluting into 20 µl of EB-Buffer (Qiagen) containing 0.05 % Tween 20 (Sigma-Aldrich). Sequencing libraries were built using 30 µl of UDG-treated DNA extract following the protocol described for the shallow sequenced dataset above.

For KNP samples, up to 30 µl of extract were UDG-treated as above followed by library preparation with a modified NEBNext DNA Library Prep Set (NEB E6076) and blunt-end modified Illumina adapters at LMAMR. In brief, 100 ng of DNA extract or 30 µl of extract was incubated for 3 h at 37°C with 10x NEB Buffer 2, 300 µM dNTP mix, 0.1 mg of BSA, 1 mM of ATP, 2 µl of T4 PNK (10 U/µl), and 3 µl of USER enzyme (NEB, 1U/µl) in a 50 µl reaction volume. Following incubation, 2.5 µl of the End Repair enzyme from the NEBNext Library Kit was added to the reaction and the sample incubated for 40 min at room temperature and 20 min at 37 °C. The reaction was purified with a MinElute column (Qiagen) and eluted with 30 µl of EB following a 5 min incubation at 37 °C. Illumina adapters (10 µM P5/P7) were ligated to the end-repaired sample with the NEBNext DNA kit in 50 µl reactions. Ligation reactions were purified with a MinElute column as described above and eluted in 30 µl of EB buffer. The adapters were filled in with 2 µl of BsT polymerase (NEBNext kit), 5 µl of reaction buffer in 50 µl reactions. Each reaction was incubated for 30 min at 37 °C and 20 min at 80 °C. qPCR was used to determine appropriate cycle number for indexing PCR. Indexing reactions were performed with unique dual 8 bp indexed primers in triplicate in 25 µl reactions with 0.25 µl of Phusion High-Fidelity DNA Polymerase (Thermo Scientific), 0.3 µM of each indexing primer, 200 µM of dNTPs and 4 µl of template. Reactions were denatured at 98 °C for 30 sec, followed by variable cycle numbers of 98 °C for 10 sec, 60 °C for 15 sec, 72 °C for 30 sec and a final 5

min elongation step at 72 °C. Triplicate reactions were pooled and purified with a MinElute column as described above. The resulting libraries were sent to the MPI-SHH for sequencing.

Per-sample details of the library construction procedure for both samples and controls are provided in External Data Repository Table R1 under 'Production' columns.

S2.4 Sequencing

S2.4.1 Dental calculus

Sequencing of all calculus DNA was performed on Illumina NextSeq 500 and HiSeq 4000 platforms at the MPI-SHH with paired-end 75 bp or single-end 150 bp cycle chemistries. Calculus sequencing depth targets for shallow sequencing data was 10 million post-merging reads (Data S1), and deep sequencing data was calculated per sample (see section S6 below) to achieve 5x coverage of selected genomes, resulting in a sequencing depth of 60-480 million post-merging reads per sample (Data S1).

S2.4.2 Controls

Extraction and library controls were sequenced on a HiSeq 4000 platform at the MPI-SHH using paired- or single-end 75bp chemistry to a target depth of 2 million reads (Data S1). A subset of controls were sequenced on a MiSeq using reagent kit v3 with paired-end 75 bp chemistry at LMAMR.

In addition to the calculus samples and laboratory controls in this study, we also re-sequenced half-UDG treated shotgun libraries of ten archaeological bone samples from a previous study (170) to serve as archaeological environmental proxy controls (see section S3.1 and section S3.6 for rationale): ARS001, ARS004, ARS005, ARS010, ARS011, ARS012, ARS013, ARS015, ARS017, and ARS020. Bone environmental proxy controls were sequenced on a HiSeq 4000 using paired-end 50bp chemistry to a target depth of 30 million reads per library (Data S1).

S2.4.3 Sequence data

Per-sample details of the sequencing results and ENA accession numbers are reported in External External Data Repository File R2. The ENA project accession for all sequence files is under PRJEB34569. The re-sequencing data of archaeological bone control samples, containing all human and non-human sequences, fall under run accessions ERR3579689-ERR3579698. All uploaded data are adapter- and base-quality trimmed (20), however are not length filtered nor merged. Present-day modern human calculus samples have had reads mapping to the modern human reference genome (HG19) removed (section S3.2).

S3. Data processing and quality filtering

Unless otherwise noted, all table-based data manipulation was performed using R v3.6.1 (173) and the 'tidyverse' v1.2.1 (174) set of packages - in particular readr (175), tidyr (176), dplyr (177), ggplot2 (178), purrr (179), magrittr (180), tibble (181), stringr (182). Other regularly used packages were data.table (v11.8) (183), RColorBrewer (v1.1.2) (184), ggtree (v1.16.5) (185), and patchwork (v0.0.1) (186). A graphical overview of data processing procedures is shown in Fig. S2. All code notebooks, scripts and commands and small data files are provided in the External Data Repository: https://github.com/jfy133/Hominid_Calculus_Microbiome_Evolution. A long-term Zenodo archive can be found at the following DOI: 10.5281/zenodo.3740493.

S3.1 Publicly available data

S3.1.1 Dental calculus

Sequencing data of additional calculus samples (1) were downloaded from the OAGR database (https://www.oagr.org.au/dataset/68) in pre-adapter clipped and merged format. In the event of unmerged

and/or orphaned singletons reads being in a separate file, these were concatenated together.

S3.1.2 Modern human microbiome and environment

Comparative sequencing data for source comparison analysis (termed here as comparative sources) were downloaded from the EBI ENA (https://www.ebi.ac.uk/ena), and NCBI SRA (https://www.ncbi.nlm.nih.gov/sra) archives. This included cave sediment data (187), 'traditional society' modern human fecal samples (188, 189), skin swabs from hand palms (190), supra- and subgingival plaque and fecal samples from the Human Microbiome Project (191), and additional 'industrialised society' fecal samples (192). SRR accessions were downloaded and converted to FASTQ with the fastq-dump function from sratoolkit (v2.8.0) (193). ERR accessions were downloaded directly in FASTQ format. Download commands and a list of ERR and SRR accessions can be seen in External Data Repository Section R5.2 and Data R4.

Finally, as an additional 'environmental proxy' control comparison source to represent the type of bacteria that colonise archaeological osteological material, we used 10 resequenced Late Bronze Age modern human femur samples (site code ARS; see above section S2.4.2). These samples, originating from Bronze Age Mongolia (Arbulag Soum, Khövsgöl, Mongolia), were selected because they had been previously shown to contain a high proportion (>90%) of environmental (soil) bacterial DNA that is typically found in archaeological skeletal material ((170) and section S3.4). All bone samples were obtained from femora, and thus are not expected to contain any oral bacteria. Tooth samples were avoided as environmental proxy controls because it was recently shown that approximately 20% of teeth are decomposed by bacteria originating from the oral microbiome (194).

S3.2 Sequencing Quality Control and Human DNA Removal

For newly sequenced samples, raw sequencing BCL files were converted to FASTQ with bcl2fastq (v2.20.0.422), and demultiplexed. Demultiplexed FASTQ files for shallow and deep sequencing libraries were then pre-processed following the EAGER pipeline (v1.92.55) (195) to remove adapters, trim low quality sequences, and merge overlapping paired reads. AdapterRemoval settings were --minlength 30 --minquality 20 --minadapteroverlap 1 with --trimns and --trimqualities. Unmerged reads and orphaned singletons were combined with merged reads and considered equivalent to merged reads. Note that the datasets from Slon et al. (187) and Weyrich et al. (1) were already clipped and merged, therefore this step was skipped for these files.

Because it is known that some bacterial reference genomes contain contaminating human sequences (196–198), we first removed all human DNA from our entire dataset of calculus samples, controls and comparative sources, by mapping our sequences to the modern human HG19 reference genome (GRCh37; NCBI RefSeq Accession: GCF_000001405) (199) . We did not map each set of samples to their respective host genomes, as the aim of this procedure was only to remove contaminating present-day human DNA that may confound analysis - i.e., present-day modern human DNA from sample and/or laboratory handling, or endogenous DNA that is similar enough to human DNA to map to contaminated regions of microbial reference genomes. EAGER settings are provided in the External Data Repository Section R6.1.2, and for human DNA removal we used the relaxed bwa aln (200) mismatch value (-n 0.01) (201), and map quality threshold at 0 to ensure all possible human reads are found. For human DNA statistics reporting, mapped reads were deduplicated with DeDup (195). BAM files with only unmapped reads were then converted back to FASTQ files using samtools fastq (202). FASTQ files with only unmapped reads were then concatenated together, when either: multiple calculus samples were derived from a single individual (e.g., TAF018.**A** and TAF018.**B**), multiple libraries were generated from the same extract with the same treatment (e.g., TAF018.A01**01** and TAF018.A01**02**), or when libraries were sequenced over multiple runs (e.g., pre-2018: TAF018.A0101**.171215** and TAF018.A0101.**180215**, or post-2018: TAF018.A0101**.SG1** and TAF018.A0101**.SG2**). For further information about library concatenation, see the External Data Repository Section R6.2.3. All preprocessing steps and scripts are described in External Data Repository Section R6.

S3.2.1 Shallow sequenced dataset

For the shallow sequencing dataset, we newly generated a total of 2,573,264,971 raw reads for ancient and historic African hominid and howler monkey dental calculus (mean: 25,477,871 ± sd 13,417,354; minimum: 159,310; maximum: 70,895,520), and 520,054,720 raw reads for present-day modern human calculus (mean: 65,006,840 ± sd 21,962,897 per individual, minimum: 31,782,890, maximum: 89,173,288). For negative controls, the mean and standard deviation were 2,166,638 ± sd 1,575,696. For environmental (archaeological bone) controls, the mean and standard deviation was 34,067,925 ± sd 9,619,220 reads. Per individual sequencing details for all newly sequenced libraries for this study of the shallow sequencing dataset are provided in Data S1 and External Data Repository Section R6 and Fig. R1).

After poly-G trimming of newly sequenced calculus (to remove NextSeq artefacts that can map to repetitive regions of the modern human reference genome), the proportion of human DNA per group was as follows: *Alouatta*, 0.737% ± sd 0.176; *Gorilla*, 5.705% ± sd 10.198; *Pan*, 0.594% ± sd 0.853; *Homo* (Neanderthal), 16.079% ± sd 21.805; *Homo* (Modern Human), 2.93% ± sd 13.901. For present-day modern human calculus, the mean and standard deviation of human DNA was 2.205% ± sd 5.599. For controls, the mean and standard deviation of human DNA was 3.168% ± sd 5.812 (External Data Repository Section R6.4 and Fig. R1C).

Following sequence quality filtering, preprocessing (adapter removal and read merging), and human DNA removal, the mean and standard deviation of ancient reads per group in the 'analysis ready' shallow sequenced dataset used for downstream analysis (including previously published calculus from Weyrich et al. (1) and Velsko et al. (2)) were as follows: *Alouatta* (*n* = 5), 11,266,889 ± sd 2,860,592; *Gorilla* (*n* = 29), 10,257,550 ± sd 7,083,471; *Pan* (*n* = 21), 11,448,782 ± sd 4,776,739; *Homo* (Neanderthal, *n* = 17), 15,572,537 ± sd 14,354,052; *Homo* (Modern Human, *n* = 34), 12,897,453 ± sd 5,488,578. The mean and standard deviation for present-day modern human calculus (*n* = 18) samples were 51,131,788 ± sd 22,278,698. Controls (*n* = 35) had a mean and standard deviation of 750,969 ± sd 615,420 reads. Additional summary figures can be seen in the External Data Repository Section R6.4 Fig. R1.

S3.2.2 Deep Sequenced dataset

For the deep sequenced dataset, we newly sequenced a total of 3,875,041,498 raw reads, with a mean and standard deviation of 203,949,553 ± sd 123,359,872.3 reads per UDG-treated ancient/historic calculus individual (minimum: 61,703,682; maximum: 479,828,768). After sequence quality filtering, preprocessing, and human DNA removal (all as above), the mean and standard deviation of reads per group in the 'analysis ready' deep sequenced dataset were as follows: *Alouatta* (*n* = 3), 110,247,421 ± sd 50,136,787; *Gorilla* (*n* = 3), 88,927,251 ± sd 59,587,055; *Pan* (*n* = 4), 155,248,637 ± sd 72,527,567; *Homo* (Neanderthal, *n* = 3), 133,092,916 ± sd 167,706,862; *Homo* (Modern Human, *n* = 6), 56,926,392 ± sd 25,792,155. Present-day modern human samples used for analysis with this dataset (JAE008, JAE014, VLC004, VLC009) are the same as for the shallow sequencing dataset above; no UDG treatment was performed as this is unnecessary for present-day samples that lack damage, and have 52,724,105 ± 16,956,171 reads. For UDG-treatment library negative controls (*n* = 6), these had an average of 497,092 ± 316,696 reads. Further preprocessing statistics for all newly sequenced samples for this study of the deep sequenced dataset are provided in Data S1. Additional information regarding library merging and sequencing statistics are provided in the External Data Repository Section R6.4 and Fig. R2.

S3.3 Taxonomic binning and classification

At present, there is no consensus on the most suitable taxonomic classifier or database for the assignment of taxonomy to ancient microbial short read data (203, 204). Because there is no one-size-fits-all classifier (204), we performed taxonomic binning and classification on all concatenated unmapped reads of the shallow sequencing dataset non-UDG libraries and SRR and ERR comparative sources using MALT, because of its previously demonstrated good overall classification performance on ancient microbial sequences (204), and because of its ability to produce useful alignment information for downstream ancient DNA characteristics assessment (205). Detailed information regarding the

commands and scripts used for the following steps are provided in the External Data Repository Section R7.1.

Our primary metagenomic binning method was to use the aligner MALT (v040) (206, 207) with the NCBI Nucleotide database ('nt', October 2017). As an additional check we also ran MALT against a custom Genome RefSeq ('RefSeq', October 2018) (208) database. The RefSeq database consisted of all bacterial and archaeon assemblies at scaffold, chromosome and complete levels, with a maximum of 10 randomly-selected genomes per species but with completeness prioritisation (i.e. reference > representative; complete > chromosome > scaffolds) and the addition of the modern human reference genome HG19 (199). Details on database construction for both nt and custom RefSeq databases can be seen in External Data Repository Section R4.1. A list of genomes used for the custom RefSeq Database can be seen in External Data Repository Data R10.

All shallow sequencing, controls and comparative source data was searched and binned using the MALT nt (External Data Repository Section R7.1.1) and RefSeq databases. For the RefSeq database, we additionally produced SAM files for downstream functional analysis (see External Data Repository Section R12.5). For the MALT alignment step, we followed Vågene et al. (207) and used a relaxed percent identity parameter of 85% to account for potentially damaged reads, and possibly more divergent endogenous microbial genomes from the modern human-biased reference entries in existing genetic databases. We also extracted the 'MinSupport set to:' field from the MALT log files to allow manual adjustment of minimum abundance thresholds. The resulting subsetted log file was then formatted to TSV, and downstream manual adjustment of reads was performed by multiplying the 'MinSupport set to' entry of the log files in the script (External Data Repository Section R7.1.2).

MEGAN6 CE (209) was used to export OTU (Operational Taxonomic Unit) tables at species and genus levels using the taxonNameToCount and summarised options. Two variants were created: one with all domains, and a variant with the option to exclude non-Prokaryotes turned on. The former were used for preservational screening purposes, whereas Prokaryotes-only tables were used for all downstream microbial analysis. The corresponding tree was then exported in newick format. The OTU tables generated using the nt database are provided in Data S2.The procedure, '.megan files', RefSeq OTU table, log file summaries, and tree files can be located in the External Data Repository Section R7.2.1 and Data R11.

For alignment to the NCBI nt database, across all individuals, laboratory controls and comparative sources, the group with the lowest taxonomic assignment rate was sediment with a mean percentage of taxonomically assigned reads (over all non-human reads) of 8.9% ± 2.4; library controls were the highest group with 52.7% ± 13.8 of reads assigned. Across ancient, historic, and present-day calculus samples, each group had a mean percentage of taxonomically assigned reads as follows: *Alouatta*: 18.7% ± 1.8, *Gorilla*: 30.5% ± 11.2, *Pan*: 17.5% ± 2.7, *Homo* (Neanderthal): 26.1% ± 17.5, *Homo* (Modern Human): 35.6% ± 15.2. For alignment to the NCBI RefSeq database, across all samples, laboratory controls and comparative samples, the lowest group was likewise sediment with a mean percentage of taxonomically assigned reads of 9.7% ± 2.8, and 60.5 ± 15.5 in library controls as the highest group. Across ancient, historic, and present-day calculus samples, each group had a mean percentage taxonomically assigned reads and standard deviation of: *Alouatta* 30.3% ± 3.4, *Gorilla* 39.1% ± 10.9, *Pan* 27.6% ± 5.3, *Homo* (Neanderthal) 35.6% ± 18.2, *Homo* (Modern Human) 50.5% ± 19.9. For more information on comparisons, see External Data Repository Section R7.3 and Fig. R3.

Comparison between the percentage of aligned sequences with taxonomic assignments showed an increase across all groups between the NCBI nt 2017 to custom RefSeq 2018 databases. Present-day gut (additional 24.1% assigned on average) and plaque (additional 33.5% on average) showed the greatest increases, whereas sediment showed the smallest (additional 0.8% on average). Estimation statistical analysis and Gardner-Altman plots were generated via the dabestr package (v.0.2.1, (210)). A paired mean comparison with 5000 bootstrap resamples showed a high confidence in the increase in the mean percentage of across all groups (additional 11.1%), and within ancient, historic, and present-day calculus only (additional 10.9%, see External Data Repository Section R7.3 and Fig. R4). The greater number of reads being taxonomically assigned in the RefSeq database is likely due to the inclusion of

many more modern human microbiome derived 'MAGs' (Metagenome-assembled genomes) since the download of the nt database, however these remain biased primarily towards industrialized societies (despite improvements in sampling diversity (211)), and therefore not directly useful for assignments to non-humans. We note that the custom RefSeq database is not directly comparable with the nt database, as the former does not include eukaryotic genomes other than *Homo sapiens*, so a portion of the metagenome (e.g., dietary and fungal DNA) is undetectable using this database. This suggests that most likely we have an underestimate in the amount of reads than can be taxonomically assigned, and, given an expansion of computational resources and a larger database, a greater fraction of reads could be potentially assigned.

For calculus and plaque specifically, in the both databases we observe that there are roughly similar percentages of assignment between calculus of present-day modern human individuals and that of present-day plaque (nt: ~50% and RefSeq: ~80%) - however this value decreases with older samples. We cannot currently state whether the decrease in the older samples is due to greater amounts of environmental background or past oral microbial diversity being not represented in modern databases. The custom RefSeq database shows an increase in the amount assigned reads (as above), however there is also a slightly increased amount of reads assigned to present-day plaque compared to present-day calculus (see External Data Repository Section R7.3 and Fig. R5). This may be due to more recently added genomes, as well as MAGs, having been sequenced from plaque rather than calculus sources, given industrialised populations tending to be the main isolation sources of such studies (211) and these populations having better dental care.

The ratio of bacterial/archaeal/viral to eukaryotic sequences, as detected by the NCBI nt database, and after removal of possible human DNA sequences during preprocessing, ranged from 0.02 to 1762.56 across all calculus samples, laboratory controls and comparative sources, with a mean and standard deviation of 170.48 ± 303.49 (see External Data Repository Section R7.3 and Fig. R6). The ratio of these alignments increased across groups: *Alouatta* 15.61 ± 5.84; *Gorilla* 32.90 ± 34.50; *Pan* 99.51 ± 81.24; *Homo* (Neanderthal) 58.86 ± 105.33; *Homo* (Modern Human) 275.15 ± 247.49. Visual inspection of OTUs in calculus samples with low ratios, e.g., BAN001.A0101 (0.17) or TAF016.B0101 (0.48) showed very high levels of eukaryotic taxa either known to have adapter-contaminated reference sequences (e.g. *Cyprinus carpio,* http://grahametherington.blogspot.com/2014/09/why-you-should-qc-your-reads-and-your.html) or to be common handling and storage contaminants, such as cotton (*Gossypium raimondii*) and environmental fungi (*Aspergillus glaucus (212)*).

Calculus from present-day modern humans generally had more bacterial/archaeal/viral alignments (494.61 ± 251.80) than eukaryotic alignments compared to present-day plaque (277.14 ± 334.42); whereas archaeological modern human calculus had a smaller ratios (Preagricultural modern humans 98.32 ± 108.54; Preantibiotic modern humans 244.71 ± 157.19, see External Data Repository Section R7.3 and Fig. R7). This suggests that the present-day calculus has a very high non-eukaryotic microbial biomass compared to ancient calculus. The low ratios in present-day plaque samples appear to be from high levels of taxa such as *Cyprinus carpio, Homo sapiens, Spirometra erinaceieuropaei* and *Onchocerca flexuosa.* The first two are well-known to have contaminated reference sequences, and the latter two appear to be specific to the Human Microbiome Project samples, and thus likely represent contaminants within the libraries used to generate those datasets.

S3.4 Preservation assessment and removal of low quality samples

Two common challenges for ancient DNA research are (1) the degradation of the original tissue of the organism of interest, and (2) high levels of environmental DNA contamination originating from postmortem burial and/or storage environments.

For ancient microbiome research in particular, decomposition and alteration of the original bacterial community can pose a major challenge. Decomposition is the postmortem process by which some microbes proliferate at the expense of others as nutrients stop being supplied by the deceased host. This process often involves a large influx and bloom of non-endogenous taxa (and occasionally the overgrowth of particular endogenous taxa (194, 213)) that invade host tissues, leading to the formation of a

'necrobiome' (214, 215) that breaks down decaying biological material (216). This microbial decomposition can also lead to further DNA fragmentation and damage of host and other endogenous (i.e., host microbiome) DNA, as the necrobiome enzymatically metabolises the components of older cells (217). However, degradation proceeds at different rates in different contexts depending on the nature of the burial microenvironment. For example, postmortem decay occurs slower in colder and drier environments (218, 219), which are less favorable for microbial growth, and this delays the breakdown of endogenous DNA and slows the overgrowth of environmental taxa.

A first task of any ancient DNA study is to screen for samples that have undergone heightened levels of decomposition such that they no longer preserve a reasonable representation of the state of the original microbiome prior to the death of the organism. As the modern human oral microbiome is relatively well-studied, we can utilise previously generated modern human oral microbiome reference metagenomes and taxonomic inventories (e.g., (191, 220)) as a benchmark to assess the prevalence and abundance of oral-associated taxa in ancient or historic calculus samples. This approach is possible because there is very little overlap at the species (and even genus) level in the microbial taxa that inhabit the modern human oral cavity and those that inhabit soil and other environmental contexts (149, 220–222), and source ambiguity typically only occurs for clades that are poorly characterised (e.g., *Actinomyces*) or paraphyletic (e.g., *Clostridia*, *Bacillus*).

Here we employ two tools to assess oral microbiome preservation to identify poorly preserved samples for removal from the dataset: (1) a new visualisation termed here as 'cumulative percent decay', and (2) SourceTracker (223). We applied these tools to dental calculus datasets from 124 individuals (106 ancient/historic, 18 present-day) and identified 89 individuals (71 ancient/historic, 18 present-day, with the NCBI nt database criteria) with sufficient ancient microbiome preservation for downstream analyses, alongside controls and comparative sources.

S3.4.1 Cumulative percent decay

To test whether a given sample retains an oral microbiome signature, we first plotted the 'decay' of a cumulative percentage of oral-associated taxa along a per-sample OTU abundance rank.

For this, we used the metagenomic binning data from MALT in the form of the OTU tables, to perform the following procedure:

1. Per sample, rank the identified taxa by abundance;
2. Compare each taxon to a database of common isolation sources of that taxa - assigning whether it has been commonly found in the oral cavity or not;
3. For each abundance rank, calculate the fraction of oral-derived taxa over all taxa that are observed up until this rank (as in step 1);
4. Plot a curve based on the percentage of oral-associated taxa (y-axis) until the rank (x-axis).

A schematic diagram of the concept can be seen in External Data Repository Section R8.1 and Fig. R8.

To generate an isolation source database, we first compiled a list of taxa as reported from the MALT alignment to the NCBI nt database of a variety of calculus samples that have been sequenced in our laboratory. We then extracted isolation sources of these taxa from two online databases: (1) the Human Oral Microbiome Database (HOMD, homd.org (220), all considered 'oral'), and (2) for those without a listing in the HOMD, from the 'isolation_source' metadata field in NCBI nt database, as reported when using the eutils package (https://www.ncbi.nlm.nih.gov/books/NBK25500/).

The results from the eutils command were then parsed in R. We then performed manual curation of the isolation sources of the two databases. Those with NCBI isolation sources matching the keywords 'oral, mouth, plaque, calculus, tartar, saliva, periodont' were listed as coming from the oral cavity. Manual checks removed certain keywords such as 'coral' or 'floral'. All other taxa were listed as 'unknown' source. We then manually curated the list to remove taxa known to also commonly reside in the general environment and certain non-oral specific pathogens (e.g., *Yersinia pestis*) that are in the HOMD. This R

code for isolation source database construction and the resulting ‘database’ are provided in the External Data Repository under Section R8.1 and Data R16.

In testing this approach using present-day comparative sources (Fig. S3A; External Data Repository Section R8.1 and Fig. R9), we observed that all non-plaque microbiome controls and bone environmental proxy sources never exceeded an oral fraction of 50% (nt) or 65% (RefSeq) across all abundance ranks. Most of the ancient and historic calculus samples exhibited similar curves to the present-day reference dental plaque - i.e., the majority of the most abundant taxa being derived from the oral cavity, and the proportion of non-oral taxa increased gradually only at lower abundance ranks. In addition, this pattern was observed across host genera, with even the outgroup *Alouatta* showing a strong oral profile using this approach, indicating that the method can be applied broadly to primate oral samples, and is not limited only to modern humans. We decided to utilise these observations to develop a threshold for the detection and removal of poorly preserved ancient and historic calculus samples prior to downstream compositional analysis.

During testing, we observed that a single extraction blank, EXB015.A2801, exceeded the oral fraction threshold value for a few of the most abundant taxa, despite the rest of the distribution matching other laboratory controls. This is likely due to low level laboratory cross-contamination in low complexity libraries, combined with the artefact effect of the higher ‘weighting’ of more abundant taxa - as the low abundance rank number leads to a small denominator and a large fraction value. Further, we observed that while many ancient/historic calculus and modern calculus samples also had the majority of the most abundant taxa having > 50% (for nt) or > 65% (for RefSeq) oral fractions, this pattern could fluctuate rapidly. There are three factors that may contribute to this effect. First, some taxa in the isolation source database, despite being true endogenous oral taxa, may be listed as ‘unknown’ in the database if the source information is incompletely reported in its metadata (missing metadata). Second, for oral taxa that are not in the database, reads from that species may exhibit a best hit to a closely related non-oral species in the database (incomplete database). And third, some samples obtained from museum collections have an overgrowth of a few environmental taxa from the storage environment that rank among the most abundant taxa, even though the remaining taxa detected in the sample originate from the original oral microbiome (overgrowth).

To account for these artefacts, we calculated a ‘burn-in’ threshold (External Data Repository Section 8.1 and Fig. R8B), then evaluated the samples. This ‘burn-in’ reduces the influence of taxonomic fluctuations from overgrowth, minor contamination, and missing database information, by focusing sample evaluation on the section of the curve where the oral fraction distribution is more stable - i.e., when the changing denominator has less of an effect. The burn-in mechanism selected was as follows, per sample:

1. In descending order, measure the difference in oral fraction between a given abundance rank and the next abundance rank.
2. Calculate the standard deviation from the mean of differences of all abundance ranks.
3. Identify the abundance rank position and onwards down the rank, the fluctuation in oral fraction differences between a position and next, does not exceed above or below the overall standard deviation of these differences.

After establishing this parameter, we then discarded any sample that did not exceed a 50% (for the nt analyses) or 65% (for the RefSeq analyses) threshold at any point after the abundance rank position identified (Fig. S3A, External Data Repository Section 8.1 and Fig. 9). Samples that had decay profiles similar to skin or sediment (and had high proportions of these two comparative sources according to SourceTracker - see below) - particularly in the Neanderthal group - were also indicated by the ‘burn-in’ threshold to be less preserved, suggesting this method performs well. Implementation of this procedure is provided as an R notebook referred to in the External Data Repository Section R8.1.

Of the 124 total individuals included in this study, 89 (71 ancient/historic, 18 present-day) remained after cumulative percent decay analysis against the NCBI nt database, and 77 (59 ancient/historic, 18 present-day) when compared against the custom RefSeq database. The percentage of samples per group that passed the filters is described in the External Data Repository under Section R8.1 and Table R1.

Comparing the two databases, there was a high degree of concordance in the estimated preservation status, although more samples were retained when using the nt database. The largest number of changes were observed for the PreagriculturalHuman_2 group (European preagricultural individuals) with 3 individuals (ECO002, ECO010, PLV001) being retained by the nt database and rejected by the custom RefSeq. Neanderthals have the fewest samples passing the threshold in both cases. A complete list of individuals remaining after these filtering steps is provided and described in the External Data Repository under Section R8.1 and Data R17. Calculus samples that did not pass the threshold set for each database, respectively, were subsequently removed for downstream compositional analysis.

S3.4.2 SourceTracker

As an additional method of preservation assessment, we used SourceTracker (v1) (223) to estimate the proportion of endogenous oral microbiome in our samples (203) and controls, through comparison to our comparative sources. Because SourceTracker was primarily designed for 16S rRNA amplicon data, we extracted 16S rRNA reads from our shotgun dataset of samples, laboratory controls and comparative sources and performed SourceTracker on this data subset, as has been performed in previous studies (2, 165). We utilise the older but less precise percent sequence similarity approach for OTU clustering (despite more recent higher-accuracy ASV techniques (224)), as we only are interested in an approximate estimation of preservation and ultra-short shotgun reads (rather than amplicon) may not be suitable for ASV approaches e.g. with DADA2 (225). Furthermore, it has been previously found that shallow-shotgun aDNA data cannot produce sufficient coverage for the DADA2 error correction methods (204).

To extract 16S rRNA reads, we mapped the shallow sequenced data against the SILVA SSU Ref Nr 99 trunc (v128) (226) database (with uracil replaced with thymine) using BWA, and then converted to QIIME (v1.9.1) (227) compatible FASTA files. The procedure is described in External Data Repository Section R8.2.1 and the number of mapped 16S rRNA reads is provided in External Data Repository File R12.

Across all samples, laboratory controls and comparative sources, the percentage of 16S rRNA mapped reads of the total processed non-human reads ranged from 0% to 0.5%, with a mean of 0.15%. Within calculus samples, each group had a mean and standard deviation as follows: *Alouatta* 0.21% ± 0.03; *Gorilla* 0.21% ± 0.08; *Pan* 0.12% ± 0.02; *Homo* (Neanderthal) 0.12% ± 0.07; *Homo* (Modern Human) 0.15% ± 0.05. (External Data Repository Section R8.2.1. and Fig. R10). These values are close to the predicted value of 0.2% for a microbial community, assuming an average 16S rRNA gene of length 1500 bp present in four copies in a 3 Mbp bacterial genome (228). Note that the percentages for Neanderthals do not include the individuals from Weyrich et al. (1) because these data were preprocessed in a different way prior to data upload.

Present-day human calculus yielded a greater fraction of 16S rRNA mapping reads (0.16% ± 0.04) than ancient/historic calculus (pre-agricultural humans: 0.12% ± 0.02; pre-antibiotic humans: 0.16 ± 0.07) (External Data Repository Section R8.2.1. and Fig. R11). This reduced fraction in the pre-agricultural ancient samples could either be due to age or taphonomic-related fragmentation leading to shorter reads that result in nonspecific mapping, or a reduced endogenous microbial biomass relative to present-day eukaryotic contamination. The latter will mean 16S rRNA reads are less likely to be ‘sampled’ from the library during sequencing.

QIIME closed-reference OTU clustering against the GreenGenes (v13.8) (229) database was then performed at 97% identity with the settings as shown in External Data Repository Section R8.2.2 Table R3. All other parameters were left at default. The procedure for closed-reference clustering is provided in the External Data Repository Section R8.2.2.

The number of identified OTUs across all samples, laboratory controls and samples ranged from 2 to 214,624 with a mean of 28,292. The lowest number of OTUs identified in a calculus sample was DLV002.A0101 with 27, which is consistent with its generally poor preservation indicated by both the cumulative frequency decay and Sourcetracker analyses. The mean and standard deviation of identified OTUs for each calculus group was as follows: *Alouatta* 16,244 ± 4,160; *Gorilla* 13,954 ± 11,885*; Pan*

8,877 ± 4,626; *Homo* (Neanderthal) 15,482 ± 20,200; *Homo* (Modern Human) 42,860 ± 52,705 (External Data Repository Section R8.2.2 and Fig. R12). A summary table of closed-reference clustering can be seen in External Data Repository File R13.

Present-day modern human calculus and plaque recovered a greater number of OTUs as represented in the GreenGenes database (External Data Repository Section R8.2.2 and Fig. R13), whereas ancient samples (pre-agricultural and pre-antibiotic modern humans) have a mean and standard deviation of 11,331.5 ± 7,524.7 and 14,254 ± 8,652.0 identified OTUs, respectively, whereas calculus from present-day individuals yielded 100,140.61 ± 53,809.9 OTUs and present-day plaque 50,9082.6 ± 26,393.7 OTUs.

After clustering, we removed any sample that did not have at least 1000 detected OTUs. This filter removed one chimpanzee (KNP004), one Neanderthal (BAN001), two preagricultural modern humans (DLV001, DLV002), and all but one library control blank (LIB0015.A0101). Next, the summarize_taxa.py QIIME utility script was used to subset the table to include OTUs at only genus level.

We do not use the 16S rRNA read data downstream for other analysis due to the low yield of these reads in many samples (External Data Repository Section R8.2.1 and Fig. R12), likely from the fragmented nature of our reads. Additionally, it is known that QIIME 1 assigns a very high number of low-abundance false-positive taxa, and is not reliable for fine scale taxonomic profiling of ancient metagenomes (204). This data therefore does not provide enough resolution for compositional analysis. However, despite the low yield and known biases, it is useful for estimating the approximate contribution of the oral microbiome in each sample using SourceTracker.

We provide the final OTU table from the clustering in biom format (230) and the post-rarefaction (see below) and cleaned OTU table in .tsv format in External Data Repository Section R8.2.2 and File R14. Code for visualisation of summary data can be seen in External Data Repository Section R8.2.2.

We ran SourceTracker (v1.0.1) (223) on the 16S rRNA QIIME OTU tables to estimate the proportion of the calculus from each individual that resembled supra- and subgingival plaque, urban and rural gut, sediment, skin and bone ‘environmental proxy controls’ (in the form of the Bronze Age Mongolian femurs). These comparative sources (indicated by SRR*, ERR* and ARS sample prefixes, Data S1 - Literature Dataset Tab) were selected based on the following criteria:

1. Data is from Illumina shotgun sequencing;
2. Data is from ‘unmodified’ samples (e.g. no treatment or disease);
3. Samples are from independent individuals (in the case of modern human samples);
4. A sample has enough data to find more than 1000 OTUs based on reads mapping to a 16S rRNA database.

SourceTracker analysis was then performed on the QIIME OTU table, which was rarefied to 1000 OTUs per individual, using calculus samples from this study and Weyrich et al. (1) as sinks and using SRR, ERR, and ARS samples as sources. While rarefaction has been described as ‘inadmissible’ (231) for most ecological analyses, it is recommended for use with Sourcetracker by its developers. For other analyses in our study we do not use rarefaction but rather apply more appropriate data normalization approaches (see SI Appendix section S4).

For analysis, we interpreted estimated contributions from any modern human gastrointestinal microbiome (i.e., dental plaque and gut) as indicative of endogenous oral microbiome preservation. This is because it has been previously shown that there is frequent misassignment between related oral and gut taxa using QIIME-based approaches, even when analyzing present-day oral samples (2, 204).

Subsequently, we visualised the results using R (see External Data Repository Section 8.2.3 and Fig. R14). The sink prediction results from Sourcetracker are provided in External Data Repository File R18. No clear threshold for a preservation cut-off was identified via visual inspection of per-calculus source proportions (Fig. S3B).

S3.4.3 Comparison of methods

Figure S3b shows a comparison between the estimated modern human gastrointestinal (oral, gut) source contributions across all calculus samples with more than 1000 OTUs based on 16S rRNA clustering and samples that passed the shotgun data-based cumulative decay threshold (Fig. S3a). The two methods show a general concordance in preservation assessment between both shotgun and 16S rRNA read based approaches, with most samples that were estimated to have low modern human gastrointestinal (oral, gut) contributions also not passing the cumulative decay threshold. In some cases, certain individuals such as MTS002 and MTM011 were unable to pass the threshold despite having a reasonable proportion of 16S rRNA reads resembling the expected endogenous content of calculus. These samples exhibited long tails of very low abundant OTUs (with very small fraction oral changes), which caused the standard deviation of fluctuation from the mean fraction to be very narrow. Subsequently, the rank where the fraction threshold fell within the standard deviation was pushed far along into the tail of very low abundant OTUs, where the tail was consistently under the 50% threshold for the NCBI nt database. These samples were discarded from our analysis. Other samples, such as MTM010 and TAF017 were retained via the cumulative decay method despite having comparatively low plaque and gut content. These were retained due to tails having minimal decay along the rank, containing many low-abundant oral taxa intermixed with non-oral taxa, rather than most of oral taxa being at the most abundant level. We retained these two samples for analysis because they do preserve oral taxa, but we acknowledge that they are borderline cases. Both cases show that further development of the cumulative percent decay method is required. Nevertheless, the two methods generally show a high degree of concordance, and the cumulative decay threshold approach has the advantage of utilizing a larger amount of available data, with fewer biases introduced by the QIIME taxonomic classification step. Overall, the cumulative decay threshold approach presents a straightforward, quantitative method for the rapid detection and removal of ancient or historic metagenomic samples with insufficient preservation for downstream compositional analysis. Code for comparison of cumulative percent decay and SourceTracker results can be located in External Data Repository Section R8.2.3.

S3.4.4 Abundance of eukaryotic content as a preservation indicator?

As noted above in section S3.3, the ratio of bacterial/archaeal/virus alignments to eukaryotic alignments was higher in present-day calculus than in ancient and historic calculus. We considered the possibility that this may be an indicator of microbiome preservation, assuming that the vast majority of DNA within calculus should originate from microbial and viral members the oral biofilm, whereas postmortem contaminants are more likely to include fungal and present-day human DNA (human DNA may persist even after filtering because of imperfect removal when using a consensus modern human reference genome). We tested this hypothesis by comparing the bacterial/archaeal/virus to eukaryotic ratios of ancient and historic calculus samples that passed the cumulative decay preservation threshold (n=71) versus those that did not (n=35). A one-way unpaired wilcoxon sum test was performed using the ‘wilcox.test’ function in R. The null hypothesis of the average ratio of ‘pass: false’ being the same average ratio as ‘pass: true’ by change was rejected, with the ‘pass: false’ average being less than ‘pass: mean’ (p = < 0.001; alpha = 0.05; U = 2175). While this metric was not used to reject samples because of the the level of overlap between the two distributions, a low bacterial/archaeal/viral to eukaryotic alignment ratio could be an additional ‘initial’ indicator of poor preservation across a whole ancient microbiome dataset before carrying out the more detailed procedures described above (External Data Repository Section R8.3 and Fig. R15). Code for statistical testing and visualisation of the eukaryotic content can be located within the External Data Repository Section R8.3.

S3.5 Ancient DNA authentication

Having removed low quality samples, we retained a set of 71 ancient/historic and 18 modern dental calculus samples whose microbial profile resembled an oral microbiome based on the MALT NCBI nt database screening. For the ancient and historic samples, we next sought to validate the authenticity of the oral microbial sequences to determine whether they exhibit appropriate molecular behavior for ancient DNA (203). To do this, we ran MaltExtract (v1.5) (205) on the shallow sequenced dataset (i.e., DNA sequences without UDG treatment) to generate statistics on well-known ancient DNA characteristics such

as C to T miscoding lesions at ends of fragments, and length distributions indicating highly fragmented DNA (217, 232). To generate these statistics, we used reference genomes selected from all oral genera determined to be 'core' across the primate groups in this study (*Alouatta:Gorilla:Pan:Homo*; see section S5 below).

To facilitate the visualization and examination of these patterns in the output of MaltExtract, we created the interactive R shiny (233) application 'MEx-IPA' (**M**alt**Ex**tract - **I**nteractive **P**lotting **A**pplication). All well-preserved samples in our analysis set exhibited indications of authentic ancient DNA across multiple core taxa. For example, Neanderthal calculus from Pešturina Cave (PES001, ~100 ka) and Grotta de nadale (GDN001, ~70 ka) exhibited frequencies of C to T mismatches at the end of reads in excess of 40% and short fragment lengths (45-55 bp) across a range of species level nodes of oral taxa (e.g., *Fretibacterium fastidiosum*, *Fusobacterium nucleatum, Tannerella forsythia*, *Treponema denticola* ATCC 35405, as shown in Fig. S4 and External Data Repository Section R8.4.2 and Fig. R16). MEx-IPA data for all samples can be seen on the External Data Repository Section R8.4.2.

To further validate these patterns using a more formalised model of ancient DNA damage against specific reference genomes, we then repeated this analysis using the tool DamageProfiler (234). For this analysis, we used the EAGER pipeline (see settings in section S3.2) to map the non-human reads of the entire shallow sequencing dataset to a selected set of reference genomes representative of core oral taxa (see section S5 below), including both Gram-positive (*Pseudopropionibacterium*, *Streptococcus*) and Gram-negative (*Tannerella*, *Treponema*) members. We then plotted the damage patterns generated by DamageProfiler in R using ggplot2, and inspected them manually to confirm the presence of DNA damage. All ancient and historic samples in our analysis set exhibited characteristic DNA damage patterns across these oral taxa, while present-day clinical samples did not. Representative examples of DNA damage patterns observed for Neanderthal (PES001), Upper Palaeolithic modern human (EMN001), and present-day modern human (JAE006) dental calculus are shown in Fig. 1C and External Data Repository Section 8.4.2 and Fig. R17. Source-code for MEx-IPA can be downloaded from the External Data Repository (https://github.com/jfy133/MEx-IPA, doi: 10.5281/zenodo.3380012). Visualisation of the DamageProfiler output was performed using an R script described in the External Data Repository Section R8.4.2.

S3.6 Contamination assessment and removal

Within our sample set (including both new and published data), we established that the dental calculus of 71 newly sequenced ancient and historical individuals show signs of being well-preserved, exhibiting both an oral microbiome-like taxonomic profile and having characteristic ancient DNA damage. However, even well-preserved ancient microbiome samples may contain trace amounts of contaminant environmental and laboratory DNA. While conservative approaches, such as removing from the samples all OTUs found in extraction and library controls have been previously used to remove potential contaminants from ancient microbiome datasets (1), this brute-force method is highly problematic and may lead to the removal of highly abundant endogenous taxa due to the presence of even a single misclassified DNA read in a control sample (235). Such extreme approaches to contaminant removal can substantially skew downstream datasets by removing entire oral clades, contributing to potentially erroneous results. Recently, more sophisticated and statistically robust methods for contaminant removal have been developed that consider the abundance and/or prevalence of potential contaminants, and we use these methods here.

To identify and remove potential laboratory and environmental contaminant OTUs from our calculus datasets, we used the R package 'decontam' (235) (https://benjjneb.github.io/decontam/vignettes/decontam_intro.html). To identify laboratory contamination, we used as proxies the extraction and library controls from this study. To identify environmental contaminant OTUs, we used as proxies ten ancient femur samples known to contain a high amount of exogenous, environmental DNA typical of archaeological skeletal remains; these samples (site code ARS; see Data S1 - Literature Dataset tab for full list) were analysed in a previous study (170), but were resequenced here to a higher depth to provide a suitable dataset for decontam processing (see section S2.4.2). We chose not to use sediment or soil samples for decontam processing because archaeological

remains contain a skewed subset of soil microbial communities (236), as not all microbial taxa colonise skeletal remains, thus present-day soil and sediment datasets are not suitable proxies for this analysis.

Due to the low biomass nature of our samples and following the recommendation of the tool creators (235), we applied the 'combined' method (frequency and abundance) to our dataset. For comparison, we performed decontam at both the species and genus taxonomic levels on both the 'nt' and 'RefSeq' MALT/MEGAN OTU tables, as well as on MetaPhlAn2 taxonomic tables that we later generated for HUMANn2 functional analysis (see section S7.1). By using the combined frequency and abundance method, we were able to account for the fact that the archaeological bone taxa would not follow the inverse abundance correlation of putative laboratory contaminants in extraction and library controls. To increase sensitivity of contaminant detection, we did not apply minimum support filtering to the MALT/MEGAN OTU tables prior to analysis. In addition, we included low-preservation samples in this analysis to improve sample size and statistical power. The probability threshold for contaminant identification (the threshold that a OTU was required to pass to reject the null hypothesis that the OTU was not a contaminant - see Davis et al. (235) for more detail) was modified here to take a conservative approach due to the expected higher level of contaminants given our degraded samples. We set this threshold by testing increasing values until well-known and common environmental contaminants such as *Pseudomonas* and *Streptomyces* were marked as contaminants, while well-known oral taxa such as *Tannerella* or *Streptococcus* were kept as 'not contaminants'. For both the nt and RefSeq MALT/MEGAN OTU tables, we set the combined method p-value to 0.99 for both the genus and species level analysis. For MetaPhlAn2 at the genus level, we also used a combined method p-value of 0.99 but relaxed this slightly to 0.9 at the species level. The final lists of potential contaminants are provided in the External Data Repository Section R8.5.1 and File R19.

In all cases more potential contaminants were identified at the species level than at genus level. The number of contaminants detected using our strict thresholds are provided in External Data Repository Section R8.5.1 and Table R4. For the MALT/MEGAN binning, approximately half of the OTUs detected in samples were considered possible contaminants at both the species and genus levels, and around one sixth of OTUs for MetaPhlan2. However, although the fraction of putative contaminant OTUs was relatively high, the actual number of reads contributing to these OTUs was very low for well-preserved samples (see External Data Repository Section R8.5.2 and Fig. R18), suggesting these contaminant OTUs are a part of the low-abundance tail that is often observed in metagenomic analyses (237). We used these OTU lists to remove likely contaminants from the MALT/MEGAN OTU tables for downstream analysis. The entire OTU 'decontamination' filtering procedure is described in the External Data Repository Section R8.5.

S4. Microbial compositional analysis

Having removed low quality samples and likely contaminants, we proceeded with compositional analysis of the 89 well-preserved (71 ancient/historic; 18 modern) dental calculus samples in this study.

As an initial step in comparing the calculus microbiome among the different African hominid groups in this study, we first tested the hypothesis that the oral microbial community of each primate genus was distinct at a compositional level. For this we performed two methods of clustering and statistical validation: (1) principal coordinate analysis with PERMANOVA, and (2) hierarchical clustering with cluster validation.

S4.1 Principal coordinates analysis (PCoA)

For distance matrix generation we performed an isometric-log-ratio transformation (ILR), to account for the compositional nature of our data (231, 238). The phyloseq (v1.26) (239) and PhILR packages (v1.8) (240) were used following the introductory tutorial of PhILR (https://bioconductor.org/packages/release/bioc/vignettes/philr/inst/doc/philr-intro.html), to normalise our data with additional consideration of microbial OTU phylogenetic similarities. OTU tables as generated by MALT/MEGAN were loaded in R and filtered for genus and species levels accordingly. To reduce noise, we removed OTUs falling below minimum-support (i.e. abundance) threshold values at the genus (0.07%) and species (0.04%) levels (see section S5.2 for threshold details). OTUs listed as potential laboratory

and environmental contaminants were also excluded, and the resulting OTU table was converted to a matrix. As a phyloseq object requires an associated taxonomy table to an OTU table, we used the 'get_uid' and 'classification' functions of the taxize (v0.9.6) package to get the NCBI taxonomy ID and taxonomic paths of each OTU in the table. The newick tree as exported by MEGAN was then loaded using the 'read.tree' function from the ape (v5.3) (241) package. Polytomies were removed using the 'multi2di' function, and internal nodes on single branches were removed using ape's 'collapse.singles' function (http://blog.phytools.org/2018/05/when-phylogeny-fails-isbinary-but-is.html). A phyloseq object was then created, and a pseudo-count of 1 was applied to remove 0s (log-ratio methods require positive values) using the 'transform_sample_counts' function. We acknowledge that others have suggested that pseudo-counts are not the most optimal zero-replacement solution (242); however comparisons using more sophisticated methods (e.g., the cumulative zero multiplication method as implemented in the 'cmultRepl' function of the zComposition package (243)) resulted in only minor topological differences while most relationships were maintained (External Data Repository Section R9.1 and Fig. R19). We elected to use the pseudo-count method for its simplicity and because it is the recommended method in the PhILR tutorial. Code for zero-replacement method comparison and visualisation is provided in the External Data Repository Section R9.1.

We performed the PhILR transform using the 'philr' function, with the part.weights parameter as 'enform.x.m.counts' and ilr.weights set to 'uniform'. The resulting transformed table was converted to an R dist object using the base R function 'dist' with method set as euclidean, and ordinated using Principal Coordinate Analysis (PCoA; also known as multidimensional scaling or MDS) by the 'ordinate' function in phyloseq.

The above procedure was applied to multiple sample sets. For example, External Data Repository Section R9.1 and Fig. R20 (panels A, C, E, and G) shows where highly decomposed calculus samples (identified for removal in section S3.4) ordinate compared with other samples and controls, while External Data Repository Section R9.1 and Fig. R20 (panels B,D,F,H) shows the same plots but with the highly decomposed samples removed. Overall, oral samples and comparative sources (ancient and present-day dental calculus and present-day dental plaque) cluster together, whereas non-oral comparative sources tend to fall away from this cluster, with distance correlating with the level of association to modern humans - i.e., non-oral human microbiome comparative sources such as gut and skin in some cases partially overlap with the plaque/calculus, whereas sediment and archaeological bone (as an environmental proxy control) fall further away. This is expected due to taxonomic similarities between microbial species across the modern human body, as well as in the case of skin from contamination during handling of ancient and historic samples. Samples detected as having low preservation generally fall in a space between calculus samples and other non-oral comparative sources, suggesting these low preservation samples contain a large fraction of non-oral metagenomic content. These patterns are consistently observed for both the NCBI nt and RefSeq MALT/MEGAN datasets.

Principal coordinate analysis of only well-preserved calculus samples indicates that samples tend to cluster by host genus, albeit with varying levels of overlap (External Data Repository Section R9.1 and Fig. R21). For both databases (nt and RefSeq), Principal Component (PC) 1 generally separates *Alouatta/Gorilla/Homo* from *Pan*, whereas PC 2 separates *Gorilla*/*Alouatta* from *Pan/Homo.* Neanderthal individuals generally fall between *Gorilla* and *Homo*. PC 3 separates differently for each database, with nt partially separating *Alouatta*/*Gorilla* from *Homo/Chimp,* whereas there is no clear separation of host genera with the custom RefSeq database. The full distance matrix, ordination procedure and code are provided in the External Data Repository Section R9.1.

S4.2 PERMANOVA

To determine whether the clustering of well-preserved calculus microbiomes by host genus is statistically significant, we performed a nonparametric ANOVA (PERMANOVA (244)) on the PhILR transformed euclidean distance MALT aligned matrix (omitting contaminants and low preservation samples), as implemented by the 'adonis' function in the R package vegan (v.2.5.4) (245). For this analysis we also removed *Alouatta* as it had a small sample size and originated from a single population. *Gorilla, Pan* and *Homo* had significantly different (alpha = 0.05; $p < 0.01$) centroids in euclidean space for both the nt and

RefSeq MALT/MEGAN datasets and at both the microbial genus and species taxonomic levels. A summary of these results is provided in External Data Repository Section R9.2 and Table R5.

It is known that PERMANOVA is sensitive to heterogeneity in the dispersion of samples from a cluster's core. While the method is robust against this when sample sizes are equal, this is not the case when there are different numbers of samples per group (246). To measure dispersion, we used the PERMDISP test (247) as implemented in vegan with 'betadisp'. The base R function 'anova' and vegan's 'permutest' were both used to check if there was a significant difference in the amount of dispersion between each group. In all cases, heterogeneity was shown to be borderline given an alpha value of 0.05 in both tests (see External Data Repository Section R9.2 and Table R6).

Because methods for correcting for beta-dispersion heterogeneity in unequal sample sizes (248) have not yet been implemented in R, we accounted for this by repeatedly subsampling (with replacement) each genus group to 10 individuals and running the adonis() function on these subsets 100 times. *Alouatta* was not included because this group only contained 5 individuals. External Data Repository Section R9.2 and Table R7 shows that the distribution of *pseudo-F* and *p* values across all runs confirm that *Gorilla*, *Pan* and *Homo* cluster around statistically distinct, host genus-specific centroids (alpha = 0.05; mean $p < 0.01$, mean pseudo-$F > 5$ across all database taxonomic level combinations). Larger sample sizes are required, as well as more preservational and storage history information to control for these additional variables, to provide enough power to test for population structure within each host genus, and therefore did not test this here. The beta dispersion and PERMANOVA procedure and code are provided in the External Data Repository Section R9.2.

S4.3 Hierarchical clustering

Although the calculus communities of each host genus were generally distinct, the PCoA visualisation showed large areas of overlap between groups along different combinations of principal coordinate axes. To further understand the relationships between each host genus, we performed hierarchical clustering to visually identify species that drive these overlaps. To prepare the dataset, we used the MALT/MEGAN OTU tables from both the NCBI nt and custom RefSeq databases (excluding low preservation samples and after removal of contaminants) and applied a low-abundance OTU tail filtering threshold of 0.04% for species, 0.07% for genus (see section S5.2 for threshold selection and taxon retention or removal). We then also applied a prevalence filter by removing any microbial taxon that was not present in at least 5 individuals across the whole dataset. To again account for the compositional nature of our data, we then applied a pseudo-count of 1 and a Centered-Log-Ratio (CLR) transform to the filtered OTU tables (following https://github.com/ggloor/CoDa_microbiome_tutorial/wiki/Part-1:-Exploratory-Compositional-PCA-biplot). Although CLR can cause some downstream issues with further statistical testing (which is why PhILR was used above) (249), ILR transformation causes loss of species specific information (i.e., single values for each species rather than balances between different species). We therefore chose to use CLR to retain this information, and we do not perform the type of downstream statistical testing that may be compromised by this approach. The resulting CLR-transformed OTU tables were converted to the euclidean distance using the 'dist' function in R, and a range of hierarchical clustering algorithms were applied to both the individuals and the microbial taxa, as offered by the 'hclust' function. To identify the optimal algorithm that best describes the original relationships in the distance matrix, the 'cophenetic' function from the vegan R package was used to inform which hierarchical clustering had the greatest correlation with the distance matrix. In all cases, either the unweighted ('average') or weighted ('mcquitty') versions of the UPGMA algorithm were selected. To estimate the uncertainty in the support of the clusters identified by the optimal algorithm, the pvclust R package (250) was used to perform bootstrapping on the hierarchical clustering with 1000 bootstraps. To extract the pvclust support values for display in ggtree, a modified version of the 'as.phylo.hclust.node.attributes' function from the fastbaps package (251) was used. Heatmaps for both NCBI nt and the custom RefSeq database are provided in the External Data Repository Section R9.3 and File R20. Clusters of microbial taxa within the heatmap were visually identified, and phenotypic metadata for all microbial taxa was collated from the BacDive database (252) (https://bacdive.dsmz.de/) via the BacDiveR R package (v0.9.0, https://doi.org/10.5281/zenodo.1308060). Of the 278 taxa in the NCBI nt database remaining after filtering, 148 had phenotypic data recorded in the BacDive database.

In general, individuals cluster within their host genus (Fig. 3; External Data Repository Section R9.3 and File R20). In both cases (nt and RefSeq), *Alouatta* and *Gorilla* cluster together as sister groups, and a subset of ancient modern humans fall with *Pan*. Neanderthals also consistently fall within the diversity of *Hom*o, slightly more often within the non-*Pan* associated *Homo* cluster. Present-day *Homo* fall in a distinct cluster either within the wider *Homo* cluster (nt), or as an 'outgroup' to all other clusters (RefSeq). The bootstrap support values for the described clusters range from 85% (*Pan*/*Pan*-associated *Homo* bifurcation) to 96% (Present-day *Homo*/Ancient *Homo* bifurcation) in nt, and 77% (*Pan*/*Pan*-associated *Homo* bifurcation) to 87% (*Pan* and *Pan* associated ancient *Homo*/non-*Pan associated* ancient *Homo bifurcation*) in the custom RefSeq heatmap (Fig. 3; External Data Repository Section R9.3 and File R20).

A collection of *Actinomyces, Streptococcus, Pseudopropionibacterium, Fusobacterium, Ottowia and Campylobacter* species are found present across all individuals. The three strongest signals across all host genera are *Pseudopropionicum propionicum*, *Actinomyces* sp. oral taxon 414, and *Ottowia* sp. oral taxon 807 (External Data Repository Section R9.3 and File R20). It is remarkable that the three most prevalent species identified in our dataset - both in present-day and ancient/historic individuals - are virtually uncharacterised and two out of three are unnamed. Further investigation is warranted to understand the role of these taxa in the oral biofilm and why they are prevalent: for example, whether this is an artifact from cross-mapping or misidentification; whether they possess phenotypic characteristics that result in unusually good preservation; or whether they play a special role in the hominid oral biofilm.

Within the heatmap, *Gorilla* and *Alouatta* share blocks of aerobic and microaerophilic taxa, and in particular they harbour a large diversity of streptococci (External Data Repository Section R9.3 and File R20). Part of the *Streptococcus* diversity in *Gorilla* and *Alouatta* is also shared with most *Homo* individuals, alongside other aerobic and microaerophilic taxa such as *Neisseria*, *Haemophilus*, *Rothia, Aggregatibacter,* and *Eikenella*. Less prevalent among *Gorilla* and *Alouatta* are taxa typically considered to be 'late colonisers', such as the anaerobic bacteria *Fretibacterium*, *Tannerella*, *Treponema* and *Porphyromonas.*

Pan has a notably low abundance and diversity of *Streptococcus* compared to other African hominids, and anaerobic bacteria are the major component of *Pan* individuals, as well as a subset of ancient modern humans. *Pan* harbours the four major late colonisers not found in *Gorilla*, as well as a wide diversity of other anaerobic species within the genera *Slackia*, *Prevotella*, *Olsenella* and *Eggerthella* (External Data Repository Section R9.3 and File R20). The '*Pan*-like' modern human individuals originate from diverse locations (Europe, as well as northern and southern Africa) and time periods (both pre-agricultural and pre-antibiotic periods), suggesting that this pattern is not related to a specific geographic or temporal origin (External Data Repository Section R9.3 and File R20). Furthermore, these samples are from both adult male and female individuals, and originate from a variety of tooth types. Although lacking the wide breadth of anaerobic diversity seen in *Pan,* these modern human individuals do share other features with *Pan*, such as having low to absent levels of early coloniser microaerophilic *Streptococcus* taxa (such as *S. gordonii*, *S. sanguinis* and *S. cristatus*), which are otherwise prevalent across *Gorilla*, *Alouatta* and most *Homo*.

Neanderthals present a very similar range of taxa to modern humans. This highlights that compositionally, there appear to be few microbial differences among members of *Homo*, albeit based on the contents of the databases currently available that are biased towards Western-industrialised populations. *Homo* calculus in our dataset in general contains a range of taxa consistent with current knowledge of the modern human dental plaque microbiome, including *Streptococcus, Actinomyces, Fusobacterium, Neisseria, Haemophilus, Campylobacter, Capnocytophaga, Prevotella, Tannerella, Treponema,* and *Porphyromonas (253)* (External Data Repository Section R9.3 and File R20). One species of particular interest is *Fusobacterium nucleatum*. It is able to coaggregate with a wide diversity of other taxa, making it a 'bridging' taxon between early and late stages of oral biofilm formation (254). We find this taxon to be ubiquitous across all host genera (slightly less with the custom RefSeq database), with a similar relative abundance in each individual. Following our filtering criteria, we note that - at the species level - we do not find members of *Corynebacterium*, a genus recently shown to play an important role in acting as a 'structural pylon' between early and late colonisers (255). However, we do detect *Corynebacterium* at the

genus level (see section S5), which suggests that we have poor species-level resolution and thus likely insufficient database representation for this genus. These patterns are found for both the NCBI nt and custom RefSeq databases.

In the late 1990s, periodontologist Sigmund Socransky classified dental plaque bacteria into a series of complexes, each with different health and disease associations (256). Of these, the 'red complex', which consists of *Porphyromonas gingivalis*, *Tannerella forsythia,* and *Treponema denticola*, was considered to be associated with periodontal disease and the species within it remain targets of clinical interest. Across the analysed African hominids and howler monkey outgroup, we found that most host genera harbour at least some alignments to these species (External Data Repository Section R9.3 and File R20). However, red complex taxa were less prevalent in *Alouatta* and *Gorilla*, and, while present in all individuals in *Pan*, they were less abundant than in *Homo*. Contrary to previous studies, it is now known that the presence of these taxa in calculus does not necessarily indicate severe disease (2, 23), even in present-day modern human populations, and we are not able to comment on the periodontal disease status of many of the individuals in this study because this information was not available or not recorded for a large number of individuals.

Although the genus-level assignments across African hominids appear to be robust, it is important to note that for species-level identifications, some of these alignments may represent mis-mappings between closely related species of the same genus that are absent from the databases. As the diversity of sequenced genomes in public databases increases, it is clear that past alignments and identifications may have been biased towards Western and clinically 'relevant' taxa (211, 257, 258). In these cases, reads from uncharacterised commensal relatives have been mis-assigned to closely related (often pathogenic) species present in the database, resulting in false positives (see for example the case of *P. propionicum* and *P. acnes (203)*). Each of the 'red complex' species belong to genera containing multiple other oral species, and it is likely that the species diversity of these genera has not yet been fully characterised. Indeed, we note that *Tannerella* sp. oral taxon HOT-286, an unnamed health-associated species (259) not included in all databases, has a similar prevalence to *Tannerella forsythia* across all samples in our analyses, suggesting that care should be taken by archaeogeneticists before reporting on oral 'pathogens'. Further evidence of this effect is discussed in section S6.3, where we show that many single genome alignments of our dataset resulted in high numbers of multi-allelic SNP calls for these haploid bacterial references, indicating the presence of multiple strains and - potentially - additional uncharacterised species.

The hierarchical clustering procedure and code is provided in the External Data Repository Section R9.3. The location of the script used to collate phenotypic metadata ('bacdive_searcher') is also provided in External Data Repository Section R9.3.

S4.4 Indicator analysis

To further investigate the characteristics of the host specific clusters defined above, we performed Indicator analysis (260) as implemented in the indicspecies R package (261) on the dataset described in section S4.3 using default settings and the nt database. Indicator analysis considers the mean relative abundance of a taxon versus the abundance across all other groups, and considers the frequency (prevalence) of the species in the given group to identify taxon that can be a 'signal' species of a given group. The indicspecies package performs this procedure on each group and on all possible combinations of groups, and generates a significance value by randomly reassigning individuals to other groups to assess the uncertainty of taxon assignment as an indicator of that group(-combination). We again used the BacDiveR package to collate phenotypic characteristics for the taxa. Whether or not a taxon was 'oral related' was determined by searching for any sample type that contained the following keywords: oral, dental, caries, plaque, perio, mouth, dentine, gingiva, and saliva. False keywords, such as "coral" or "floral", were excluded.

In this analysis, 140 taxa were identified as statistically significant indicator species (alpha ≤ 0.05). *Alouatta* and *Pan* each had 23 distinct species that were considered to be indicators for each group,

whereas *Gorilla* had 6 and *Homo* had 7; the remainder indicated various host combinations. Of the total 140 indicator species, 100 had phenotypic data in the BacDive database. Aerobes and microaerophiles made up the largest fraction of indicator species for *Alouatta* and *Gorilla*, as well as combinations between *Alouatta* and *Gorilla* with *Homo*. In contrast, all indicator taxa for *Pan* were anaerobes, as were the majority of all combinations involving *Pan*. Few overall patterns were observed for *Homo*, but the seven *Homo*-specific indicator taxa included three species of *Capnocytophaga* (a group of bacterial anaerobes with gliding motility), three species of *Methanobrevibacter* (a group of archaeal methanogens), and *Selenomonas sputigena* (a motile bacterium that coaggregates with *Actinomyces* and *Fusobacterium*). All three groups have been associated with subgingival plaque and periodontal disease in modern humans (262, 263), and of these, members of the genus *Methanobrevibacter* were indicators of *Homo* exclusively, suggesting that the modern human oral biofilm provides a particularly favorable environment for the growth of these methanogens, which only grow in a highly reduced and anaerobic (H_2 and CO_2 rich) local environment, such as that found in mature subgingival plaque biofilms (264). The list of all taxa identified by indicator analysis and associated metadata is provided in External Data Repository Section R9.3.2 and File R21. The code used to perform indicator analysis is also provided in the External Data Repository Section R9.3.2.

S4.5 Are dietary practices associated with oral microbiome composition?

Adler et al. (265) and Weyrich et al. (1) have previously reported phylum-level ancient oral microbiome structure correlating to broad dietary (subsistence) patterns in modern humans. However, in each study sample sizes were small and population factors were not controlled. Specifically, Weyrich et al. (1) defined four oral microbiome clusters using UPGMA hierarchical clustering of Bray-Curtis distances of shotgun data: (1) Forager-gatherers (Chimpanzees, African pre-pastoralist individuals and two the Neanderthals from El Sidrón), (2) Ancient agriculturalists (Individuals from the Industrial revolution period, Medieval England and Germany, and Neolithic Linear Band Keramik [LBK] individuals from Germany), (3) Hunter-gatherers (Polish Mesolithic individuals, the well-preserved Neanderthal from Spy, African pastoralist period individuals and an individual from the LBK), and (4) a single present-day individual. However, there were outliers in these clusters that did not match the expected dietary practices of the individual's time period, such as an LBK individual in the hunter-gatherer group (1).

Using our larger dataset, improved preservation assessment tools, more appropriate analytical methods for compositional data, and more balanced sampling strategy, we attempted to replicate these results. We divided individuals from our data into four groups that primarily differ by temporal period, but also by broad dietary practices: (1) Neanderthals, (2) Pre-agricultural modern humans (hunting and gathering), (3) Pre-antibiotic modern humans (traditional agriculture and pastoralism), and (4) Present-day individuals (industrialised diets). To perform the test, we replicated our principal coordinate analysis above at genus level, but only on individuals from the members of *Homo*. We found large overlaps between all four groups (Fig. S5E-F and External Data Repository Section R9.4 and Fig. R22). No variation in beta-dispersion was observed, and PERMANOVA as performed by the 'adonis' function showed a significant difference in centroids of each group ($\alpha = 0.05$, $p = 0.001$, pseudo-$F = 4.98$, $R^2 = 0.23$). However, pairwise adonis functions found a significant difference only between the present-day individuals and Pre-agricultural and Pre-antibiotic era modern humans; all other group comparisons were not significantly different.

To more closely approximate the hierarchical clustering approach used Weyrich et al. (1), we applied the UPGMA (average-linkage) hierarchical clustering method using the hclust() R function on the euclidean distances of the PhILR transformed OTU table (again accounting for the compositional nature of our data). This also did not result in any clear clustering by group (Fig. S6, External Data Repository Section R9.4, and Fig. R23). Indeed, while individuals from each time period and region (Europe or Africa) appear to somewhat group together, clusters were always interspersed with individuals from other temporal or geographic areas. Programmatically dividing the outcome of the hierarchical clustering into four clusters (corresponding to our four groups, performed by the 'cutree' R function), and comparing to our defined temporal/dietary clusters using the adjustedRand and Jaccard metrics, again showed low correspondence between the 'expected' and actual clustering (adjusted rand index (HA) = 0.017, Jaccard = 0.29). The two values being closer to 0 than 1 suggest random clustering rather than following the

expected grouping. Thus, we were unable to replicate the previously described pattern of oral microbiome clustering by dietary practice, as reported by Weyrich et al. (1), at least at the present resolution of the data.

Although we observe no large-scale compositional differences among modern humans based on dietary practices, there may be more subtle differences that fall below the limits of our detection, perhaps driven by low abundance taxa. However, due to variable levels of environmental contamination in ancient DNA datasets, it is currently difficult to separate true and false endogenous low abundance taxa. Although we do not follow exactly the approach described by Weyrich et al. (1), we believe our analytical approach is robust, having a larger sample size, using more balanced population groupings, and employing distance calculation methods more suited for compositional data (ILR transform). In addition, we performed our analysis at the genus level rather than at the phylum level, which should be more sensitive to compositional differences, and we believe it unlikely that a phylum level analysis would provide any further insights as it would only represent a data reduction.

One problematic aspect of previous ancient DNA reports of oral microbiome structuring by diet (1, 265) is that these studies did not have a systematic approach to managing environmental contaminants or other artifacts related to preservation. As such, the clustering they observed could have been driven more by preservational factors rather than diet. It is noteworthy that studies of present-day oral microbiome variation to date have found little evidence of systematic compositional differences among groups with differing diets (266–268) or related to the intake of specific food and drink (269–272). On the basis of our evidence, we do not find support for higher taxonomic-level structuring of the oral microbiome by diet within *Homo*. Rather, we believe such structure, if present, should be sought at the level of species or strains, a level that currently lies at the very limits of detection for ancient microbiome studies using available technologies. The locations of the scripts and procedures for generating PCoA, PERMANOVA and hierarchical clustering by subsistence strategy analysis are provided in the External Data Repository Sections R9.4.

S5. Core microbiome

As shown by PERMANOVA (section S4.2), the calculus oral microbiomes of African hominids (*Gorilla*, *Pan*, *Homo*) are generally distinguishable by taxonomic composition. However both PCoA (section S4.1) and hierarchical clustering (section S4.3) demonstrate that large taxonomic overlaps between each group are observed when considering the most abundant taxa in each of the samples. Furthermore, it was noted that despite the fact that the reference genomes used for taxonomic assignment have a strong bias towards those found in present-day modern humans from Western, industrialised nations, the cumulative frequency decay plots and Sourcetracker plots produced during preservational assessment (Fig. S3) were still able to clearly detect a distinctively oral signal even in evolutionary distant host species, such as *Alouatta*. Together, this suggests that there is a level of similarity in the compositional makeup of the oral microbiome of each host genus. We therefore decided to explore these shared microbial features and their evolutionary time depth by performing a core microbiome analysis of the host genera in this study. But first, we review the current state of knowledge regarding oral biofilm spatial organization and development in modern humans.

S5.1 Biofilm spatial organization and development: current state of knowledge

Dental plaque is a highly structured and spatially organised (255, 273–278) biofilm composed predominantly of bacteria (149) but it also contains viruses (279), archaea (149), and eukaryotes (280). Early studies on dental plaque development both *in vitro* and *in vivo* revealed that there is a temporal progression of the species present (with species being classified as “early” and “late” colonisers (256, 281) and that plaque structure is related to microbial composition (273). Much of the work on biofilm plaque development was performed before DNA sequencing was commonly available, and biochemical traits were used to identify species. This, along with updated nomenclature and taxonomic classification, as well as the description of novel species in the intervening years, has resulted in some inconsistencies in our knowledge of the species that form the early biofilm.

In brief, modern human dental plaque biofilm formation is initiated by the adsorption of salivary proteins and glycoproteins to the enamel surface of teeth, forming the 'acquired enamel pellicle' (AEP) (282). Oral bacteria cannot bind directly to the hydroxyapatite mineral of enamel, but a subset of bacteria - notably some members of *Actinomyces* and *Streptococcus* - can bind to proteins in the AEP, and thereby initiate dental colonization. Proteins and glycoproteins making up the AEP include salivary amylase, proline-rich proteins (PRPs), statherin, histatins, mucins (MUC5B, MUC7), cystatins, lysozyme, and lactoferrin, among others (283). Two particular saliva-derived components of the enamel pellicle, PRPs and salivary α-amylase, may be especially important in the formation of the oral biofilm because of their specific interactions with oral bacteria. Proline-rich proteins (PRPs) are a family of salivary proteins that are constitutively expressed in humans and make up the majority of salivary proteins (284). Plant tannins, which can cause toxicity in mammals, have a high affinity for PRPs (285), and ingestion of dietary tannins induces expression of PRPs in rodent models (284), although PRPs may be continuously expressed in primates (8). Proline-rich proteins appear to protect against dietary tannins by binding and precipitating them, thereby neutralizing their effects. Proline-rich proteins also bind to hydroxyapatite, the main mineral in both tooth enamel and dental calculus. Binding to a hydroxyapatite surface can induce conformational changes in PRPs that reveal novel epitopes, to which oral bacteria bind (286). Such epitopes, called cryptitopes, are masked when PRPs are in solution but are exposed on binding. PRPs are major salivary proteins in primates, with the exception of gelada baboons, which have low tannin, grass-based diets and do not express PRPs in saliva (287). Salivary α-amylase is the most abundant enzyme in saliva (288), and it breaks down dietary starch into maltose and dextrins. Salivary amylase complexes with the abundant salivary mucin MUC-5B (formerly MG1), which is strongly attracted to hydroxyapatite, and may make up a large proportion of the enamel pellicle depending on the location in the mouth (288), but not of the mucosal pellicle (289). Salivary α-amylase is distributed throughout the pellicle with no obvious pattern of localization (290). The enzyme remains active after adsorption (288), and although salivary α-amylase spontaneously desorbes from the tooth surface, constant amylase activity can be detected at the pellicle surface (Hannig 2005). Salivary α-amylase expression in primates varies substantially between species (Pajic et al. 2019), and dietary tannins inhibit salivary α-amylase activity (288). Members of both *Actinomyces* and *Streptococcus* are capable of binding to PRPs in the AEP, while only members of *Streptococcus* are known to bind to salivary α-amylase (291, 292). For an extended discussion of the special relationship between *Streptococcus* and salivary α-amylase, see section S5.6.1.

Plaque development can be classified into distinct stages based on the timing and characteristics of the organisms present (281, 293, 294). Early biofilms are dominated by Gram-positive cocci for 1-2 days, followed by a proliferation of fusiforms and filaments, and finally a proliferation of Gram-negative species around 5-7 days. Temporal succession from predominantly Gram-positive species in the early plaque biofilm to predominantly Gram-negative species in mature plaque biofilms is well-reported in both microscopic and culture-based studies of plaque development. Beneath a surface layer of aerotolerant taxa, the bulk of the mature dental plaque contains predominantly Gram-negative anaerobic species, while the deepest layers of plaque that develop earlier contain mainly Gram-positive facultative anaerobes. This transition is readily seen in Gram-stained modern and ancient dental calculus sections, in which clearly defined calcified layers each repeat a pattern of a layer of Gram-positive aerotolerant biofilm initiators followed by an overlying layer of Gram-negative, largely anaerobic cells that proliferate within the growing biofilm (221, 277).

Early biofilms are dominated by members of *Streptococcus* and *Actinomyces*, which is related both to their ability to bind to the AEP and to their metabolism as carbohydrate-fermenting facultative anaerobes (295), but other organisms including *Neisseria* and *Nocardia* are also present (296–300). In addition to being able to bind to components of the AEP, both *Streptococcus* and *Actinomyces* are also able to coaggregate with each other (286, 301, 302), which is important for initiating biofilm formation on the tooth surface.

Proliferation of the species *Fusobacterium nucleatum* in the biofilm allows the transition to a mature plaque profile. *Fusobacterium nucleatum* acts as a "bridging organism", meaning that it is able to bind to many other biofilm species, both early colonisers and late colonisers, that often cannot otherwise interact with each other (302). Interestingly, this function is disrupted in the presence of lactose (303–306), which may explain why young children rarely develop periodontitis (307) and do not generally show stable

colonization by some late coloniser bacteria, such as *Aggregatibacter actinomycetemcomitans* and *Porphyromonas gingivalis*, until late childhood or adolescence (308).

The growth and stable binding of *Fusobacterium nucleatum* to multiple bacterial partners signals a transition to a mature biofilm profile, in which the predominantly anaerobic late coloniser species rise in abundance. Late colonisers are typically anaerobic and proteolytic, and include taxa frequently associated with periodontal disease (274–277). Following this change, the biofilm community remains stable, with no additional structural changes (281, 294). Biofilm formed above the gum line is referred to as supragingival plaque, while that below the gum line is known as subgingival plaque. Due to the anaerobic and serum-rich environment within the gingival crevice, subgingival plaque generally contains a greater abundance of anaerobic and proteolytic bacteria than supragingival plaque (309).

Plaque thickens as bacteria proliferate, forming morphologically distinct clusters called microcolonies (273). Large clusters, termed ‘hedgehogs’ arise from a core set of genera that are highly abundant and prevalent in plaque (255). The architecture of hedgehog structures is related to the functional needs of the participating species, in which different groups occupy different layers according to, for example, oxygen tolerance, an organizational trait previously noted in dental plaque (255, 274–277). One distinct morphology in dental plaque is the ‘corn cob’ structure, made of long thin filamentous species coated with small cocci, and resembling an ear of corn-on-the-cob. Pairs of early and intermediate coloniser species making up these various ‘corn cobs’ include *Corynebacterium matruchotii-S. sanguinis* and *F. nucleatum-S. sanguinis (255, 310)*. Certain microbial genera appear to play especially important roles in biofilm structuring, including *Corynebacterium,* a filamentous clade of bacteria that form the structural basis of hedgehog structures, and *Capnocytophaga gingivalis,* a motile species that can move non-motile species throughout the biofilm as though transporting cargo (255, 311). Bacterial co-adhesion is also important for introducing late colonisers into plaque biofilms. Species that do not bind well or at all to the enamel pellicle, such as *Porphyromonas gingivalis*, bind readily to other species that do, such as *S. gordonii, F. nucleatum*, and *C. matruchotii (273)*.

Calculus formation occurs through the periodic mineralisation of the biofilm community, although the initiation mechanism remains unclear (312). Ancient and historical dental calculus, collected from skeletal material without soft tissues, cannot be reliably classified as supra- or subgingival, and the difference may be indistinct in part because such biofilms reach a high degree of maturity (i.e., experience a proliferation of anaerobes) before calcification (2). Following calcification, biofilm formation starts again through the same process of salivary protein adherence to the mineralised biofilm surface. Abundant plaque buildup, and the presence of calculus, are strongly linked to inflammation of the gingiva, called periodontal disease.

S5.1.1 Periodontal disease and associated microbial taxa

Early work to understand the relationship between plaque and dental diseases such as caries and periodontal disease considered the multispecies nature of the biofilm to be at the heart of plaque pathogenicity (313). This community-driven disease initiation later developed into the ‘ecological plaque hypothesis’ (314). When research on specific organisms that could be isolated from plaque and grown in culture in labs became routine, the focus on the pathogenic nature of plaque switched from community-centered to a specific species-centered view, called the ‘specific plaque hypothesis’ (313). Until the application of Next-Generation Sequencing (NGS) technologies to the study dental plaque, the specific plaque hypothesis dominated thinking about plaque pathogenicity, and much research focused on individual species that are abundant at sites of disease, such as *Streptococcus mutans* in caries and *Porphyromonas gingivalis* in periodontal disease. There is now a growing return to the idea that pathogenicity is driven by changes in the entire plaque community rather than the actions of specific species (314, 315).

Certain abundant species could be detected or cultured out of disease-site plaque that were rarely cultured from healthy-site plaque, and because of this, specific species were considered causative agents of dental caries and periodontal disease (256, 274, 309). *In vivo* animal infection studies confirmed the virulent potential of these species, and much periodontal microbial research has focused on

understanding their pathogenicity (275). Like most microbiology work prior to the development of NGS, these studies were based on culturing and biochemical characterization, which limited research to only those species that could be grown in the lab. The now classic study by Socransky and colleagues (256) used DNA-DNA checkerboard hybridization to detect 40 culturable species in >12,000 subgingival plaque samples. This study reduced the complex nature of plaque into a more easily comprehensible collection of 5 complexes, named after colours, consisting of species that were statistically significantly associated with health or disease states and with each other. In particular, the group of three species making up the "red complex", *Porphyromonas gingivalis*, *Treponema denticola*, and *Tannerella forsythia,* were significantly correlated with each other and with clinical parameters of periodontal disease (256), and are commonly and popularly referenced in periodontal microbial research. These "red complex" species are all fastidious, asaccharolytic, obligate anaerobes that are relatively easy to grow in culture, compared with other species in their genera. However, NGS technologies have allowed researchers to detect many more species in plaque, including those that could not be grown in culture, and from these studies numerous other species have been subsequently associated with periodontal disease (309, 316–320). Supragingival plaque samples from healthy teeth contain low levels of species associated with periodontal disease (321). Culturing these biofilms in media that mimics conditions of inflamed gingiva (protein-rich or added serum) enriched many species associated with periodontal disease, several of which were undetectable in the initial inocula, all detected by shotgun metagenomics. This supports the ecological plaque hypothesis, in which the entire community composition changes to promote disease (321).

Metabolic interactions between members of the "red complex" enhance the growth of these species. All three species grow better when co-cultured or when grown in spent media from one another than when grown alone or in fresh media (322–325), suggesting that the metabolic products of one species can be used as nutrition by the others. It has also been shown that metabolic pathways that are not fully covered by the genome of one species can be completed by another to enable full pathway activity in culture (326). Although these three species are rare in healthy tooth-site plaque of present-day populations and strongly associated with periodontal disease, they are much more prevalent and abundant in the dental calculus of historic populations from both healthy and periodontitis-affected teeth (2, 89). Because of the clinical interest in the "red complex" and other periodontal disease-associated species, we present a brief introduction to several of these species below.

Porphyromonas gingivalis *Porphyromonas gingivalis* is the most heavily-studied member of the "red complex". It is a Gram-negative, non-motile, asaccharolytic, obligately anaerobic coccobacillus. Growth media for *P. gingivalis* is supplemented with heme for optimal growth, and colonies grown on blood agar turn dark brown to black from accumulation of heme on their surface, indicating that the species acquires iron from the host, a pathogenic trait. *P. gingivalis* is detectable in early biofilms by DNA-DNA checkerboard hybridization, but only at very low levels (327), and it is considered a late coloniser. However, it is able to grow as a biofilm with the early colonisers *S. gordonii*, *A. oris*, *and Veillonella* sp., as well as with *S. oralis* together with *S. gordonii (328)*, which may explain its low-level presence in early biofilms. It can also form corncob-like interactions with filamentous species (255) in the part of the biofilm closest to the gingiva (farthest from the tooth surface). Dipeptides are the preferred nutrient source of *P. gingivalis (275)*, and it expresses several proteases, most notably the gingipains (275), to release peptides from larger proteins. Lipopolysaccharide structure varies depending on the environment (329, 330), but it is less immunogenic than the LPS of enteric species.

Treponema denticola *Treponema denticola* is a Gram-negative, motile, asaccharolytic, obligately anaerobic spirochete, and the second member of the "red complex". It is found at the surface of plaque biofilms (277, 293), in anaerobic conditions with readily-available host-derived proteins. The presence of *T. denticola* correlates with periodontal disease progression, even when its abundance is low (331, 332). Growth in culture requires media supplementation with iron as well as short-chain volatile fatty acids. In turn, *T. denticola* produces a variety of short-chain volatile fatty acids including lactate, succinate, and formate, as well as methyl mercaptan and hydrogen sulfide from protein metabolism, all of which contribute to halitosis (333, 334). The proteases cystalysin and dentilisin, major surface protein MSP, chemotaxis proteins, and lipooligosaccharide are all virulence factors that promote pathogenicity towards host immune cells, epithelial cells, and fibroblasts (275, 335, 336). Direct interaction between *T. denticola*

and *T. forsythia* was shown through the leucine-rich repeat proteins LrrA of *T. denticola* and BspA of *T. forsythia (337)*.

Tannerella forsythia *Tannerella forsythia* is a Gram-negative, asaccharolytic, obligately anaerobic bacillus, and the third member of the "red complex". The genus and species name has changed twice, from the originally designated *Bacteriodes forsythus* to *Tannerella forsythensis* to the now-accepted *Tannerella forsythia*. Much less is known about this species than the other members of the "red complex" because it is fastidious and difficult to grow in laboratory settings (338). Growth in culture requires addition of N-acetylmuramic acid, a component of peptidoglycan, which suggests that *T. forsythia* scavenges this molecule from the biofilm rather than synthesizing it. Like *P. gingivalis*, *T. forsythia* acquires iron from the host by binding heme (339), and cleaves dipeptides from host proteins for nutrition (340). Virulence factors include the leucine-rich repeat protein BspA (341), the cysteine protease PrtH (342), and sialidases NanH and SiaHI (343). *Tannerella forsythia* is coated in a unique protein-based capsule called the S-layer that appears to be characteristic of the genus *Tannerella (344)*, although it is found in several Gram positive and Gram negative genera (345). The S-layer contributes to *T. forsythia* virulence (275) by assisting in adhesion to epithelial cells and aiding the bacterium in evading the immune system (346, 347). The two glycoproteins that make up the S-layer, TfsA and TfsB, have no known homology to any other prokaryotic S-layer proteins, and they are among the most abundant proteins expressed by *T. forsythia* cells. Glycosylation of the S-layer proteins modulates immune cell recognition of the organism (348), as well as its role in biofilm formation (349).

Fretibacterium fastidiosum *Fretibacterium fastidiosium* is a recently described species that is associated with periodontal disease (350). It is a member of the phylum Synergistetes, one of the many understudied and poorly characterized phyla. It is a Gram-negative, asaccharolytic, motile, obligately anaerobic, curved rod, yet relatively little is known about this species, since it is only recently possible to grow it in the lab. *In vitro* cultivation depends on co-culture with other oral taxa such as *F. nucleatum (350)* or supplementation of growth media with siderophores (351). Acetic acid, propionic acid, hydrogen sulfide, and proteases are produced by *F. fastidiosum* during culture, similar to other characterised periodontal taxa (350). It is more abundant in plaque of patients with chronic periodontitis than healthy patients (319, 352, 353), and is more abundant in moderate to deep periodontal pockets than in shallow pockets (354), which is concordant with its designation as an obligate anaerobe.

S5.2 Parameter selection of core microbiome analysis

A core microbiome is roughly defined as the microbial genes or taxa that are present across all (or nearly all) individuals of a given group, population(s) or species (355). Defining the core microbiome of different modern human body sites has been a primary goal of many studies (276) including the Human Microbiome Project (356), with the reasoning that these taxa are fundamental for the formation and maintenance of health within a given microbiome. Deviation from this core structure is often described as dysbiosis and is frequently associated with disease (357, 358). While research into 'core' microbiomes at highly variable body sites, such as the gut, has shifted from a focus on core taxa to a focus on core functions (359, 360), the plaque biofilm differs in that it is a stable (361), stratified system with a known developmental process (253, 255, 277, 282). A consistent set of genera have been shown to dominate dental plaque communities compared to other oral surfaces (253, 255, 277, 282), and a growing body of evidence suggests that there is a core oral microbiome present at healthy oral sites (276). To date, however, this research lacks an evolutionary framework and has focused almost exclusively on dental plaque samples from modern human subjects in urban centers from Western industrialised societies. To better understand and define the core human oral microbiome in a broader, evolutionary context, we compared the microbial species and genera found within the dental calculus of the African hominid groups in this study. We sought to determine whether a core microbiome could be defined for each group, whether the core was phylogenetically coherent, and whether some members of the core were specific to certain host groups. We also sought to determine what roles core members play within the biofilm, and whether certain functions are more phylogenetically conserved than others.

Because our study includes both present-day and archaeological samples, we devised a strategy that is robust to minor variations in sample preservation. We selected a three-level threshold approach to

identify potential ‘core’ OTUs in the MALT/MEGAN taxonomic assignments, and a flowchart diagram illustrating this selection process is provided in Fig. S9A and External Data Repository Fig. R24. We then identified a provisional core microbiome for each host genus separately, then in all host-genus combinations following the branching of the host phylogenetic tree, as well as controls (e.g., *Pan:Homo*, *Gorilla:Pan:Homo, Alouatta:Gorilla:Pan:Homo*). Specific details of this process are as follows, with technical descriptions in the scripts described in External Data Repository Section R10.1.1.

First, using the high-quality, filtered dataset described in Sections S3.4 and S4.6, we assigned a taxon as ‘present’ if it had a non-zero value in the OTU table, and ‘absent’ if had a zero value (see below for the minimum fraction of reads required to be a non-zero value). We then required that the taxon be present in a minimum of 50% of samples of a given host genus ‘population’ to be considered core to this population. We kept this first threshold relatively low for three reasons: 1) small samples sizes for some populations, 2) known variability in sample preservation and limited knowledge of taxon-specific preservation characteristics (194), and 3) knowledge that periodontal disease can cause shifts in abundant taxa, making core ‘health’ associated taxa less detectable.

Second, of taxa that we defined as ‘core’ to a host genus population, we further defined the taxon as being core to the host genus overall if it was designated a core taxon in at least two-thirds (66%) of all populations of that host genus. This ensured that taxa that were specific to only one particular population, geographic region, or subspecies were not counted as core for the genus overall. *Alouatta* was exempt from this parameter as only a single population was available. We then intersected the host genus core lists in all possible combinations using the ‘upset’ function from the UpSetR (v1.3.3) package (362, 363).

Third, we observed that when using taxonomic tables generated with a lower minimum abundance thresholds (i.e., the proportion of alignments a taxon has over the total alignments of the sample), well-known environmental contaminants such as *Pseudomonas* and *Streptomyces* were reported as being ‘core’ to African hominids. The finding of such taxa at low abundance in our samples is not unexpected, given that these highly robust environmental species are nearly ubiquitous in burial soils, as well as in museum collections, storage facilities, and even laboratories (364, 365). For example, we have successfully cultured numerous live *Pseudomonas* species and strains (>40) from archaeological bone, and we believe that low-level surface colonization by such taxa is typical for most archaeological and museum collections. Wet lab methods (EDTA wash and UV irradiation) remove some of these contaminants, but not all. Additionally, the ‘decontam’ filter that we applied in section S3.6 was also not sufficient to remove these highly ubiquitous contaminants due to the fact that they were identified in more samples and at higher abundance than other contaminant taxa, in part because of their overrepresentation in reference databases.

To address this, and to increase the confidence that taxa identified as ‘core’ are truly endogenous oral-related taxa, we repeated the analysis with increasing minimum support values until most of these well-known laboratory and environmental contaminants were removed but most well-known oral taxa were retained. Categorization of taxa as oral vs. environmental was based on the predominant reported isolation sources as identified in literature reviews and databases (e.g., human oral microbiome database (HOMD), PubMed, and BacDive). Specifically, we increased the minimum support threshold until there were no host genus-combinations that included the ‘control’ category (consisting of all extraction blanks, library blanks, and bone environmental proxy samples) for the following monophyletic human lineage groups: *Homo, Homo:Pan, Homo:Pan:Gorilla, Homo:Pan:Gorilla:Alouatta*). This ensured that the microbial taxa identified as ‘core’ were highly abundant across most samples and above the ‘noise’ threshold of contaminants present in controls. The corresponding cutoffs were found to be 0.07% and 0.04% for genus and species taxonomic levels, respectively. Setting higher minimum abundance thresholds did not improve the results, and only led to the removal of lower abundance well-known oral taxa. These thresholds were valid for both NCBI nt and custom RefSeq databases, and we therefore applied these cut-offs prior to the compositional analyses described above.

The effect of increasing the minimum support value from baseline (0.01%) to the final selected values (0.07% for genus and 0.04% for species) on host group core microbiome assignment is shown in Fig. S7, External Data Repository Section R10.1.2, and Fig. R25 for the refseq database (see External Data

Repository File R23 for tabular comparisons). Notably, the number of taxa shared by controls and all 4 host genera is much higher than the number shared by controls and 2 or 3 host genera, suggesting these are indeed widespread background contaminants, and that these contaminants are successfully removed by increasing the minimum support value. More taxa are lost when increasing the minimum support value from 0.01% to 0.04% than when increasing it from 0.04% to 0.07%, which is expected given the 'long-tails' of very low abundance (often spurious) taxa typically assigned to large metagenomic datasets (237), a classification artifact that is also known to affect MALT-derived taxonomic tables (204). There are few reassignments of taxa to new host-genera combinations when increasing the minimum support value from 0.04% to 0.07%, suggesting these assignments are robust. These reassignments consist of 2 changes for the nt database at genus level, 4 for nt species, 4 for custom RefSeq genus, and 3 for custom RefSeq species. After this point, the assignments are generally stable, and are not improved by stricter thresholds.

In this step, the taxa that are reassigned from a core combination with a control, to a core combination with host genera only, are *Streptococcus*, *Actinomyces* and *Corynebacterium*. All three of these genera are host-associated and contain many different species that reside in different parts of the modern human body (366), most notably the oral cavity and skin. Without filtering, these genera are likely identified within controls due to low-level skin contamination. However, by increasing the genus minimum abundance threshold to 0.07%, this background skin contamination signal is removed, leaving only the highly abundant oral assignments.

With regard to other reassignments during this step, most represent minor, predictable reclassifications. For example, both reassignments of taxa from *Alouatta:Gorilla:Pan:Homo* to the subclade *Pan:Homo* were anaerobic bacteria (*Porphyronomas*, *Treponema*). This is expected, given the much greater abundance of these taxa in *Homo* and *Pan*, compared to *Gorilla* and *Alouatta*, which have a lower abundance of late colonisers in general and these genera in particular (See section S4.3).

Among the taxa remaining after the application of this filter, all have been previously identified in the oral cavity, with the exception of *Mycobacterium* (External Data Repository Section R10.1.5 and File R22). *Mycobacterium* is only identified as a core member at the genus taxonomic level when using the nt database. This genus consists of a highly diverse set of primarily environmental taxa (367), of which only a small subset causes disease in mammals and may be present in the oral cavity (368). Such diversity is a common obstacle in ancient DNA analysis, where it is difficult to distinguish pathogenic *Mycobacterium tuberculosis* from closely related environmental species (369), especially because, being clinically relevant, pathogenic *Mycobacteria* are substantially overrepresented in reference databases (especially the nt database) than environmental *Mycobacteria*. Because our core assignments do not take into account abundance (beyond filtering by the minimum support value), *Mycobacterium* is likely retained in our core dataset as an artifact of its combination of relatively high abundance in control samples and variable abundance in samples, rather than it being a true oral taxon. Indeed, External Data Repository Section R10.1.5 and Fig. R27 shows that the sources with the greatest number of *Mycobacterium* alignments are from sediment, ancient bone and skin samples by an order of magnitude. The most common *Mycobacterium* detected at the species level were *M. kansasii* and *M. avium* - both taxa that are most commonly isolated from water (370). Because this genus is almost certainly a contaminant, and not a true oral member, we excluded it from downstream analysis. Following the removal of *Mycobacterium*, all remaining core taxa are known to be oral associated, having been previously isolated from the oral cavity or identified through genetic screens.

At the species level, we accepted three taxa as core that failed to meet the above requirements on technical grounds, namely because they either belong to poorly described microbial families that do not yet have accepted genus- and species-level designations (Candidatus Saccharibacteria TM7x and Anaerolineaceae bacterium oral taxon 439) or because their current genus is paraphyletic and undergoing reorganization ([Eubacterium] minutum). All three taxa meet species-level threshold requirements, despite not having an accepted genus designation. We have included these taxa in all downstream reporting.

Overall, we found that applying a minimum support threshold prior to core microbiome assignment was critical for removing spurious assignments due to background environmental contamination. We selected our final thresholds (genus 0.07%; species 0.04%) based on the fact that they are the minimum values required to remove most known contaminants without also removing known oral taxa. By setting a minimum support threshold, some oral taxa previously core to larger clades were reassigned to subclades; however, this affected relatively few taxa. Specific changes included reassignment of *Desulfovibrio, Olsenella,* and *Tannerella* from *Alouatta:Gorilla:Pan:Homo* to *Gorilla:Pan:Homo*; reassignment of *Porphyromonas* and *Treponema* from *Alouatta:Gorilla:Pan:Homo* to *Pan:Homo*; and reassignment of *Mogibacterium* from *Gorilla:Pan:Homo* to *Pan:Homo*.

As a final check to ensure that the core microbiome assignments are robust and not strongly influenced by sampling or preservation effects, we examined the effect of our initial sample filtering step (section S3.4) on core microbiome assignments. We found that following the removal of low quality samples, one *Gorilla* population was reduced to a single individual. We then checked for the influence of this individual on the calculation of the core taxa (Fig. S8, External Data Repository Section R10.1.3, and Fig. R26, with a tabular comparison in File R24). Very few changes were observed at the genus level, and for both the genus and species levels, only two reassignment combinations occurred within our human lineage core combinations of interest (i.e., *Alouatta:Gorilla:Pan:Homo, Gorilla:Pan:Homo*, *Pan:Homo*, and *Homo*). In both cases (*Capnocytophaga* at genus level, and *Streptococcus constellatus* at species level), this consisted of a reassignment from *Alouatta:Gorilla:Pan:Homo* to the *Alouatta:Pan:Homo* combination. Outside of the primary core combinations of interest, all combination reassignments at genus level consisted of *Alouatta:Gorilla* assignments being reassigned to just *Alouatta.* Equally, all species affected by reassignments are combinations including *Alouatta* that result in *Gorilla* being removed from the combination. Two *Staphylococcus* species that were core to *Gorilla* were removed entirely from the core calculations. Due to the similarity of *Alouatta* and *Gorilla* in all of our analysis, we decided to retain the single individual population for this calculation. Furthermore, the two taxa that are reassigned away from core combinations of interest are known oral taxa in modern humans and are otherwise ubiquitous across all other host genera - thus suggesting that they are true core oral taxa, even if at lower prevalence across the two multi-individual *Gorilla* populations.

A table of all taxa that are considered a part of the host genus phylogenetic core at each minimum support level is provided in Data S3. Final summary statistics and core taxa can be seen below in section S5.4.

S5.3 Comparison among African hominid core microbiomes

In order to explore the identified core microbiome(s) in more detail, we used UpSet (362, 363) plots to visualise the number of core taxa in all possible host group combinations at both the genus and species levels and for both the NCBI nt (Fig. 2A-B) and custom RefSeq databases (External Data Repository Section R10.1.6 and Fig. R28).

In all cases we observe that the number of unique core taxa in *Pan* and *Alouatta* is higher than in *Homo* and *Gorilla.* This suggests that *Homo* and *Gorilla* each share a high number of core taxa with all of the other host genera. At the genus level, the number of taxa shared across *Alouatta:Gorilla:Pan:Homo* is the next highest, highlighting the similarity of all of the primate oral microbiomes included in this study. However, this pattern is less clear at the species level, suggesting that while there is a general conservation of the core oral microbiome, there are host-specific differences at the microbial species level.

To assess the support for core microbiome assignments, we compared the assignments made using the NCBI nt and NCBI RefSeq databases for corroboration (External Data Repository Section R10.1.6 and File R25). Of the 75 microbial genera assigned to the core microbiome of a host genus, 26 taxa were in agreement between databases, 14 taxa had different host genus assignments, and 35 taxa did not pass the prevalence thresholds in one or the other database (18 assigned in the nt database and not assigned in the custom RefSeq, and 17 vice versa). Of the 223 core-microbiome species assignments, 34 were in agreement, 30 had different host genus assignments, and 159 did not pass the prevalence thresholds in

one or the other database (81 assigned only in the nt database, and 78 in vice versa). In both cases more than half of the taxon combination assignments were not shared between the two databases, highlighting the importance of database selection and representation. While the core taxa corroborated by both databases are likely highly robust, the status of those in disagreement may change over time as a wider diversity of taxa (especially from living primates and more diverse modern human groups) are sequenced and databases improve (258). Revisiting this data in the future will likely improve the identification of core taxa shared across hominids, as well as identify additional core taxa for each host genus.

When looking specifically at host phylogenetic clade combinations, corroboration of core microbiome assignments between the two databases was much higher at both the genus level (≥ 50% of assignments corroborated for *Pan:Homo*, *Gorilla:Pan:Homo*, and *Alouatta:Gorilla:Pan:Homo*) and species level (all ≥ 60%) (External Data Repository Section R10.1.6 and File R25). Of the combinations that do not corroborate, none involve genera being reassigned to other host-groups and only 4 result from the reassignment of a microbial species to a different host-group. Of these, 3 species (Anaerolineaceae bacter oral taxon 439, *Olsenella* sp. oral taxon 807, *Tannerella forsythia*) were reassigned from a hominid combination containing *Gorilla* in the nt database assignments to the same combination but without *Gorilla* in the custom RefSeq database, and one (*Campylobacter gracilis*) was reassigned from the *Alouatta:Gorilla:Pan:Homo* group in the nt database to the *Gorilla:Pan:Homo* group in the custom RefSeq database. The remaining 40 non-corroborating combinations were due to a given taxon not passing prevalence thresholds in one or the other database (9 at genus level, 31 at species level). Given the differences between the two databases, this pattern is not entirely unexpected and may be attributed to evolutionary divergences between the actual microbial species present in each host and the closest reference sequence in the database, as well as to the problem that databases with large numbers of reference genomes can reduce the specificity with which specific (especially conserved) reads are assigned at the species level, thereby spreading such reads across many closely related species (258).

To further add support to our core microbiome assignments, we performed bootstrapping to assess the stability of taxon assignments to particular host-genus combinations. For each host-genus, we randomly subsampled individuals (with replacement) and re-performed the core calculation for a total of 1000 replicates. To assess the effect of including controls, we ran the bootstrapping procedure on datasets for both databases (NCBI nt and RefSeq) and microbial taxonomic levels (species and genus) with and without controls (excluding blanks from both). The majority of *Alouatta:Gorilla:Pan:Homo* combinations with the NCBI nt database were robust when considering a minimum bootstrap support of 75%, with only *Capnocytophaga* falling below this, at 71.6%, when excluding controls (External Data Repository Section R10.1.4 and Table R8). This adds robusticity to our core microbiome calculations, despite small sample sizes for each host genus, by showing relatively low stochasticity in the assignment.

The procedures used to perform core microbiome calculations, alluvial plot visualizations, *Mycobacterium* analyses,UpSet plot visualizations and bootstrapping are described in the External Data Repository Section R10.

S5.4 Evolutionary and ecological relationships among core oral microbes of African hominids

In total we determined 23 genera and 5 independent species to be core to the human oral microbiome (Fig. 2C), of which 19 genera and 4 species (in both cases excluding *Mycobacterium*, see section S5.2) fall exclusively within expected host phylogenetic groupings: ((((*Homo*)*Pan*)*Gorilla*)*Alouatta*). Nearly half of these human core taxa (10 genera) are core members of all four host genera in this study (*Alouatta:Gorilla:Pan:Homo*), suggesting a high level of taxonomic conservation in the primate oral microbiomes in this study. The spatial location of these organisms (*Actinomyces, Streptococcus, Campylobacter, Capnocytophaga, Corynebacterium, Fusobacterium, Ottowia, Prevotella, Pseudopropionibacterium, Selenomonas*) spans the entire biofilm, from the basal to external layers (255, 274–277, 371–375). Many of these core genera play key roles in biofilm architecture. For example, *Actinomyces* and *Streptococcus* are responsible for initial colonization and biofilm initiation (see section S5.1). *Corynebacterium* form long filaments that co-aggregate with other taxa to produce 'corncob' and 'hedgehog' scaffolds that act as 'structural pylons' in biofilm architecture and create habitats for other bacteria to grow and exchange nutrients (255, 376, 377). *Capnocytophaga* are capable of gliding motility

and serve as microbial shuttles for non-motile taxa throughout the biofilm (378), and *Fusobacterium* is a critical bridging taxon that expresses numerous surface adhesins and links to early and late biofilm colonisers (379). *Selenomonas* use their flagella to form interwoven web-like structures that are thought to fulfill architectural functions in the biofilm (375). Less is known about the specific roles of *Campylobacter*, *Ottowia*, *Pseudopropionibacterium*, and *Prevotella* in oral biofilms. We note that the genus name *Pseudopropionibacterium* was mis-applied and may soon be reassigned to *Arachnia (380)*.

To more closely examine the spatial relationships, biofilm roles, potential coevolutionary partners of core microbiome taxa, we developed a schematic representation of the biofilm (Fig. 2C) based on previously reported species and/or genus localization determined by a variety of *in situ* imaging techniques (255, 274–277, 371–375). A three-layer model of dental plaque based on microbial succession has been described using FISH (255, 277), and we followed this model in our analyses. In Fig. 2C, we have coloured each microbial genus according to its respective host group, focusing on those groups cladal to humans (*Alouatta:Gorilla:Pan:Homo*, *Gorilla:Pan:Homo*, *Pan:Homo*, and *Homo*). Four taxa were represented by species rather than genera, either because they lacked valid genus designations (Anaerolinaceae bacterium oral taxon 439 (phylum Chloroflexi); [*Eubacterium*] *minutum*; TM7x) or because the genus had insufficient genome representation in the database (*Filifactor alocis*) to meet genus-level thresholds. No spatial information was available for several core taxa (Anaerolineaceae bacterium oral taxon 439, *Desulfomicrobium, [Eubacterium] minutum*, *Gemella* sp. oral taxon 928, *Mogibacterium*, *Olsenella*, and *Ottowia*), and so we have omitted these taxa from the figure. In addition, four genera were found to be core to paraphyletic host groups, but were included in the Fig. 2 because they are core microbiome members of *Homo*: *Leptotrichia* (*Alouatta:Homo*); *Aggregatibacter, Eikenella,* and *Neisseria* (*Alouatta:Gorilla:Homo*). These taxa were shaded gray. For *Gemella* and *Filifactor*, there was an inconsistency in the core host group between the bacterial genus and species: *Gemella* (genus) and *Filifactor* (genus) are core to *Pan* only, but *Gemella* sp. oral taxon 928 (species) is core to *Alouatta:Gorilla:Homo* while *Filifactor alocis* (species) is core to *Pan:Homo*. For these 2 species, we have included them in the Fig. 2 with their species core host groups.

Importantly, the core *Alouatta:Gorilla:Pan:Homo* taxa make up all but two of the genera forming the basal layer of the oral biofilm. This suggests that the earliest stages of biofilm development are conserved, while the metabolic processes stemming from the growth and interactions of these organisms likely can support a larger and more variable community of additional species in their hosts. Together, members of the *Alouatta:Gorilla:Pan:Homo* core and the *Alouatta:Gorilla:Homo* core (*Aggregatibacter, Eikenella, Neisseria*) are largely associated with oral health in present-day modern humans, while those in the *Gorilla-Pan-Homo* core (*Fretibacterium, Olsenella, Tannerella*) and especially the *Pan-Homo* core (*Desulfomicrobium, Mogibacterium, Parvimonas, Porphyromonas, Treponema*) are more typically associated with oral disease, particularly periodontal disease (256, 309), in present-day modern humans. Unfortunately, we do not have periodontal health assessments for the majority of teeth sampled in this study, and thus we are unable to comment on their disease status. However, it is possible that the strong clinical association of species such as *Porphyromonas*, *Treponema*, *Filifactor*, *Mogibacterium*, etc., with periodontal disease in humans may stem from their more recent evolutionary association or from extensive modern oral hygiene procedures that disrupt the biofilm community. This suggests a disconnect between the biofilm species profile and the immune response, rather than these species being inherently pathogenic (381). Specifically, the gingival immune response evolved to respond to biofilms that developed undisturbed, and reached a particular balance of microbes and microbial gene expression. The immune response tolerates these homeostatic conditions, while disturbances in relation to it trigger a response. Now through modern hygiene practices and antibiotic use, biofilm development is frequently disturbed, and may not reach the same balance of microbes and microbial gene expression. Consequently, the metabolic signals produced by the community may be eliciting a strong inflammatory immune response, in part because the immune system of the host has yet not adapted to these new conditions.

The only core genus that was found to be *Homo*-specific was *Veillonella*. Species in this genus have a close metabolic dependency on *Streptococcus* species, whereby *Veillonella* use *Streptococcus*-produced organic acids as a primary carbon source (382), possibly explaining why it is uniquely core in *Homo*. Certain *Streptococcus* species, specifically those considered to be characteristically oral taxa, are more

abundant and prevalent in *Homo* than in other host genera, as shown in section S5.6, and these streptococci may be necessary for *Veillonella* to flourish. Although other *Streptococcus* species are prevalent and abundant in other host genera, the relationship between *Streptococcus* and *Veillonella* may have uniquely developed in *Homo*.

S5.5 Evolution of oral disease virulence factors in late-colonisers

We observed above that several classic periodontal disease-associated taxa (256) (see section S5.1.1) were sufficiently prevalent and abundant to be considered core to *Gorilla:Pan:Homo* or *Pan:Homo*. Even when not core to a particular group, we observed that reads matching these taxa were sometimes relatively high in a few individuals. Because many modern human pathogens evolve from non-disease-causing commensal relatives (383), we wished to determine whether the reads assigned to these taxa in non-human primates included the specific virulence factors that are associated with pathogenicity in modern humans. To test this, we focused on a set of relatively well-characterised virulence factors from *Tannerella forsythia* and *Porphyromonas gingivalis (275, 338)*. We did not include *Treponema denticola* in this analysis because the genus was not found to be sufficiently represented, even at low levels, in members of *Alouatta* or *Gorilla* to allow meaningful comparison (see section S6.2). We also excluded *Fretibacterium fastidiosum,* as this is a more recently identified periodontitis-associated taxon for which specific virulence factors are not well-characterised.

Using the deep sequencing dataset, we mapped all well-preserved samples to the reference genomes of *Tannerella forsythia* and *Porphymonas gingivalis* (see section S6.2), and we calculated the depth and breadth of coverage for all annotated genes in the two genomes using bedtools coverage (v.25.0) (384). To account for possible high divergence of virulence gene sequences across host genera, as well as low coverage mappings with insufficient information, we discarded all annotated genes that did not have a breadth of coverage of ≥ 70%. This resulted in a clear differentiation between individuals with low and high coverage genomes (External Data Repository Section R12.1 and Fig. R35), and we next discarded all individuals who had fewer than 500 annotated genes passing the breadth of coverage filter. For the remaining mappings, we calculated median gene depth coverages for each individual. To remove likely outlier genes with unusually high or low coverage we then discarded any gene with a depth of coverage that fell outside of the range of ± 75% of the median. This ensured that samples with normal or nearly normal distributed gene coverage distributions retained most of their genes, while some of the more outlier genes were removed from samples with highly skewed or uniform distributions (*in lieu* of a better consistent threshold across heterogeneous coverage distributions). We did this because very high gene coverage can be an indication of multicopy genes or of taxonomic crossmapping of highly conserved (i.e., housekeeping) genes, while very low coverage may not afford sufficient information for analysis. For each virulence factor gene (External Data Repository Section R12.1 and File R35), we then calculated the ratio of the average depth coverage of all well-supported genes to the depth coverage for each virulence factor, with the assumption that a virulence factor that is truly present should be found at a ratio of 1:1 with other single copy genes. The results of this analysis are provided in Fig. S9B-C and External Data Repository Fig. R36.

We find that *P. gingivalis* and *T. forsythia* virulence factor genes appear to be present within the oral microbiomes of all African hominids and the howler monkey outgroup. Thus, neither these taxa nor their virulence factors are specific to modern humans. All of the *T. forsythia* virulence genes were identified in the majority individuals harboring the species, and at mostly similar coverage depths as other genes in the genome. *P. gingivalis* virulence genes exhibited higher variability, and in particular the *Mfa* series of genes was absent in three quarters of *Pan* mappings, both *Alouatta* mappings, and one *Homo* mapping. By contrast, gingipains (*kgp*, *rgpA* and *rgpB*) were present across all individuals carrying *P. gingivalis*, although at a reduced coverage in *Aouatta* and *Gorilla*, suggesting that not all strains in these groups have these virulence factors. The shared presence of virulence factors in most individuals harboring *P. gingivalis* and *T. forsythia* suggests that the acquisition of these genes, and the pathogenic lifestyle they enable (381), has a time depth long predating modern humans. The pathogenic potential of such genes in non-human hosts has been demonstrated by their *in vitro* expression in animal models (385).

Although most virulence genes were identified in non-human primates, other genes systematically differed between host groups, suggesting host genus-specific strain differences. For example, we found that several regions of the *P. gingivalis* reference genome were not covered by chimpanzee samples. Ozga, et al. (386) mapped their chimpanzee dental calculus to the same *P. gingivalis* reference genome that we used, and also found a similar pattern. Although they did not further investigate the missing genes, the concordance in the findings of the two studies supports the notion that chimpanzees and modern humans carry different host-specific clades of *P. gingivalis*, with different gene content and different virulence potential. The code for generating virulence gene coverage ratios is provided in the External Data Repository Section R12.1.

S5.6 Host-associated differences in early-coloniser Streptococcus species

Streptococcus is a large and diverse genus of Gram-positive bacteria. They have undergone numerous name changes and re-classifications based on biochemical and phylogenetic analyses, and current nomenclature is still under debate (2, 387, 388). Broad categories of *Streptococcus* were described in early classification, such as the modern human-ubiquitous oral streptococci (formerly the viridans group), and the pyogenic group, which includes species infrequently isolated from modern humans, except in the context of disease (i.e., *S. pyogenes* and *S. agalactiae*).

In modern humans, members of *Streptococcus* are the most abundant species in the mouth (149), and they are universally present in all sites of the oral cavity (389). They are the dominant genus found in saliva and on soft tissues of the mouth (389), and they are one of the most abundant genera in modern human supragingival plaque (390) and healthy-site subgingival plaque (309). Many *Streptococcus* species have been isolated from dental plaque, the most common of which comprise two oral *Streptococcus* groups, the mitis group and the sanguinis group (390, 391). Members of these groups, together with a third group known as the salivarius group, have been found to express one or more salivary α-amylase-binding proteins that allow them to derive nutrition from dietary starches (392). They are the only oral species known to have this ability. Further information about these streptococcal groups with α-amylase activity is provided in section S5.6.1.

During hierarchical clustering analysis, we observed that *Pan* has a very low prevalence and relative abundance of *Streptococcus*, while *Homo* has a very high prevalence and abundance of this same genus (see section S4.3). Consequently, we wished to examine the relative abundance and distribution of different *Streptococcus* species across the host genera in more detail. To do this, we first filtered the species-level MALT OTU table that was used for core microbiome analysis (Section 5.2) to include only *Streptococcus* species assignments. Based on species level 16S rRNA gene and whole genome sequences (391, 393, 394), we then assigned each *Streptococcus* species to a broader *Streptococcus* group (389): mitis, sanguinis, anginosus, salivarius, mutans, pyogenic, or other. The group assignments are provided in External Data Repository Section R12.2.1 and File R33. We then calculated the fraction of alignments to each *Streptococcus* group over total alignments to *Streptococcus* references at species level, and the results are shown in Fig. 5B and External Data Repository Section R12.2.1 and Fig. R64 (for sample labels).

To further verify these patterns, we then repeated this analysis on our much deeper sequencing dataset, which we aligned to a concatenated *Streptococcus* super-reference genome (see section S6.1 for details) that included one reference genome per NCBI species-level name in the genus *Streptococcus* (External Data Repository Section R12.2.1 and Fig. R37). We performed this analysis to account for the taxonomically imbalanced range of reference sequences in the NCBI nt database, as well as to account for the lower sequencing depth of the shallow sequencing dataset. This test is described in detail in the External Data Repository Section R12.2.1. Note that the larger fraction of unknown in the super-reference plot is likely due to the absence of a minimum support value being applied - i.e., many of the unknown group species have few aligned reads by MALT, and thus are removed when the increased minimum support value (0.04%) is applied as in Fig. 5B.

Overall, we observe high concordance between both analyses. In both cases, we observe that streptococci within *Homo* oral biofilms are particularly dominated by alignments to the sanguinis group.

By contrast, this group makes up a much smaller fraction of streptococci in *Alouatta* and *Gorilla*, and - with the exception of a single individual - it is entirely absent in *Pan.* Although *Alouatta* and *Gorilla* have fairly balanced distributions of the different Streptococcus groups, *Pan* differs from the other groups in having very low streptococcal diversity consisting mostly of anginosus and pyogenic group streptococci, with a few individuals also having mutans group streptococci. We note that a subset of the modern human individuals appear to have similar distributions of taxa to *Pan*, i.e., the lack of sanguinis, mitis and salivarius groups. These individuals overlap with those appearing '*Pan*-like' in taxonomic profiles (see Section 4.3), who span both Europe and Africa, pre- and post-agricultural lifestyle transitions, biological sex and teeth type. We do not have in this study consistent health status information to infer health related factors for this pattern. However the very low streptococcal diversity observed in *Pan* may be related to the overall much lower relative abundance of aerobic and microaerophilic early colonisers (including *Streptococcus*) observed in chimpanzee oral biofilms by hierarchical clustering and indicator analysis.

Of all the host genera, *Alouatta* has the highest diversity of private streptococcal species, an observation that is consistent with the high number of unique and associated 'combinations' also identified by indicator analysis (see section S4.4) and by hierarchical clustering heatmap analysis (see section S4.3). This high level of diversity may either be due to an actual greater diversity of *Streptococcus* species in the oral cavity of *Alouatta*, or to the absence of sufficiently related reference sequences in the database for true *Alouatta* streptococcal species. If reference sequences are too distant from the true sequence for adequate mapping, this can lead to a seemingly 'random' spreading of read assignment across other species in the genus. This artifact, which often manifests as a diverse distribution of alignments over less well-characterised species and strains, may explain the relatively high fraction of 'unknown' and 'unnamed' *Streptococcus* group assignments for both *Alouatta* and *Gorilla*.

S5.6.1 Streptococcus α-amylase-binding proteins as an indicator of host-microbe coevolution

Several oral streptococci are capable of binding salivary α-amylase, which provides easy access to nutrition and may assist with biofilm formation (392). These taxa use amylase-binding protein genes to 'hijack' the high levels of α-amylase enzymes in modern human saliva for two purposes: (1) to access the simple sugars released by enzymatic activity of α-amylase on dietary starch, and (2) to assist in the anchoring of the bacterial cells to the enamel tooth surface (onto which salivary α-amylase is adsorbed) during plaque biofilm initiation (395). Salivary α-amylase is the most abundant enzyme in modern human saliva (288), where it breaks down dietary starch into maltose and dextrins, and it is an important component of the acquired enamel pellicle (AEP). It is also among the most abundant modern human proteins found in ancient and historic dental calculus (2, 89, 396).

Salivary α-amylase in the hominid lineage evolved from pancreatic amylase by a gene duplication event followed by the insertion of a gamma-actin pseudogene and later the insertion of an endogenous retrovirus (397–401). Salivary α-amylase is a multi-copy gene in some apes (33, 46), but humans have the highest reported copy numbers (mean ~5-8 copies) as well as the highest copy number variation (up to 30 diploid copies) (33, 34, 155, 156). Copy number expansion of the modern human salivary α-amylase gene occurred after the split with chimpanzees, within the last 5-7 my (400)Neanderthals and Denisovans appear to only have two diploid copies of the salivary α-amylase gene, suggesting that high *AMY1* copy number variation may be a distinctly modern human trait that emerged after the modern human lineage split with the Neanderthal and Denisovan lineages (402). Among modern human populations, those that consume more dietary starch appear to have higher average copy numbers than populations that consume low levels of dietary starch (33, 46), suggesting that salivary α-amylase copy number variation may track changes in starch consumption during recent human evolution. In addition, salivary α-amylase copy number has also been correlated with salivary microbiome richness (156), suggesting a link between diet-related evolutionary change and microbiome composition. Taken together, promoter region variation combined with copy number variation and microbiome associations make salivary α-amylase a unique marker for understanding the relationship between diet and hominid evolution.

Human salivary α-amylase concentrations are higher than other apes, as well as most Old World and New World monkeys (25, 34, 35, 403). Only baboons have higher salivary amylase than humans (despite having a low *AMY1* copy number), which may be related to their unique use of cheek pouch food storage (47). Humans have the highest measured salivary amylase activity of the apes, while gorillas and orangutans have higher activity than chimpanzees or bonobos, which are comparable (34, 35, 404). Secretion of amylase has been shown to increase with increasing salivary flow rate, but solubilized starch solutions have minimal effect on stimulating salivary secretions in the parotid gland (405). Salivary α-amylase is detectable in low levels in saliva of a variety of herbivores and omnivores that consume plants that are high in starch, but not in carnivores (406). Domesticated dogs exhibit convergent evolution in having also undergone amylase copy number expansion, but in pancreatic β-amylase (34, 407, 408). Across mammals, salivary α-amylase expression is highest in animals with broad range diets (34), and although salivary α-amylase expression is highly variable among primates, it is thought to be related to diet (406).

A variety of human-associated *Streptococcus* species*,* many of which are abundant early biofilm colonisers (409, 410), are able to bind salivary α-amylase. They use a variety of α-amylase-binding proteins termed **a**mylase-**b**inding **p**rotein A (AbpA), AbpB, AbpC, and several other poorly-characterised proteins (392). These different α-amylase-binding proteins share no homology, but rather appear to confer a similar phenotype through convergent evolution. Amylase-binding activity has been reported in nine oral *Streptococcus* species: *S. australis*, *S. infantis*, *S. mitis*, *S. oralis*, and *S. parasanguinis* (the mitis group), *S. cristatus*, *S. gordonii*, and *S. sanguinis* (the sanguinis group) and *S. salivarius* (the salivarius group) (392). Amylase-binding activity is not universal in any species, and there is considerable strain variation for this trait (411, 412). Amylase-binding proteins have not been reported in any bacteria outside the *Streptococcus* mitis, sanguinis, and salivarius groups.

<u>*abpA*</u> Amylase-binding protein A (*abpA*) in *S. gordonii* was the first genetically identified streptococcal amylase-binding protein (413). It is 588 nucleotides long, contains a signal peptide (413, 414) and is active at the cell surface near sites of cell division (414, 415) , but it is also detectable in cell culture supernatants (412, 414, 416). The gene is co-transcribed with a downstream sortase *srtB* that anchors ApbA in the cell wall (414, 417). The gene is expressed in cultures of *S. gordonii* only in the presence of both starch and salivary α-amylase (418), and is likely regulated by the breakdown products of starch (418). Salivary α-amylase stimulates the growth of AbpA-containing *S. gordonii* strains, and it also increases their resistance to low pH (419). Phylogenetic analysis of the N-terminal sequence of AbpA-like sequences, identified in a variety of *Streptococcus* species by BLAST searches, cluster the sequences into 5 groups, which were suggested to have evolved independently (392). The two genes *abpA*-*srtB* appear to be horizontally-transferred together as a genetic island (392), but the origin of these genes in *Streptococcus* is not known and homologs have not been found in other taxa (however, Haase et al. argue there may be an ortholog in *Gemella haemolysans* M341 (392)).

<u>*abpB*</u> Amylase-binding protein B (AbpB) was first described in *S. gordonii (420)*, is 1,959 nucleotides long and shares homology with dipeptdyl-peptidases of other species, but not with AbpA (421). Unlike AbpA, AbpB does not bind α-amylase to the cell surface (420), and instead cleaves collagen and fibrinogen, and interacts with glycosyltransferase G (Gtf-G), yet is needed for optimal colonization of tooth surfaces by *S. gordonii (421)*. A complex of AbpA, AbpB, and Gtf-G, which produces extracellular glucans from sucrose, can be precipitated from saliva by α-amylase (420–422), suggesting a close association *in vitro*. Sucrose breakdown and production of alpha-linked glucans by Gtf-G is elevated in the presence of AbpB but not AbpA (421), and therefore AbpB may participate in bacterial tooth colonization and biofilm formation by complexing AbpA and Gtf-G to bring dietary starch metabolism and bacterial glucan production in close proximity (421).

<u>*abpC* and others</u> Amylase-binding protein C (AbpC) and other putative amylase-binding proteins are to date poorly characterised (392), and are not further considered in this study. AbpC differs greatly from AbpA and AbpB, and shares domain sequence similarities to glycosyltransferases and choline-binding proteins in *S. oralis (419, 423)*. It is found thus far in strains without AbpA and AbpB, and the protein appears to be associated with the cell wall (419). Other putative amylase-binding proteins also exhibit possible homology to choline-binding proteins and ABC-transporter substrate-binding proteins (392);

however their functional significance remains to be determined. Another recently identified group of putative amylase-binding proteins form fimbriae or pilus structures capable of binding to surface-bound salivary α-amylase (419). These proteins are found in the sanguinis and mutans *Streptococcus* groups, and appear to be more involved in biofilm initiation and adhesion to teeth than in nutrient harvest.

S5.7 Detection of amylase-binding protein genes abpA and abpB

During compositional analysis, we observed that among early colonisers, *Streptococcus* species that are known to have amylase-binding capability in modern humans were found at relatively high abundance in the calculus of *Homo* (Fig. 5b; excluding the *Pan*-like individuals, see section S4.3). These species were found at lower abundances in *Alouatta* and *Gorilla*, and were effectively absent from the calculus of *Pan*, even though strains of *Streptococcus oralis* (a member of the mitis group) have been previously isolated from chimpanzee saliva (wadges) (424). The observation in our dataset of greater proportions of these taxa in *Homo* is interesting due to the much higher copy number of salivary α-amylase genes in the modern human genome compared to other primates (46). The ability to procure and digest starch has been proposed to confer an evolutionary advantage to *Homo (46, 157, 425)*. We therefore wanted to further investigate amylase-binding protein genes in our calculus dataset and determine whether they may assist in understanding the evolution of hominid diets and the role of dietary starches in shaping human evolution.

To identify potential amylase-binding-protein genes in the *Streptococcus* super-reference (see below section S6.1), we downloaded from NCBI the corresponding Genbank files for each of the reference genomes in the concatenated super-reference, and removed any that did not contain annotations. We then ran panX (v1.5.1) (426) with default parameters to cluster annotations. We loaded the output of the panX analysis into a local instance of the panX visualisation web server. We searched for the gene clusters containing '*abpA*' (amylase-binding-protein A) and '*abpB*' (amylase-binding-protein B; both derived from the *S. gordonii* reference genome, NCBI Assembly Accession: GCA_000017005.1), and downloaded the corresponding FASTA alignments from the sequence alignment table. Visual inspection of the alignments showed some divergent and/or low quality alignments. To filter for confident alignments similar to the well-characterised *abpA* and *abpB* sequences in *S. gordonii*, we first calculated a pairwise distance matrix using the 'dist.alignment' function in the seqinr R package (v3.4.5) (427). - Clusters of identity similarity were observed at different levels (e.g., sequences clustering at ~50%, 63% and 73% identity, respectively). Therefore to retain more confident alignments, for abpA we retained alignments with ≥ 80% identity, and for abpB ≥ 70%. As a sanity check, we cross-referenced the taxa containing these alignments with those showing in vitro evidence of amylase activity as reported by Haase et al. (392), by extracting taxonomic information from the assembly using the Rentrez R package (v1.2.2) (428). Most publicly available Streptococcus genomes do not have an annotated abpB CDS. In our search, often the genes that were aligning to the *abpB* reference sequence were annotated as a dipeptidase. Therefore we retained annotations labelled "abpB-like_dipeptidase_lipoprotein", due to very high sequence similarity to the *abpB* gene sequence itself, and to maximise the identification of as many potential *abpB* reads possible (given the low coverage nature of our genomes).

The procedure for the identification of amylase-binding-protein gene-like annotations is provided in the External Data Repository Section R12.2.2. The identified annotations with corresponding NCBI assembly ID are provided in External Data Repository File R34. The procedure for identifying the location of these annotations in the *Streptococcus* super-reference is also provided in External Data Repository Section R12.2.2. We used bedtools intersect (see section S5.5)to extract read counts for these regions from each individual. We then normalised the number of *abpA* and *abpB*-like reads to the total number of *Streptococcus* reads identified by alignment to the super-reference, as described in External Data Repository Section R12.2.2.

As shown in Fig. 5C and External Data Repository Section R12.2.2 and Fig. R65 (for sample labels), we find that the relative abundance of *abpA*-like reads in non-human primates is negligible, in contrast to *Homo* where they were identified in 8 of 10 modern humans and 1 of 3 Neanderthals. More reads were found to align to *abpB* overall, but far greater levels were detected in *Homo* than in non-human primates. Among *Homo*, *abpB* was detected in all individuals, and at moderate to high levels in 3 of 3 Neanderthals

and 9 of 10 modern humans. Mann-Whitney U tests applied to the ratios of mapped *abpA* and *abpB* reads to total *Streptococcus* reads, respectively, showed significant (α = 0.05) differences between Homo and non-Homo groups (albeit with wide confidence intervals due to the small sample sizes). The results are as follows: *abpA* - *U* = 112, *p* = 0.001657995, effect size of *r*: 0.663, 95% confidence interval (CI) = 0.398-0.861; abpB - *U* = 128, *p* = 0.000006992604, effect size of *r* = 0.815, 95% CI = 0.686-0.851. Effect sizes were calculated using the rcompanion R package (429). Additionally, a randomisation test of Mann-Whitney U tests applied to 100 random shuffles of sample group assignments showed the *p*-values of the true results fell well-outside the distribution of the random shuffles (see External Data Repository Section R12.2.2 and Fig. R63), supporting that the difference between *Homo* and non-*Homo* groups is not due to random chance.This pattern of absent-to-low *abpA* and *abpB* in non-human primates and conversely moderate to high *abpA* and *abpB* in modern humans is striking in its similarity to known AMY copy number differences between non-human primates and humans. If both are connected to starch consumption patterns, the *Streptococcus* amylase-binding proteins would present a good target for investigating host and commensal microbe evolution, with the advantage that ancient dental calculus provides evolutionary scale calibration points.

Although only three Neanderthals had sufficient sequencing depth for this analysis, they exhibited an interesting, contrasting pattern with respect to amylase-binding proteins. Overall, Neanderthals had little representation of *abpA*, which was found at very low proportions, not dissimilar from non-human primates. However, in contrast, abpB was found among Neanderthals, and the proportion of *abpB* reads falls within the range seen for modern humans.

This observation for *abpB* is interesting in light of Inchley et al. (94), who proposed that the selective sweep of the salivary α-amylase locus in modern humans occurred after the split from Neanderthals. The high levels of *abpB* observed in Neanderthals suggests that dietary starch may have been important in the diet of *Homo* prior to the split with Neanderthals. However, the presence of *apbB* without equivalent amounts of *apbA* is difficult to interpret, as AbpB is also a peptidase, and amylase-binding may have initially emerged as a secondary function. Further research is therefore required to confirm the presence of different types of amylase-binding-proteins (and their genes) in Neanderthals to better understand potential microbial responses to starch consumption.

S5.8 Salivary amylase, cooked starches, and the Expensive-Tissue Hypothesis of human evolution

Homo differs substantially from other great apes in many features, including by having a greatly enlarged brain and a reduced gut, but an otherwise normal basal metabolic rate for their body size (430). This paradox has long posed a puzzle to evolutionary anthropologists because the brain is an energetically expensive organ, requiring a substantial amount of glucose to function (accounting for 20% of adult basal metabolic expenditure (425, 431)), and yet a reduction in gut size would seemingly result in less - not more - nutrient availability to support a large brain. In 1995, Aiello and Wheeler proposed a potential solution to this problem, which they called the Expensive-Tissue Hypothesis (430). Their hypothesis posited that the encephalization seen in *Homo* coincided with a major dietary shift towards higher quality (i.e., more energy dense) foods that were more easily digested, and thus did not require a large hind gut for extended fermentation and digestion. Under this model, higher quality diets relax the metabolic constraints on encephalization and allow the gut to reduce in size, which also reduces the metabolic cost of maintaining the gut, and the gain in brain size and cognitive function enables more complex foraging patterns, which further increases the quality of the diet. Recently, a modification of this theory was proposed, whereby the role of reducing adipose fat reserves was examined as an important factor in encephalization (432). During the course of hominin evolution, brain size tripled between early australopithecines and humans, with a period of major growth associated with the emergence of Middle Pleistocene ‘archaic’ *Homo* species that expanded both within and outside of Africa, leading to the colonization of Eurasia beginning ca. 2 Ma (430, 433), followed by a subsequent expansion ca. 800 ka (434). Meat acquisition through hunting or scavenging, a widespread human trait, has been proposed as one driver of this process (430, 435), as has a transition to more fat-rich dietary resources (e.g., bone marrow and brain tissue) (436), and more efficient food preparation skills, such as cooking (430, 437).

Pleistocene stone tools and cut marks on animal bones confirm that *Australopithecus afarensis* and - to an even greater degree - early *Homo* processed and consumed animal products (438–440), but the targeting of energy dense underground storage organs (USOs), such as roots, tubers, and bulbs, has also long been proposed as an increasingly important (albeit archaeologically invisible) food source during hominin evolution (437, 441, 442). Together with nuts and seeds, USOs are widely sought after by modern forager groups, and in most cases they represent a more reliable and consistent source of calories than hunting (443–446). Many of these plant resources are notably high in starch, an energy-dense food and a ready supply of the glucose needed to support a large brain. In addition, cooking and the controlled use of fire would have dramatically enhanced the nutrient availability - and thus quality - of many animal products (437, 447, 448) and especially USOs (437, 448), although it is not necessary to efficiently access fats (marrow, brains). Even very short forms of cooking, such as brief tuber roasting, provide benefits by easing mastication (449) or reducing toxins (450). As a form of external digestion, prolonged cooking denatures proteins and gelatinises starches, rendering both more accessible to human digestive enzymes (i.e., proteases and amylases) for hydrolysis and absorption (87, 442, 448, 451, 452). It has thus been argued that cooking, which emerged during the Middle Pleistocene, may have been the driving factor behind the final encephalization leap observed in later *Homo erectus* and other 'archaic' *Homo* species (453).

Dating the origins of cooking necessarily requires a good understanding of when hominins first began controlling fire (454). Clear evidence for the controlled use of fire, including hearths and pyrotechnic hafting technologies, has been documented at Middle Pleistocene Neanderthal sites in Europe from approximately 400 ka onwards, and evidence of fire and cooking is frequently observed at Late Pleistocene Neanderthal and modern human sites in both Europe and Africa (455, 456). However, some cold climate Neanderthal sites lack evidence of fire, which to date remains unexplained (455). Although evidence of fire use is widespread after 400 ka, such finds do not likely indicate the earliest uses of fire. Sporadic evidence of fire has been found as early as 1.5-1.0 Ma at Oldowan and Acheulean sites such as Chesowaja in Kenya (457) and Wonderwerk Cave in South Africa (458), and strong evidence for the controlled fire usage has been documented as early as 780 ka at the Acheulean site of Gesher Benot Ya'aqov (GBY) in Israel (459), where palaeoethnobotanical analysis has also revealed a complex foraging strategy that included more than four dozen species of seeds, fruits, nuts, vegetables, and plants producing underground storage organs (460). Based on the inferred digestive abilities of lower Pleistocene *Homo* as well as these and other finds, an early date for regular fire use and cooking by *Homo* has been proposed, beginning approximately 2 Ma and correlating with the encephalization and transcontinental expansion of Middle Pleistocene *Homo* species (i.e., *Homo erectus*) (453). Although there is not yet consensus within the field about the exact timing of the origins of cooking - whether it became widespread 2 Ma, 750 ka, or 400 ka - it is clear that the technology predates *Homo sapiens* and was likely widely, although perhaps not uniformly, practiced by Neanderthals and other similarly aged Middle Pleistocene 'archaic' *Homo* species (454).

Cooking has been shown to dramatically improve the digestibility and nutrient bioaccessibility of both animal and plant foods (442), and the inherent palatability of cooked foods is demonstrated by the general preference of great apes for cooked rather than raw meats and tubers (461). Although *Homo* meat and fat consumption can be inferred from the abundant evidence of cut-marked and fractured animal bones in the archaeological record, plants preserve less well and have likely been greatly underestimated at Palaeolithic sites (87, 462–464). Recent technological improvements in macrobotanical recovery (flotation) and the emergence of new methods for recovering microbotanical remains (pollen, phytoliths, and starch granules), however, are greatly improving the recovery and identification of specific plants and plant classes consumed by Middle and Late Pleistocene *Homo (85, 87)*. For example, in addition to the palaeoethnobotanical studies conducted at GBY, water flotation of Middle Palaeolithic sediments from the Neanderthal site at Kebara Cave in Israel produced more than 4,000 charred seed and fruit fragments belonging to at least 48 different species of plants, including legumes, acorns, pistachios, and a variety of taxa that produce edible USOs (83). Wild olives and pine nuts have been recovered at the presumed Neanderthal site of Gorham's Cave in Gibraltar (465), and tree fruits have been found in Neanderthal layers at Douara Cave in Syria and at Amud Cave in Israel (87). Grass seed phytoliths, plant starches, and other plant microfossils have been identified in sediments and on artifacts at Neanderthal sites in Israel, the Crimea, and France (84, 464, 466), and morphologically diagnostic date

palm phytoliths and starch granules from grass seeds (Triticeae, Andropogoneae), legumes (Fabaceae), USOs, and other plants have been identified in Neanderthal dental calculus from more than a dozen sites in the Near East and throughout Europe (85, 86, 88, 467). Some of these starches also show evidence of deformation that is consistent with thermal processing, and specifically heating in the presence of water (467), as is characteristic of cooking techniques such as roasting/baking or boiling (451, 468, 469), although admittedly some of these morphological changes may also derive from taphonomic processes (470).

Evidence for plant consumption and cooking at modern human sites is widespread (85, 117, 118, 471), and wild plant foraging, with a focus on starch-rich plants that were further improved by cooking, was probably a general feature of early *Homo sapiens*. In addition, recent experimental work in modern humans suggests that cooking is a necessary biological trait of modern humans, meaning that modern humans are unable to maintain reproductive fitness (regular cycles of ovulation and pregnancy success) on a diet of raw wild foods (437, 454, 472). When such a dependence on cooked foods emerged, and whether it is specific to *Homo sapiens* or also extended to Neanderthals, Denisovans, or other 'archaic' *Homo*, is unknown - and possibly unknowable.

It is clear from the growing body of evidence on Middle and Late Pleistocene *Homo* diets that both cooking and starch consumption likely played important roles in *Homo* evolution, although perhaps to different degrees. A major selective sweep and copy number expansion of the salivary α-amylase gene *AMY1* in *Homo sapiens* (up to 30 diploid copies) has been interpreted by most researchers as clear evidence for an increasing reliance on starch-rich foods (USOs and seeds) by modern humans (46, 94, 425), and this expansion has been dated using genetic modeling to just after the split ca. 550-765 ka between *Homo sapiens* and the *Homo* branch leading to Neanderthals and Denisovans (94). Although some have speculated that this salivary α-amylase copy number expansion in modern humans had a non-dietary cause (155), as the enzyme also plays minor roles in reducing tannin condensation (36, 473, 474) and stress response (35, 475), the convergent evolution of pancreatic β-amylase copy number expansion in domesticated dogs (407, 408) makes this argument doubtful. Moreover, salivary α-amylase has independently evolved from pancreatic β-amylase in multiple mammal groups (e.g., primates and rodents), and amylase genes have repeatedly undergone copy number expansions in animals consuming starch-rich diets (e.g., humans, baboons, mice, rats, pigs, and dogs) (34).

There is general consensus that *AMY1* copy number expansion in humans is a response to a starch-rich diet, and high levels of α-amylase are expressed by modern humans, where it is the most abundant protein in saliva (151). In the oral cavity, α-amylase contributes to the partial hydrolysis of the two major constituents of starch (amylose and amylopectin), and it also contributes to taste perception (476), including increased sweetness sensation through minor maltose perception (477) and altered starch viscosity and mouth feel (478). Cooking has been found to further increase the hydrolysis rate of α-amylase, and this effect scales with both α-amylase expression levels and length of mastication (479). Interestingly, α-amylase is expressed in modern human saliva from birth, whereas expression of pancreatic β-amylase is low in modern human infants and only reaches adult levels around 16 months of age (480, 481). Although the period of breastfeeding in modern humans is extended, lasting on average more than 2 years in modern forager societies, plant foods are nevertheless generally introduced within the first year (482, 483). Enhanced salivary α-amylase production may thus be advantageous during the early stages of plant food introduction in young children's diets, especially until pancreatic β-amylase reaches adult levels. This may complement salivary α-amylase provided by adults through pre-mastication, a widespread practice performed by adults to soften foods for young children before they have sufficient dental development to efficiently chew foods on their own (484).

Beyond salivary α-amylase copy number variation, other loci in the modern human genome hint at early diet- mediated evolutionary changes within the genus *Homo*, such as the pseudogenization of the masticatory myosin gene (*MYH16*), two bitter taste receptor genes (*TAS2R62* and *TAS2R64*), the major urinary protein genes (*MUP*), and the enzyme CMP-sialic acid hydroxylase (*CMAH*), all of which are involved in food perception, mastication, or digestion (402, 485). In addition, seven additional genes found to have differential expression patterns in mice fed cooked and raw diets show evidence of recent selection in humans and all were found to have changed between the chimp-*Homo* split and not between

Neanderthal/Denisova and modern *Homo (485)*. Relatively few diet-related genetic changes are known to have occurred in the modern human lineage after the split with the branch leading to Neanderthals and Denisovans; however, these include a functional change in the aryl hydrocarbon receptor gene (*AHR*) that alters the toxicity of cooking-related polycyclic aromatic hydrocarbons (PAHs) (486). Overall, timing the adaptation of *Homo* to dietary starches using genetic data is challenging, in part because the evolution of mammalian genomes is slow, many genes have inadequately understood functions, and polygenic phenotypic traits are generally poorly understood (487–490).

Commensal microbes of the modern human microbiome represent an underutilised and independent source of information about host evolutionary changes (491, 492). With generation times orders of magnitude shorter than their hosts and the ability to acquire new functions through horizontal gene transfer across distantly related groups, microbes represent a particularly dynamic and temporally resolved system for investigating changes in host behavior. The finding made here of an expansion of *Streptococcus* groups capable of binding and utilizing salivary amylase in both Neanderthals and modern humans supports hypotheses for an early importance of dietary starches in *Homo* evolution. Moreover, because some *Streptococcus* amylase-binding proteins, such as AbpA, are only robustly expressed in the presence of *both* salivary α-amylase and dietary starch (395, 418), it again emphasises the connection between these microbial genes and host diet, and not merely to host stress response or other minor functions of α-amylase. In addition to their dietary function, amylase-binding proteins also facilitate *Streptococcus* colonization of the oral biofilm (either through direct binding to the AEP, as in AbpA, or through competitive advantage, as in AbpB) (419), and this may explain why *Streptococcus* has become the most abundant genus in modern human oral biofilms, whereas it is one of the least abundant core genera in the chimpanzee oral biofilm.

Our finding that amylase-binding proteins are present in *Homo* (Neanderthals and modern humans), but not other African hominids nor the howler monkey outgroup, and that the *Streptococcus* clades in which these genes are found have greatly expanded in abundance in *Homo* compared to other taxa, suggest that dietary starches played an important and early role in *Hom*o evolution prior to the split between Neanderthals and modern humans. Moreover, because the period during which starches are present in the oral cavity is transient (with only small amounts of starch persisting after the swallowing of the food bolus), salivary α-amylase would be of much greater benefit to resident oral microbes if starches were consumed in an already gelatinised (i.e., cooked) form, which makes them much more susceptible to rapid hydrolysis by α-amylase (479). Thus, selection for amylase-binding proteins, and the selective advantage of microbial clades containing these genes within the oral biofilm, likely intensified after cooking became habitual in *Homo*.

Considering each amylase-binding protein separately provides further evolutionary insights. Of the two best studied amylase-binding proteins, AbpA and AbpB, it is interesting to note that AbpB shares sequence similarity with other oral *Streptococcus* dipeptidases (including from *S. gordonii*) (420) and thus likely evolved from pre-existing genes in this genus, whereas AbpA bears no homology to other known bacterial proteins (392, 419), and thus likely emerged within modern human oral biofilms through horizontal gene transfer from an as yet unknown source (but possibly related to *Gemella haemolysans* M341 (392)). We find that *abpB*-like genetic sequences are present in both Neanderthal and modern human oral biofilms, while *abpA* is only detectable at appreciable levels in modern humans. The possible later evolution of *abpA*, which directly participates in AEP binding and physically docks α-amylase to the bacterial cell surface for nutrient harvest (414, 418), may signal an even greater reliance on cooked dietary starches by modern humans compared to other members of *Homo*, a dietary change that is further supported by the modern human-specific copy number expansion of *AMY1*. Thus, evolutionary and ecological changes in the oral microbiome of African hominids suggest that dietary starches first became important during the early evolution of *Homo*, followed by an even greater use of cooked dietary starches by modern humans. To refine the chronology and tempo of this process, future focused research on the *abpA* and *abpB* genes in *Homo* using capture enrichment technologies is needed to improve gene coverage depth and breadth.

S5.9 Dating expansion of starch-associated streptococci using amylase binding proteins

To further explore if commensal oral bacterial evolution mirrors human evolution, we examined the presence and distribution of *abpA* and *abpB* in our samples to determine if we had sufficient data to observe an expansion of the gene through time. We downloaded the nucleotide sequences of all protein sequences listed in the supplementary tables of (392) for *abpA* (15 sequences) and *abpB* (14 sequences) and used these as references. To capture additional sequence variation in these two genes, we performed a BLASTn search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) of each reference gene against a database of all *Streptococcus* genomes downloaded from NCBI RefSeq in May 2018. The exceptions were *S. pneumoniae*, *S. suis*, *S. agalactiae*, *S. pyogenes*, and *S. equi*, which contained, respectively, 8,093, 1,197, 911, 338, and 279 genomes, and each was randomly subsampled down to 50 genomes because amylase-binding activity has not been demonstrated in these species. All sequences that had a query coverage of 100% were selected and compared, and only unique sequences were selected for further analysis. In total, we found 41 unique *abpA* sequences and 47 unique *abpB* sequences (including the reference sequences). We built these into a maximum likelihood tree using Geneious v8.0.5 (https://www.geneious.com) and selected one representative of each clade to use for our BLAST search and mapping (see External Data Repository Section R12.2.3 and File R36).

We searched for *abpA* and *abpB* in the shallow sequencing dataset as well as a published set of historic calculus samples from the Radcliffe Infirmary in England (2), and the Human Microbiome Project plaque samples (191) , independently with a 2-step BLAST search followed by a mapping approach, and further by additional mapping with alternative parameters to confirm our findings, as described below. Two BLAST databases were built, one with each reference sequence of *abpA* and the other with each reference sequence of *abpB*, and we performed a nucleotide BLAST (493, 494) (v. 2.7.1+) search of all of our shallow sequencing data, as well as the Radcliffe dental calculus dataset, against both databases. Parameters are provided in the External Data Repository Section R12.2.3.

Reads in each sample that hit any amylase-binding protein reference sequence were selected out of the original sequence file and placed into a specific fasta file, one for *abpA* and one for *apbB*. Next, each reference sequence was indexed with bwa (v0.7.12-r1039), and the subsetted fasta file of reads identified by BLAST was mapped against each reference sequence. Duplicate reads were removed with samtools, and alignments with greater than 10 mapped reads were visually inspected for coverage, read distribution, and mis-matches characteristic of aDNA damage. To confirm the specificity of the mapping, and to ensure that we had not missed more distantly related sequences, we performed additional mapping of the shallow sequencing data using much less strict mapping parameters without the preliminary BLAST search. Parameters are provided in the External Data Repository Section R12.2.3.

For each sample mapped with the parameters above, we counted the number of reads that mapped and visually inspected the alignments to determine if they contained mis-matches characteristic of ancient DNA damage. For all calculus samples, no more than 3 additional reads mapped compared to the initial mapping with strict parameters. This suggests that we were not missing highly diverged sequences in our samples, and that the samples with no reads mapping to either gene did not have the gene, or that the abundance was too low to be detected from the depth of sequencing performed in this study.

S5.9.1 abpA

The only reference gene sequence to which ancient and historic samples mapped with greater than 10 reads was from *S. gordonii* str. Challis substr. CH1, and only alignments to this reference were assessed going forward. Visual inspection of the mapping with IGV (495) (v. 2.4.14) showed even read distribution and expected mismatches at the ends of mapped reads typical of aDNA damage patterns. Consensus sequences were exported from IGV for samples that had at least 40% of the gene covered at least 1X, thirty-six in total. Of the 36 samples with evidence of *abpA*, 4 were pre-agricultural modern humans, 8 were pre-antibiotic modern humans (7 from the Radcliffe data), and the remaining were present-day calculus (10 total) and dental plaque (14 total). An alignment of the consensus sequences as well as the reference sequence were used to generate an xml file in BEAUTi for Bayesian skyline analysis with

BEAST2 v2.5 (496, 497). All parameters were set at default with the GTR (498) substitution model, a strict clock, and coalescent Bayesian skyline model (499). The MCMC chain length was 800000000 with sampling every 80000 states. The chain converged but inspection of the tree structure in DensiTree showed no structure. For this reason, we did not proceed with BEAST analysis to date the expansion of *abpA*. Insufficient sample size, unequal distribution of samples (only 12 historic samples vs. 24 present-day calculus and plaque) and a short gene length may be responsible for this. Future work will focus on increasing sample size and coverage, as well as inclusion of contiguous *srtB* sequences.

S5.9.2 abpB

Reads from ancient and historic samples mapped to five reference sequences of abpB, with at least 10 sequences per sample. Four of these references were from strains of *S. gordonii* and the fifth an unnamed species isolate. The reads that mapped to each of these references were checked and found to be consistent (i.e., the same set of reads mapped to each abpB reference sequence). We therefore chose to work with only alignments to reference *S. gordonii* str. Challis substr. CH1 going forward, to be consistent with the *abpA* analysis. Visual inspection of the mapping with IGV showed an even read distribution and expected mismatches at the ends of mapped reads typical of aDNA damage patterns. Consensus sequences were exported from IGV for samples that had at least 40% of the gene covered at least 1X, fifty-three in total. Of the 53 samples with evidence of *abpB*, 4 were pre-agricultural modern humans, 14 were pre-antibiotic modern humans (12 from the Radcliffe data), and the remaining were present-day calculus (15 total) and dental plaque (19 total). Although Neanderthals had a relatively high number of *abpB* reads, they fell just below this coverage threshold.

The alignment of the consensus sequences as well as the reference sequence were used to generate an xml file in BEAUTi for Bayesian skyline analysis with BEAST2 v2.5. All parameters were set at default for the GTR substitution model, a strict clock, and coalescent Bayesian skyline model. The MCMC chain length was 50000000 with sampling every 5000 states. The chain converged and inspection of the tree structure in DensiTree showed structure where the oldest samples were more basal than the present-day calculus and plaque. External Data Repository Section R12.2.3 Fig. R38 shows the Bayesian skyline plot of population expansion in *abpB*. However, we chose not to pursue dating the expansion of *abpB* with the data we have because of small sample sizes in older samples (only 4 pre-agricultural samples) and uneven distribution of samples (18 pre-modern calculus vs. 34 present-day calculus and plaque), as well as the additional peptidase activity of this gene that may have evolutionary pressures independent of amylase-binding activity. Therefore, we do not draw any conclusions from this plot, but suggest that future work with more samples and better temporal representation may be able to determine whether there has been a population expansion of the amylase-binding protein genes in *Streptococcus,* that coincides with periods of proposed increased starch consumption.

S6. Microbial phylogenetics

Given that the core microbiomes of African hominids are distinct at a compositional level, we wanted to test whether individual genomes of oral bacterial species reflect the evolutionary phylogeny of their hosts. For this we selected shared species in the core microbiomes of African hominids and the howler monkey outgroup (see section S5.4).

A major challenge in analyzing microbial ancient DNA is identifying short reads belonging to a single microbial strain against a highly diverse microbial background of both other endogenous related species and environmental contaminants (203). Cross-mapping from these other species to the reference genome of interest frequently results in large numbers of multi-allelic positions (sometimes referred to as ‘heterozygous positions’) in what should otherwise be haploid genomes. This makes SNP calling for phylogenetic analysis more difficult, as it is not clear which ‘true’ allele comes from which strain, and thus can confound calculating genetic distances and mutation rates between genomes.

Recent methods developed for this problem typically perform strain separation post-SNP calling, and these typically require coverage depths of ~10-20X to separate strains, as well as well-characterised comparative databases of strain diversity within the species (500–503). In contrast, the highly fragmented

nature of ancient DNA and reduced proportion of the original microbiome of ancient and historic samples (against a larger background of environmental contamination), often requires a prohibitive sequencing effort to reach sufficient genomic coverage for the application of these methods. Furthermore, many of the commensal taxa that make up the bulk of the original microbiome do not have sufficient comparative genomes available in existing databases, unless they are of clinical interest. This makes frequency based SNP profiling very difficult due to the inability to perform probabilistic tests based on frequency and genotyping quality.

In an attempt to mitigate the effects of cross-mapping from related taxa, we explored an approach to reduce multi-allelic positions prior to SNP calling. The procedure consisted of simultaneously mapping to a 'super-reference' genome that included one genome of each species of a given genus (allowing reads mapping to multiple places to be placed randomly), extracting mapped reads to the single genome of interest (typically the one with highest coverage across the super-reference), and running SNP calling on this. As reads from other species would map to the regions in those other species, theoretically, the remaining reads on the genome of interest should be most likely derived from that given species or a non-represented close relative. In cases where a read could still map equally to two places on the reference, we utilised the behaviour of bwa to randomly assign to one of the sites, thereby spreading the generic 'baseline' or 'background noise' of the shared genomic regions across all species. Explained another way, when using this approach, a read from a different species that could still reasonably map to the species of interest when mapping solely to that reference genome, would more likely be assigned to the correct genome or to another closely-related taxon, than be falsely assigned to the species of interest.

To improve chances of obtaining reasonable genome coverages while reducing false-positive SNP calls, all data used for building phylogenies are from the deep sequencing data set of well-preserved calculus samples. These samples were selected by mapping the shallow sequencing dataset to a range of known oral microbes of interest that were prevalent across individuals. We then estimated for each species and individual the number of reads required to reach 5X coverage. Due to the low number of PCR duplicates, we were unable to adequately estimate the necessary sequencing using tools such as PreSeq (504). Therefore, we linearly extrapolated the sequencing depth by multiplying the number of mapped reads after deduplication by the target coverage divided by the current coverage. We then filtered out any sample-target pairs that would require more than 100 million reads of additional sequencing to reach 5X coverage, as well as samples with cluster factors (also known as duplication factor: mapped reads before deduplication / mapped reads after deduplication) above 1.2. After filtering, for most individuals, only one or two target genomes were projected to reach the desired depth. We therefore selected the samples for which three targets could be sequenced to the desired 5X coverage with the fewest number of additionally sequenced reads. We verified that the remaining samples passed the taxonomic preservation filtering above (section S3.4). This procedure is also described in External Data Repository Section R11.1.

S6.1 Super-reference construction

We downloaded the NCBI 'assembly_summary_genbank.txt' file from ftp.ncbi.nlm.nih.gov/genomes/ASSEMBLY_REPORTS/assembly_summary_genbank.txt on 2018-11-26 and the Batch Genomic Meta Information file from the Human Oral Microbiome Database (HOMD) (220) on 2019-05-16 (for the addition of *Campylobacter* and *Ottowia*). Both files were cleaned, standardised and merged in R using tidyverse packages. The genomes were filtered to include only those corresponding to 'core' genera in the *Alouatta:Gorilla:Pan:Homo*, *Gorilla:Pan:Homo*, and *Pan:Homo* calculus microbiomes (excluding *Mycobacterium*, see section S5.4). In the case of the genus *Pseudopropionibacterium* (soon to be renamed back to *Arachnia (380)*), additional closely related genera were also included due the recent reorganization of this group to closely related skin (*Cutibacterium*, *Acidipropionibacterium* and *Propionibacterium*) and oral (*Pseudopropionibacterium*) taxa (505). For each genus, a single representative assembly per unique species name was selected. In cases of multiple candidates, a selection hierarchy was applied to ensure inclusion of better quality assemblies on the basis of various NCBI metadata criteria:

1. Assembly level: Complete Genome > Chromosome > Scaffold > Contig

2. Refseq category: reference genome > representative genome
3. Sequencing status: Complete > High Coverage > Survey

In cases of multiple candidates, one was randomly chosen using the 'sample_n' function from the R package dplyr. Genomes labelled 'sp. UBA', 'sp. HMSC' and 'sp. CAG', were removed because many of these appear to be unvalidated 'metagenomic-assembled-genomes' (MAGS), and appear to be unnamed duplicates of already published assemblies. Next, for each assembly we used the NCBI entrez package to query whether an isolation source was listed and/or whether host metadata was available. Such data were downloaded and again loaded in R. Where possible, unique isolation source entries were then visually checked for obvious non-host related isolation sources such as 'sediment', 'seawater', 'crude oil', etc. If found, such genomes were removed because these assemblies would likely correspond to highly divergent relatives of mammalian-associated taxa. A full list of the assemblies that were selected are provided in External Data Repository Section R11.2 and File R26. Associated metadata tags that were excluded are listed in the R notebook described in the External Data Repository Section R11.2.

The final list of selected genomes per genus were then downloaded in FASTA and GFF formats using the links supplied in the External Data Repository Section R11.2 and File R26. For downstream processing, FASTA headers were renamed to include the species name in addition to the NCBI accession ID. Additionally, all genomes for a given genus were concatenated into a single contig, and coordinates were recorded in a separate bed file using as described in External Data Repository Section R11.2. The resulting FASTA file was then indexed with bwa index, samtools faidx, and picard CreateSequenceDictionary using the corresponding version in the EAGER pipeline (see above, section S3.2).

S6.2 Alignment and species selection

The non-human reads available in our deep sequencing libraries (see section S3.2) were next aligned to each super-reference independently using the EAGER pipeline with the settings described in External Data Repository Section R11.3 and File R27. Once completed, bedtools coverage (v2.27.1) was run on each BAM file in order to generate the breadth (default) and depth (with -mean flag) coverage statistics for each assembly based on the corresponding BED file generated during fasta collapsing (see above).

The resulting bedtools results files were loaded into R and combined together. A variety of mapping metrics were then calculated per genome.

1. Total number of species-aligned reads (multi-contig assemblies were summed)
2. Total number of bases covered (multi-contig assemblies were summed)
3. Mean depth coverage (multi-contig assemblies were averaged)
4. Percent species-aligned over all genus-aligned reads (calculated on summed data)
5. A competitive mapping score (calculated on summed data, species-aligned reads - genus-aligned reads / sequencing depth; modified from Andrades Valtueña et al. (506))

The mapping results for the super-reference are provided in External Data Repository Section R11.3 and File R28.

For phylogenetic analysis, we wished to select species that were prevalent across as many individuals as possible, but which also had enough data for sufficient SNP calling (i.e., depth and breadth coverage). This would then allow the most optimal balance on having enough comparative data across the host genera to explore relationships, but with a higher confidence in SNP calling. We applied two methods of species selection - an 'automated' selection procedure and a 'visual inspection' as a sanity check.

For the 'automated' selection, for each genus the average and standard deviation of each metric was calculated. The standard deviation was used as a threshold to indicate when the metric exceeded the 'noise' generated by the random assignment of reads that were able to equally map to multiple places (see above). Assemblies that did not reach the mean metric + standard deviation filter, were removed. Next, the number of times each reference had passed the filter for each metric was summed, and those

only passing a single metric were removed (External Data Repository Section R11.4 and Table R9). In most cases this led to a clear candidate - i.e., a single reference that passed most of the metrics (accounting for breadth and depth). In cases where there were multiple possible candidates, either a named species (i.e., laboratory-isolated) was selected, or the more complete assembly was selected. In the case of *Campylobacter, C. gracilis* was chosen over *C.* sp. AAUH-44UCsig-a, as *C. gracilis* is named and known to be isolated from the oral cavity. For *Ottowia* however, *O.* sp. oral taxon 894 was selected over *O.* sp. Marseille-P4747 because the latter was isolated from stool rather than the oral cavity.

For the visual inspection, the metric results were then displayed as heatmaps and species were selected based on the two criteria described above. Furthermore, the consistency of the results across metrics were checked with comparison. We noted that in some of the microbial genera, the mean-depth metric was not consistent with the other methods. This often appeared to be due to the presence of uncharacterised species assemblies (i.e. 'sp.') that had very short assembly lengths compared to other taxa. For example, in all metrics for *Pseudopropionicum* except mean depth, *Pseudopropionibacterium propionicum* had the greatest prevalence and metric score. However, for mean depth, *Propionibacterium* sp. JGI 0001002-M06 had greater prevalence, but this assembly only had a length of 44,239 bp (https://www.ncbi.nlm.nih.gov/assembly/GCA_000463995.1/), which is unusually short compared to the other *Propionibacterium* related references in the dataset (mean: 2,466,726 ± 670,480 bp). Therefore, this metric was excluded from further consideration during visual inspection. In certain cases, it was not possible to decide between different species that had similar abundances and/or prevalences using visual inspection. These were instead marked as 'unsure'. Comparison between the two methods showed a general concordance in the selected species (External Data Repository Section R11.4.1 and File R29), with 12 of 15 giving the same results. *Selenomonas*, *Fusobacterium* and *Corynebacterium* did not have a clear candidate from visual inspection, and the automated selection species was selected for downstream analysis and vice versa. The species selection procedure is outlined in detail in the External Data Repository Section R11.4.

S6.3 Performance of super-reference vs. single reference mapping

Once selected, the deep sequencing dataset was re-mapped with the same settings as the super-reference mappings, but to each corresponding reference genome alone. Summary statistics for these mappings can be seen in External Data Repository Section R11.4.1 and R32. A summary of differences between the two mapping strategies in terms of depth of coverage, breadth of coverage, and number of mapped reads is provided in External Data Repository Section R11.4.1 and File R30. Because the objective was to compare the extent cross-mapping (resulting in 'multi-allelic' SNPs) between the two methods, MultiVCFAnalyzer (v0.87, https://github.com/alexherbig/MultiVCFAnalyzer) (507) was used to perform SNP calling and generate single-allelic position statistics. Detailed information regarding SNP calling using this tool is provided in the External Data Repository Section R11.5.

The percentage of SNPs called as single-allelic vs. multi-allelic were then calculated and visualised in a modified manner following the strategy of Zolfo et al. (508) (see Fig. S10; External Data Repository Section R11.5, Fig. R62 and File R31). Overall, the percentage of multi-allelic sites was reduced by a mean of 5.19% when using the super-reference mapping strategy compared to the single genome mapping strategy; however, the differences ranged from an increase of 32.10% (in *Capnocytophaga gingivalis)* to a reduction of 43.55% (in *Prevotella*), highlighting a large amount of variability across both samples and reference species. In most cases, the two mapping strategies resulted in either few differences in the number of multi-allelic sites, or a reduction in multi-allelic sites but with a large corresponding reduction in the number of overall usable SNPs called for downstream analysis (likely due to a reduction in coverage, for example in *Selenomonas*). Note that *Fretibacterium* is not included in this assessment because it has only a single representative genome in the super-reference, so any difference is simply an artefact of the alignment method. In some cases, such as in *Capnocytophaga gingivalis* in individual DJA002, using the super-reference mapping strategy increases the number of multi-allelic SNPs. This is likely due to a reduction in coverage at sites in conserved regions (which are randomly distributed elsewhere when using the super-reference), thus reducing the 'denominator' of number of reads (or possible nucleotides) covering the site when calculating the proportions of alternate nucleotides.

The very modest outcome of our attempts to reduce cross-mapping using a super-reference strategy highlights the difficulties of strain separation from low-coverage ancient DNA data. We conclude that upstream methods of microbial species enrichment (e.g., much deeper sequencing) are required to generate sufficient genome depth coverage for adequate strain-separation. In the context of ancient pathogen research, recent in-solution capture techniques have been successfully applied for this purpose (207, 506, 509). Given that super-reference mapping did not markedly improve the quality of our data, we decided to focus on single reference genome mapping for downstream analyses.

S6.4 Single reference genome mapping statistics

Across individuals, single reference genome mapping of the selected core microbiome taxa resulted in a mean depth coverage (after duplicate removal) ranging from 39.0x for *Actinomyces dentalis*, to 0.6x for *Selenomonas* sp. F0473. The distribution of mean fold coverage in deeply sequenced calculus samples is provided in External Data Repository Section R11.4.1 and Fig. R29A. The highest coverage observed was for *Ottowia sp.* oral taxon 894 in present-day modern human JAE008, with a mean depth of coverage of 219x and 91% of bases covered at ≥1x. Of the non-present day individuals, the highest coverage was observed for *Actinomyces dentalis* DSM 19115 in chimpanzee KNP004, which had a mean depth coverage of 177x and 86% of bases covered at ≥1x. The highest covered genome for a Neanderthal was *Pseudopropionibacterium propionicum* F0230a in GOY005, which had a mean depth coverage of 55.5x and 95.7% of bases covered at ≥1x. While the deep sequencing data resulted in some high coverage ancient genomes, of the 285 mappings from ancient and historic samples to the 15 reference genomes, only a third exceeded a depth of coverage of 3x, 17% exceeded 10x, and 8% exceeded 20x. However, the average cluster factor (or PCR duplicate rate) remained low at 1.2, with a maximum of 3.4, and 89% of all mappings had a cluster factor less than 1.5 (External Data Repository Section R11.4.1 and Fig. R29b), which indicates that complexity in these libraries is sufficient for even deeper sequencing and thus higher genome coverage could be achieved with greater sequencing effort.

The low cluster factors generally observed in our samples highlight the genetic richness of the dental calculus in this study, whereby even after deep sequencing, further sequencing is still possible without a reduction in the number of unique reads being gained. For example, in the individual with the highest number of non-human reads in this study, Neanderthal GOY005 (non-human reads: 326,113,634), the highest coverage genome (see above), still had a cluster factor of only 1.4. The extremely high complexity of this sample draws attention to the need for developing enrichment techniques for sequencing target species of interest in ancient dental calculus, that go beyond the simple brute force method of deeper sequencing. Final mapping statistics and visualization for those used for downstream variant calling are provided in External Data Repository Section R11.4.1 and File R32.

S6.5 Variant calling and single-allelic position assessment

Due to the generally low genome coverage and high levels of cross-mapping, we opted to take a ‘low-resolution’ approach to understanding the phylogenetic relationships between individuals based on mappings to single microbial genomes. SNP calling at low coverage is made complicated by the presence of cross-mapping from other taxa; estimating the ‘true’ SNP at each position carries a high false positive risk - e.g. two reads out of three containing the alternative allele does not provide high confidence in establishing the genotype. Furthermore, imputation of very low to no coverage regions of the genome through haplotype calling (phasing) is not possible due to the very short and fragmented nature of aDNA. Typically, few aDNA reads will span two or more potential SNP sites, which is necessary to confirm a relationship between the two sites.

In light of these challenges, we reasoned that we could instead assume that each mapping represents a sampling of a ‘population’ of related taxa (rather than a single strain), with the given reference as a representative for that population. Genetic distances between these populations could still be informative when looking at evolutionary distances between the host genera, with the caveat that mutation rates for dating estimation would not be possible as the assumption of relationships between all SNPs undergoing a similar evolutionary processes would not be met. For this analysis, we did not consider indels due to large variation in the breadth of coverage across the mappings.

To account for the low coverage, we again used MultiVCFAnalyzer as above, but changed the minimum coverage to 2X, and the 'homozygous' and 'heterozygous' minimum frequency both to 0.7. The selection of the minimum frequency parameter was performed by finding the largest majority peak (i.e., >50%) in the distribution of multi-allelic SNP frequencies (see External Data Repository Section R11.5), and was generally consistent across most of the mappings and core microbial species used as references (External Data Repository Section R11.5 and Fig. R30). The only exception was *Actinomyces,* which did not have a clear peak, and therefore 0.7 was applied here for consistency with the rest of the dataset. This meant that at the minimum accepted coverage (2x), SNPs were only called when they had a minimum of 2 independent reads supporting the call, and at 4x if three independent reads supported this call. These parameters therefore provide a reasonable level of confidence that the SNP is 'abundant' in some form within the microbial population represented by the mapping. The resulting SNP alignment from MultiVCFAnalyzer was then used for phylogenetic analysis. The resulting FASTA files are provided in External Data Repository Section R11.6.

S6.6 Phylogenetic trees

We loaded the SNP alignment files into R using the 'fasta2DNAbin' function in the adegenet package (v2.1.1) (510, 511). In order to retain sufficient information and overlap with other mappings in the alignment, we removed from the alignment any mapping that had less than 1000 called positions for each microbial genus. To maximise the number of positions used to calculate distances between each pair of the generally low-coverage mappings, we then calculated genetic distances through pairwise-deletion of the alignments, using the Jukes-Cantor 69 nucleotide evolutionary model (512) in the 'dist.dna' function from the R package ape (v5.3) (241). Across all SNP alignments for the fifteen representative taxa, the number of pairwise overlapping SNPs that informed the genetic distances between individuals ranged from a low of 29 (Tannerella: PES001 & OME002) to 246,305 (*Pseudopropionibacterium*: VLC004 & VLC009). As expected, higher numbers of pairwise overlapping bases were observed for genera with higher median depths of coverage (External Data Repository Section R11.7 and Fig. R32). However, *Capnocytophaga* also showed a large number of overlapping positions despite having a low median average fold depth across these mappings. We believe this may be due to the use of a suboptimal representative genome (due to the unavailability of more suitable genomes), and that these overlapping positions largely derive from highly conserved region(s) shared across taxa and individuals. Neighbour joining clustering of genetic distances was carried out with the 'nj' function from ape. The resulting tree was then bootstrapped with 100 resamples via ape's 'boot.phylo' function to estimate the confidence in the calculated bipartitions. Detailed information for the phylogeny creation procedure is provided in the External Data Repository Section R11.7.

In the case of *Porphyromonas*, trees failed to be generated or bootstrapped due to problems with distance calculations. This was caused by highly divergent strains in one or more samples, which violated the JC69 model assumption of equal base composition across samples during bootstrapping. Inspection of the pairwise mismatch versus match position counts revealed that the *Porphyromonas* reference genome mapping for OME003 had a very high number of mismatches. During bootstrapping, this sample would lead to a base composition difference with the remaining samples, and cause the distance calculation to fail. This individual was subsequently excluded, and *Porphyromonas* trees were generated without the individual. Furthermore, for *Porphyromonas* we increased the minimum required number of called positions to 2000 (rather than 1000) because for this taxon we found that a higher number of pairwise combinations of lower-coverage individuals contained no overlapping regions, which caused the 'dist.dna' function to fail as no distance could be calculated.

Trees (Figs. S11, S12 and External Data Repository Section R11.7 and Fig. R31) were rooted where possible on the *Alouatta* branch (when *Alouatta* individuals were monophyletic). As a New World monkey species that had diverged from apes 31-56 Ma, we assumed that *Alouatta* would be an appropriate outgroup for oral microbial strains/species. In cases where *Alouatta* did not have enough positions in the alignment to pass the minimum threshold, we performed a midpoint rooting of the tree. Due to varying robusticity in tree quality, possibly due to variability in the SNPs sampled from the microbial population (as seen by the high rates of multi-allelic SNPs in some microbial species and hosts), we then filtered for

trees where two deepest branches had bootstrap values of ≥70%. Details regarding the generation of and visualisation of trees is provided in External Data Repository Section R11.7.

Of the 8 phylogenetic trees with well-supported deep internal branching, *Homo* (n = 6) and *Alouatta* (n = 7) formed monophyletic clades in most of them. *Pan* also formed monophyletic clades for all but the *Pseudopropionibacterium* tree, and internal topologies were generally well-supported (≥ 75% other than in *Ottowia*). *Gorilla* microbial topology was more variable. *Tannerella* and *Fretibacterium* in *Gorilla*, for example, formed monophyletic clades falling sister to *Pan*, but for other microbial species one individual fell in an 'ancestral' position to the remaining *Pan* and *Gorilla* sister clades: MTM009 (4 trees) and EBO003 (1 tree) - both Western Lowland gorillas (*G. gorilla gorilla*). The difference in the level of monophyly between *Pan* and *Gorilla* microbes may be due to the *Pan* clade consisting of only individuals from the same forest region. In the *Streptococcus* tree, *Pan* and *Gorilla* each only yielded one individual with a sufficient number of positions to be considered. In most microbial trees, *Homo* tended to form two clades - one containing modern humans and another containing Neanderthals, with the latter also often including one or two ancient/historic modern humans in different combinations. The internal topology of the Neanderthal clade was generally well-supported in most trees (≥ 90% bootstrap support, other than with *Fretibacterium*). However, for *Ottowia* and *Pseudopropionibacterium*, the two *Homo* clades did not form sister monophyletic taxa, but rather fell basal to mixed-host species clades containing *Pan* and *Gorilla*. Within modern humans, high strain diversity (i.e., cross-mapping) observed for these two poorly studied microbial genera may be responsible for this pattern (see Fig. S10 or External Data Repository Section R11.5 and Fig. R62).

In analyzing the tree topologies, we noticed an interesting pattern relating to an Upper Palaeolithic individual from El Mirón from Spain directly dated to ~18.6 ka (513) (EMN001; also known as the 'Red Lady'). In every tree, she fell either immediately ancestral to or within the Neanderthal *Homo* clade. This individual was the only deeply sequenced Upper Palaeolithic European included in the phylogenetic analysis, and we found it noteworthy that all Upper Paleolithic *Homo* individuals consistently clustered together, while the majority of Holocene and later individuals tended to cluster with present-day modern humans (with some variation). Although EMN001 is substantially younger than the Neanderthals demise, it is known that modern humans and Neanderthals interbred (71, 128, 514), providing opportunities for Neanderthal oral bacterial strains to cross over to Upper Palaeolithic modern humans in Europe. Subsequent Upper Palaeolithic European groups underwent population turnover events (130, 131), and were ultimately largely replaced by the spread of Near Eastern hunter-gatherers during post-glacial warming (130, 131), Anatolian farmers during the Neolithic (142, 515), and Western steppe herders during the Bronze Age (516). During the Upper Palaeolithic, the Red Lady (EMN001) gives the type-name to the 'El Mirón cluster' of Magdalenian-related individuals who spread across large parts of Europe after the Last Glacial Maximum (ca. 25-19 ka). Prior to the arrival of farming, the El Mirón cluster was found to be largely (517) replaced at least from 14 ka by other hunter-gatherer individuals related to the so-called Villabruna cluster, who show a higher affinity to Near Eastern populations (130).

We hypothesise that the clustering patterns of EMN001 oral microbial strains may be a reflection of ancestral European Upper Palaeolithic oral microbial diversity, which may have included Neanderthal strains. This would explain the consistent clustering of EMN001 with Neanderthals, compared to contemporary non-European (TAF, Taforalt (518)) and later Mesolithic European individuals (ECO, El Collado (135)), who typically cluster with calculus samples from present-day individuals. At present, we do not have sufficient evidence to suggest that there is a host genetic relationship between European Neanderthals and EMN001; however, the pattern nevertheless raises interesting questions with respect to whether commensal microbial genomes may be useful for inferring past modern human population dynamics and migration events (see, for example, (519)). Additional data from well-preserved and deeply sequenced Upper Palaeolithic European modern humans are required to confirm the broader microbial relationships suggested by the individual from El Mirón.

Because large amounts of missing data and/or reference bias may result in tree artefacts such as 'long branch attraction' (an error that causes two highly diverged taxa to appear closely related), we checked that the clustering of EMN001 with Neanderthals was not due to an artefact of low coverage or reference bias. First, for each host genus, we compared the median number of positions that EMN001 shared with

each of the Neanderthals to the shared position distribution of all pairwise combinations of modern humans (External Data Repository Section R11.8.1 and Fig. R33). We find that the median number of shared bases between EMN001 and Neanderthals (dashed red line) does not fall outside the range of shared positions between EMN001 and other modern humans, but rather that it in fact generally falls near the EMN001-modern human median (solid orange line). Therefore, there does not appear to be a bias due to missing data; rather, the phylogenetic clustering of EMN001 with Neanderthals appears to be related to the sequence similarity of the microbes themselves.

We next considered the possibility of reference bias - i.e., that the microbial strains present in EMN001 and the Neanderthals could be different from those in the other modern humans in a way that makes them not directly comparable. However, the consistency of the pattern across all the trees makes this scenario unlikely, especially given that other ancient and historic individuals from both pre-agricultural and post-agricultural groups also sometimes cluster with this clade. Further information regarding data analysis is provided in the External Data Repository Section R11.8.1. Alternative approaches, such as possibly examining the presence/absence of marker genes, could potentially help evaluate the similarity of the strains within each microbial genus and to give stronger indicators as to the cause of the clustering.

To lend further support to the tentative pattern observed of pre-14 ka European individuals clustering separately from post-14 ka European individuals, we attempted to recreate the eight 'well-supported' microbial phylogenies generated above, but using the larger shallow sequencing dataset (Fig. S12). Although more shallowly sequenced, this dataset adds two additional pre-14 ka European individuals (PLV001, from Czech Republic and a member of the older Gravettian-associated 'Věstonice' cluster that predates El Mirón; and RIG001, from France and a younger member of the 'El Mirón' cluster (130)), and an additional post-14 ka individual (OFN001, from Mesolithic Germany belonging to the 'Villabruna' cluster (130) that succeeds El Mirón). We re-generated the phylogenies using the same references and mapping settings as for the deeply sequenced dataset above. We expected the phylogenies from this dataset to be of lower-quality due to the presence of aDNA damage in this dataset (which artificially increases variation; however, this was mitigated by requiring a minimum of 2x coverage) and because of this dataset's lower overall coverage, which leads to fewer positions being available for calculating genetic distance. However, we aimed to see if the clustering pattern of the European individuals remained consistent when adding additional individuals. Further details about this analysis and tree generation are provided in External Data Repository Section R11.8.2.

The results of the replicated phylogenies from the shallow sequencing data are shown in Fig. S12, External Data Repository Section R11.8.2. and Fig. R34. As expected, bootstrap support was generally lower, and clade relationships were often not as clear and sometimes followed different patterns than in the original phylogenies. However, despite the presence of DNA damage and fewer shared positions, we again observed the same consistent pattern of EMN001 falling with Neanderthals in all trees (when Neanderthals are also present in the trees). Furthermore, the two additional pre-14 ka individuals - RIG001 and PLV001 - also both fall within the group containing EMN001 in every case. In contrast, OFN001 fell in groups containing other European Mesolithic individuals from El Collado (ECO) and with non-European individuals, similar to the patterns seen in Fig. S11 and External Data Repository Section R11.7 and Fig. R31.

We note that in some trees the African pre-pastoralist individual OAK002 and the 'Epipalaeolithic' individual TAF008 also sometimes fall within the 'pre-14 ka' European groups. We cannot at present explain this pattern. However, OAK002 often has very long branches suggesting that there may be a potential problem or bias with this sample. Despite this, the consistency of the pattern within Europeans in both the deep and shallow sequencing datasets is striking. These findings provide additional microbial support for the population turnover in Europe at around 14 ka as originally described by Posth et al. (131).

Overall, the microbial genomic sequences we obtained from African hominid dental calculus generally formed clades corresponding to the host genus; however, the overall tree topologies do not closely reflect host genus phylogenetic relationships. Rather, *Pan* and *Gorilla* tended to form a monophyletic clade that was either sister to *Homo* (*Actinomyces, Fretibacterium, Fusobacterium, Tannerella*) or nested within *Homo* (*Ottowia, Pseudopropionibacterium*), rather than the expected pattern of *Pan* falling as a sister

clade to *Homo,* and *Gorilla* falling more basal. In fact, this expected pattern was only observed once in the *Streptococcus* tree, but this tree contained very few *Pan* (n=1) and *Gorilla* (n=1) individuals. Within *Homo*, Neanderthals consistently clustered together and they showed an affinity for Upper Palaeolithic Europeans, who either fell within the Neanderthal clade or immediately outside it.

We wish to acknowledge several caveats in this analysis that may have blurred or reduced our ability to identify additional evolutionary signals. First, it is possible that the reference genomes we selected were not the most suitable for assessing phylogenetic differences between highly divergent strains across distant host genera, and a bias may have been introduced by using reference genomes obtained from modern human isolates. Second, accurate calculation of genetic distance for tree building requires highly confident SNP calling; however, in this study the genomic coverage was low, and evidence of cross-mapping was observed for most species; thus, these SNP calls carry a certain amount of noise. Finally, the prohibitive cost of deep sequencing meant that we were unable to generate sufficient data to analyse all individuals, but rather we relied on a smaller subset of deeply sequenced individuals for the majority of our phylogenetic inference.

Despite these caveats, however, the shared sequence variation within each host genus and even host groups (i.e., Neanderthals) suggests that primate oral microbes do carry an evolutionary signal, which warrants further investigation. Further research in this direction will require technical developments in improving the methods for genetic enrichment of particular species in order to obtain sufficiently high genomic coverage for more robust strain separation and SNP calling, which may also allow phylogenetic dating.

S7. Functional and metabolic pathway analysis

The metabolic pathways present in a microbial community represent the range of molecules that can be processed and/or produced by the members of that community. This tells us how microbes interact with their environment, and can be used to predict how a microbial community contributes to host biochemistry (520). Much of the metabolic activity of bacteria remains unknown, with large numbers of genes encoding hypothetical proteins of unknown function in most genomes. Even for *Streptococcus mutans* UA159 (NC_004350.2), a heavily studied oral species involved in dental caries, approximately one third of its genes remain uncharacterised, with the open reading frames are annotated simply as “hypothetical protein” in the genome assembly. However, the potential metabolic repertoire of a community, which is inferred from the total gene content of a metagenomic sample, can offer insights into biofilm ecology and function that cannot be understood by focusing on taxonomy alone. Differences in gut microbiome metabolic functions have been shown among modern human groups with different dietary practices (452, 521). Because modern human diets are known to have changed over the course of hominin evolution, there is the potential for the metabolic functions present in the oral microbiome to have changed as well. We examined the potential metabolic functional profile of our genetic data using two approaches: classification of reads by HUMAnN2 (520) to MetaCyc pathways and KEGG orthologs (using legacy KEGG database v. 56), and classification of reads to SEED categories (522) by MALT using the AADDER tool packaged with MEGAN 6 CE (v6.12.6) (209). Using Principal Component Analysis (PCA), we found that each method consistently separated *Homo*-derived samples from *Pan*-derived samples along PC1, and *Gorilla*-, and *Alouatta*-derived samples along PC2, while *Gorilla*- and *Alouatta*-derived samples consistently clustered together in function, and *Pan*-derived samples clustered separately. This pattern follows the distinctions we saw with taxonomic classifications (see section S4). However, the enzymes with top loadings in these PCAs were different when analysed using KEGG or SEED orthologs.

S7.1 HUMAnN2

Downstream functional analysis with HUMANn2 (520) required input from the taxonomic classifier MetaPhlAn2 (523). We used MetaPhlAn2 (v2.7.1) with default parameters because it has been previously shown that DNA damage does not lead to major biases in taxonomic assignment (204). We generated two versions of the OTU table: one with the estimated relative abundances and one with the number of reads. Detailed information about the analysis is provided in External Data Repository Section R12.3.1, as well as summarised readcount and read_mapped profiles in External Data Repository File R15.

The number of reads assigned by MetaPhlAn2 ranged from 32 to 1,518,861 with a mean of 83,154 across the entire dataset. The plaque group had the greatest mean percentage of reads assigned at 2.8% ± 0.3. This is not unexpected, as MetaPhlAn2 uses a marker gene-based approach that is only expected to assign a small number of reads within the total dataset (203). The mean percentage of reads assigned for each calculus group was: *Alouatta* 0.57% ± 0.16; *Gorilla* 0.73% ± 0.27; *Pan* 0.74% ± 0.24; *Homo* (Neanderthal) 0.85% ± 0.53; *Homo* (Modern Human) 1.2% ± 0.73 (External Data Repository Section R12.3.1 and Fig. R39). Among the African hominid and howler monkey groups, we observe a positive relationship between the average percentage of reads assigned and the evolutionary closeness of the group to modern humans. Across all samples, the two non-host comparative sources had the smallest fraction of reads assigned, with bone environmental proxy controls at 0.02 ± 0.004 and sediment at 0.16% ± 0.11. These observations together suggest a bias in the MetaPhlAn2 database towards modern human-associated bacteria, and that it may under-report or under-represent taxonomic diversity in more diverged host species.

The number of genera identified by MetaPhlAn2 ranged from 1 to 407 with a mean of 49, and for species the range was from 1 to 770 with a mean of 80 (External Data Repository Section R12.3.1 and Fig. R40). Among groups, skin had the highest mean genus-level OTU count at 125 ± 29, while library controls had the lowest with 5 ± 6. The mean number of genera identified for each group was: *Alouatta* 39 ± 10; *Gorilla* 56 ± 19; *Pan* ± 46 ± 16; *Homo* (Neanderthal) 39 ± 21; *Homo* (Modern Human) 54 ± 16. The mean number of species identified for each group was: *Alouatta* 45 ± 16; *Gorilla* 78 ± 29; *Pan* 55 ± 20; *Homo* (Neanderthal) 54 ± 31; *Homo* (Modern Human) 97 ± 36.

The low number of taxa and bias towards modern human-based microbial genomes did not provide sufficient resolution for taxonomic compositional analysis using MetaPhlAn2 compared to the MALT-based analysis, and thus the MetaPhlAn results were not used for taxonomic analyses. We ran HUMAnN2 (v0.11.1) (520) with default parameters but using the MetaPhlAn2 profiles generated above. This used the default UniRef90 protein database (524) bundled with HUMANn2, and the ChocoPhlAn database (2018-07-30). See the External Data Repository Section R12.3, for more details about the MetaPhlAn and HUMAnN2 analyses.

S7.1.1 Pathway abundance

We initially looked at the pathway abundances to determine if the presence or abundance of pre-identified pathways were distinct between host genera. All analyses for HUMAnN2 are reported in detail in the External Data Repository Section R12.3.2. We performed a PCA to look for sample clustering by host genus (External Data Repository Section R12.3.2 and Fig. R41A). The initial PCA including controls demonstrated that the calculus samples clustered predominantly with each other and plaque, but a few samples plotted with extraction and library blanks or bone environmental proxy controls. The samples that plotted with controls except plaque were removed as outliers with poor preservation (External Data Repository Section R12.3.2 and File R37, and the PCA performed again (External Data Repository Section R12.3.2 and Fig. R41B). To see if oral samples clustered by host genus, the control samples were removed and the PCA performed again, which showed a separation of Plaque-*Homo*-*Pan* across PC1 and *Homo*-*Gorilla/Alouatta* across PC2 (External Data Repository Section R12.3.2 and Fig. R41C). Finally, we removed plaque and performed a PCA to see if the calculus samples clustered any more distinctly (External Data Repository Section R12.3.2 and Fig. R41D), but the pattern was highly similar to the plot including plaque. The host genera do separate from each other, but there is overlap between samples at the edges of the clusters, demonstrating that some samples share substantially overlapping metabolic pathway profiles.

S7.1.2 KEGG ortholog distribution

Since the differences in metabolic pathway profiles distinguish host genera by PCA, we asked if the species that contributed the pathways differed by host genus. However, many of the pathways identified by HUMAnN2 were not attributed to a specific species. We were therefore unable to determine how taxonomy related to function at the pathway-level with HUMAnN2 data, and chose instead to look at the

individual gene families, which are reported in reads per kilobase. The percentage of reads assigned to a UniRef90 gene family in *Homo* samples ranged from <5% to >60%, with plaque samples having the highest average of over 60%, while the average for *Alouatta*, *Gorilla*, and *Pan* was <20% (External Data Repository Section R12.3.2 and Fig. R42A). This suggests a substantial proportion of reads in the non-human samples are from genes with unknown functions. To assess the differences in enzyme distributions in our samples, we grouped the HUMAnN2 UniRef90 gene families into KEGG orthologs, and >90% of the UniRef gene families could be grouped into KEGG orthologs for nearly all samples from all host genera (External Data Repository Section R12.3.2 and Fig. R42B). This indicates that even though there is a large amount of unknown potential functional variation represented in our samples, especially from *Pan, Gorilla*, and *Alouatta*, the identifiable gene families are represented in curated databases such as KEGG.

The carbohydrate composition of modern human diets is believed to have changed substantially during human evolution, first with the utilization of carbohydrate-rich underground storage organs and the invention of cooking early in human evolution (46, 525) (and see above section S5.8), followed by the invention of agriculture ca. 10 ka, which provided a much greater proportion of dietary starches (resulting in greater incidence of caries) (526), and then finally with the onset of the industrial revolution, which introduced a large amount of refined simple sugars into modern human diets (527). We therefore chose to look at the distribution of KEGG orthologs found in the 15 carbohydrate pathways listed under KEGG PATHWAY Metabolism (00010, 00020, 00030, 00040, 00051, 00052, 00053, 00500, 00520, 00620, 00630, 00640, 00650, 00660, 00562) in our samples to see if there are different distributions of carbohydrate-processing enzymes in the oral microbiomes of hosts with different inferred carbohydrate intake. We found that the total number of KEGG carbohydrate pathway orthologs slightly differ between the host genera (External Data Repository Section R12.3.2 and Fig. R42C), but are significantly different, e.g., comparing *Homo* to *Pan* and *Homo* to *Gorilla* ($p < 0.05$ by pairwise Wilcox test; For individual comparison values see External Data Repository Section R12.3.2 and Fig. R42C). The abundance of KEGG carbohydrate pathway orthologs (External Data Repository Section R12.3.2 and Fig. R42D) was much more variable between groups, and there are significant differences comparing *Homo* to *Pan*, *Homo* to *Gorilla*, *Homo* to *Alouatta*, *Homo* to Plaque, *Pan* to *Gorilla*, and *Pan* to *Alouatta* ($p < 0.05$ by pairwise Wilcox test; for individual comparison values see External Data Repository Section R12.3.2 and Fig. R42d). *Pan* has the lowest number of KEGG carbohydrate orthologs of the four host genera, and KEGG carbohydrate orthologs are least abundant in *Pan*, which suggests that there are uncharacterised carbohydrate-processing enzymes in the oral microbiome of *Pan* that are not represented in the databases used by HUMAnN2. Comparisons to *Alouatta* may be underpowered, as we have only five *Alouatta* samples.

We next asked if the presence and abundance of KEGG orthologs in calculus are sufficiently different between host genera to separate the samples by host genus in PCA. We followed the same steps for PCA as we used with the pathway abundance data, where we plotted all samples and controls (External Data Repository Section R12.3.2 and Fig. R43A), then removed outliers (External Data Repository Section R12.3.2 and File R40) to ensure we did not include samples with signatures of poor preservation (External Data Repository Section R12.3.2 and Fig. R43B). We then removed the controls except plaque and saw that the host genera separate in a pattern similar to that seen in the pathway abundance PCA (External Data Repository Section R12.3.2 and Fig. R43C), where *Pan* and *Gorilla/Alouatta* separate along PC1 and *Homo* separates from *Pan/Gorilla/Alouatta* along PC2. Finally we removed the plaque samples and removed all KEGG orthologs present at <0.01% abundance (External Data Repository Section R12.3.2 and Fig. R43D). Permanova with adonis2 produced an R^2 value of 0.303. Removing plaque samples did not alter the spatial arrangement of samples from different host genera. We filtered out the KEGG orthologs present at <0.01% abundance to ensure that the data that went into the PCA was not a long tail of low-abundance orthologs possibly deriving from spurious alignments, in the same manner that we filter low-abundance taxa. The sample clustering by host genus we observed in the KEGG ortholog PCA suggested that there are host genus-specific differences in the abundance and distribution of KEGG orthologs. The two *Gorilla* samples that overlap with *Pan* do so in all functional analysis PCAs, with HUMAnN2 and AADDER data, as well as with the taxonomic profiles as seen in External Data Repository Section R9.1 and Figures R20 and R21.

We investigated the most influential KEGG orthologs in the PCA of calculus samples, visualizing them as bi-plots showing the ten orthologs with the strongest positive and negative loadings in PC1 (External Data Repository Section R12.3.2 and Fig. R44A) and PC2 (External Data Repository Section R12.3.2 and Fig. R44C; External Data Repository Section R12.7 and Table R9). The first PC primarily separates *Pan* (negative loadings) from *Gorilla*/*Alouatta* (positive loadings). The KEGG orthologs with strongest negative loading in PC1 are primarily involved in membrane transport, while those with strongest positive loading in PC1 are primarily involved in carbohydrate/energy metabolism and amino acid metabolism (External Data Repository Section R12.3.2 and Fig. R44a). To ensure that the results we observed were not an artefact of including present-day modern human calculus in our analyses, we repeated the PCA and biplots excluding all present-day modern human samples. We saw the same pattern of host genus separation whether or not present-day modern human samples were included (External Data Repository Section R12.3.2 and Fig. R44B,D), although PC2 had flipped to have *Homo* falling in positive coordinates instead of negative coordinates. Nine of the 10 KEGG orthologs with strongest positive PC1 loadings, and 9 of the 10 KEGG orthologs with strongest negative PC1 loadings, were shared both with and without present-day modern human samples. This demonstrates that the present-day modern human samples were not substantially skewing the separation of samples in PC1.

The second PC primarily separates *Homo* (negative loadings) from *Pan*/*Gorilla*/*Alouatta* (positive loadings). The KEGG orthologs with strongest negative loading in PC2 are primarily involved in membrane transport and DNA replication/repair, while those with strongest positive loading in PC2 are primarily involved in carbohydrate/energy metabolism and membrane transport (External Data Repository Section R12.7 and Table R9; External Data Repository Section R12.3.2 and Fig. R44a,c). To aid visual comparisons of the PCAs including and excluding present-day modern human samples, we reversed the y-axis (PC2) in the PCA that lack present-day modern human samples. This leads to some confusion when comparing the KEGG orthologs with the strongest loadings in PC2, which we hope to clarify as follows. In the PCA including present-day modern human samples, *Homo* plots predominantly in negative coordinates, but in the PCA excluding present-day modern human samples, *Homo* plots predominantly in positive coordinates. Therefore, we compare the strongest negative PC2 loadings of the PCA with present-day modern human samples to the strongest positive PC2 loadings of the PCA without present-day modern human samples. And the reverse, we compare the strongest positive PC2 loadings of the PCA with present-day modern human samples to the strongest negative PC2 loadings of the PCA without present-day modern human samples. Keeping this in mind, 8 of the 10 KEGG orthologs with strongest PC2 loadings in the direction characterizing *Homo* were shared both with and without present-day modern human samples, while 7 of the 10 KEGG orthologs with strongest PC2 loadings in the direction characterizing *Pan/Gorilla/Alouatta* were shared both with and without present-day modern human samples. This suggests there was more bias of present-day modern human samples separating the genus *Homo* from the other genera in PC2 than in PC1; however the orthologs that characterise *Homo* are largely conserved throughout the evolution of modern humans.

We wanted to visualise the distribution of the 10 KEGG orthologs with the strongest loadings in PC1 and PC2 in the host genera to see how they differ between hosts. We generated heat maps with the CLR-transformed read copies per million assigned to each ortholog in each sample, for both the PCA including modern *Homo* (External Data Repository Section R12.3.2 and Fig. R45) and for the PCA excluding modern *Homo* (External Data Repository Section R12.3.2 and Fig. R46. The orthologs with strongest loadings in PC1 positive values are more abundant in *Gorilla/Alouatta* and modern *Homo* compared to *Pan* (where they are almost absent), historic *Homo*, and Neanderthals (External Data Repository Section R12.3.2 and Fig. R45A). *Gorilla/Alouatta* have the lowest abundance of orthologs with the strongest PC1 negative loadings (External Data Repository Section R12.3.2 and Fig. R45B), which is expected since the *Gorilla/Alouatta* samples plot in positive PC1 values (External Data Repository Section R12.3.2 and Fig. R44A). Neanderthal samples have the lowest abundance of orthologs with strongest PC2 positive loadings (External Data Repository Section R12.3.2 and Fig. R45C). *Homo* samples have the highest abundance of orthologs with strongest loadings in PC2 negative values while *Pan* have the lowest abundance (External Data Repository Section R12.3.2 and Fig. R45D), which is concordant with most *Homo* samples plotting in negative PC2 values and *Pan* samples plotting in positive values (Fig. S13A,C; External Data Repository Section R12.3.2 and Fig. R44).

We next generated heatmaps to visualise the distribution of KEGG orthologs with strongest loadings in PC1 and PC2 from the PCAs excluding present-day modern human samples. The orthologs with strongest positive PC1 loadings are more abundant in *Alouatta/Gorilla* than *Homo* and least abundant in *Pan* (External Data Repository Section R12.3.2 and Fig. R46A), while those with strongest negative PC1 loadings are slightly more abundant in *Homo* than *Gorilla/Pan* and are least abundant in *Alouatta* (External Data Repository Section R12.3.2 and Fig. R46B). *Homo* samples have the highest abundance of orthologs with strongest PC2 positive loadings, which separate *Homo* from *Alouatta/Gorilla* and *Pan* (External Data Repository Section R12.3.2 and Fig. R46C), while the orthologs with strongest PC2 negative loadings are slightly more abundant in *Alouatta/Gorilla* than *Homo*, especially Neanderthals (External Data Repository Section R12.3.2 and Fig. R46D). There is a clear opposing pattern of ortholog abundance in External Data Repository Section R12.3.2 and Fig. R46C (low *Alouatta/Gorilla/Pan*, high *Homo*) and External Data Repository Section R12.3.2 and Fig. R46D (high *Alouatta/Gorilla/Pan*, low *Homo*) that is in line with the sample separation on the PCA (Fig. S13A,C; External Data Repository Section R12.3.2 and Fig. R44C,D).

S7.1.3 Species contributions to KEGG orthologs

To determine which microbial species were most responsible for driving functional differences among African hominids, we extracted from each calculus sample the relative abundance of each microbial species contributing to each KEGG ortholog with the top 10 strongest positive and negative loadings on PC1 and PC2, both with and without present-day modern human samples, from the HUMAnN2 gene families table. This allowed us to see if certain species were responsible for the distribution of KEGG orthologs in each host genus. We summed the abundance of KEGG orthologs from each species of each microbial genus to get a genus average. We plotted the average contribution of each genus that contributed a relative abundance of more than 12% to each ortholog, with any that had <12% grouped together as Other (Fig. S13B,D; External Data Repository Section R12.3.2 and Fig. R47-50). In External Data Repository Section R12.3.2 and Fig. R47-50, the results from PCAs including present-day modern human samples are shown in panels A and B, while results from PCAs excluding present-day modern human samples are panels C and D. We plotted the microbial genus distribution from each host genus grouped by ortholog (panels A,C) to look for microbial genus biases within orthologs, and the microbial genus distribution of each ortholog grouped by host genus (panels B,D) to look for microbial genus biases within host genera.

In the KEGG orthologs with strongest positive loadings in PC1 (External Data Repository Section R12.3.2 and Fig. R47), a large proportion of the orthologs are not attributable to specific species, as often >25% is unclassified. There are no clear biases in microbial genera that contribute to the orthologs, as we do not see specific species shared across host genera for each ortholog (External External Data Repository Section R12.3.2 and Fig. R47A,C). However, different species characterise each host genus, and many of the same microbial genera contribute to multiple orthologs in each host (External Data Repository Section R12.3.2 and Fig. R47B,D). For example, in *Homo*, the microbial genera *Streptococcus*, *Rothia*, and *Neisseria* are the most prevalent and abundant genera contributing to the orthologs, and *Streptococcus* contributes >25% to 6 orthologs (External Data Repository Section R12.3.2 and Fig. R47B,D). *Propionibacterium* contributes to 4 orthologs in both *Homo* and *Pan* including present-day modern human samples, but to 7 and 5, respectively, when excluding present-day modern human samples. In *Gorilla, Neisseria* contributes to all of the orthologs, and *Rothia* contributes to 3 of the 10. *Alouatta* has the highest proportion of KEGG orthologs from unidentified species, while *Aggrebatibacter* contributes to 4 and *Streptococcus* to 3 orthologs in *Alouatta*.

For the KEGG orthologs with strongest negative loadings in PC1 (External Data Repository Section R12.3.2 and Fig. R48), most of the orthologs are contributed to by identified genera, in contrast to the orthologs with positive loadings, which had a higher proportion of reads that were not assigned to a species and are listed as unclassified. There are no clear biases in microbial genera that contribute to the orthologs, as we do not see specific species shared across host genera for each individual ortholog (External Data Repository Section R12.3.2 and Fig. R48A,C). Each host genus has a wider variety of species contributing to the orthologs than for the PC1-positive orthologs (External Data Repository Section R12.3.2 and Fig. R48B,D), and some possible biases can be seen when grouping the orthologs

by host genus. Here *Fretibacterium* is represented in K01999, K02005, and K02034 in *Homo*, *Pan*, and *Gorilla*, both with and without present-day modern human samples (panels B,D). *Streptococcus* contributions were prevalent and abundant in *Homo* both with and without present-day samples, while *Propionibacterium* contributes to multiple orthologs in all four host genera. Some of the reads assigned to *Propionibacterium* may come from *Pseudopropionibacterium* in *Homo*, as this genus is highly abundant in modern human dental calculus but the MetaPhlAn2 database we used includes the reference genome for *Pseudopropionibacterium propionicum* under the its old name of *Propionibacterium propionicum*.

For the KEGG orthologs with strongest loadings in PC2 characterizing *Pan/Gorilla/Alouatta* (External Data Repository Section R12.3.2 and Fig. R49), (positive including present-day modern human samples, negative excluding present-day modern human samples) there is high variability in whether the ortholog is contributed by known or unclassified species. Orthologs K00794, K01759 (External Data Repository Section R12.3.2 and Fig. R49A), K01955, K03046 (External Data Repository Section R12.3.2 and Fig. R49c), and K00027, K03043 (External Data Repository Section R12.3.2 and Fig. R49A,C) have high proportions represented by unclassified species, but these are conserved enzymes and it may not be possible to identify the species of origin from short DNA fragments. *Homo* and *Pan* have a wider variety of species contributing to the orthologs than *Gorilla* and *Alouatta* (External Data Repository Section R12.3.2 and Fig. R49B,D). Some possible biases can be seen when grouping the orthologs by host genus. *Fretibacterium* contributes to 7 of the 10 KEGG orthologs in *Pan* (panels B,D), and to 7 of the 10 orthologs in *Homo* without present-day samples (External Data Repository Section R12.3.2 and Fig. R49D). *Aggregatibacter* contributes to 6 orthologs in *Gorilla* both when including (External Data Repository Section R12.3.2 and Fig. R49B) and excluding present-day modern humans (External Data Repository Section R12.3.2 and Fig. R49D), and in *Alouatta, Aggregatibacter* contributes to 7 orthologs when including present-day modern human samples and 6 orthologs when excluding present-day modern human samples.

For the KEGG orthologs with strongest loadings in PC2 characterizing *Homo* (Fig. S13B,D; External Data Repository Section R12.3.2 and Fig. R50), (negative including present-day modern human samples, positive excluding present-day modern human samples) there are biases in the microbial species that contribute to specific KEGG orthologs (External Data Repository Section R12.3.2 and Fig. R50A,C), and to the orthologs in a single host genus (Fig. S13B,D; External Data Repository Section R12.3.2 and Fig. R50B,D). *Propionibacterium* is the primary contributor in all 4 host genera of orthologs K02026 (Fig. S13D; External Data Repository Section R12.3.2 and Fig. R50D) and K02027 (Fig. S13B,D; External Data Repository Section R12.3.2 and Fig. R50B,D), both multiple sugar transport system permease proteins, and contributes nearly 50% of K01990, an ABC-2 type transport system ATP-binding protein, in all 4 host genera. In *Homo, Streptococcus* and *Propionibacterium* together contribute over 75% of all 10 orthologs, with *Streptococcus* contributing >50% in 8 orthologs. *Propionibacterium* contributes to 9, 8, and 4 orthologs in *Pan*, *Gorilla*, and *Alouatta*, respectively, when including present-day modern human samples, and to 10, 10, and 4 orthologs in *Pan*, *Gorilla*, and *Alouatta*, respectively, when excluding present-day modern human samples. In *Pan*, *Olsenella* contributes to 5 and 4 of the orthologs when including and excluding present-day modern human samples, respectively, and *Fretibacterium* to 4 orthologs both including and excluding present-day modern human samples. In *Gorilla*, *Rothia* contributes to 4 of the orthologs, and in *Alouatta, Aggregatibacter* contributes to 5 of the orthologs, both including and excluding present-day modern human samples (Fig. S13B,D; External Data Repository Section R12.3.2 and Fig. R50B,D).

S7.1.4 Metabolic category PCAs

Finally, we wanted to know if the orthologs in specific metabolic pathways involved in processing major biomolecule classes (carbohydrates, amino acids, and lipids) were individually distinctive between the host genera. The orthologs in all pathways of the KEGG metabolism categories carbohydrates (297 orthologs), amino acids (262 orthologs), and lipids (94 orthologs) were subsetted from all of the orthologs identified in the samples, and used to run PCAs (Fig. S10, External Data Repository Section R12.3.2 and Fig. R51). Samples clustered by host genus similar to other functional classification PCAs presented above. This suggests that there are host genus-specific differences in the functional profiles of multiple major biomolecule processing pathways. Carbohydrate processing orthologs (External Data Repository

Section R12.3.2 and Fig. R51A) did not cluster the host genera more tightly or distinctly than amino acids processing orthologs (External Data Repository Section R12.3.2 and Fig. R51B) or lipid processing orthologs (External Data Repository Section R12.3.2 and Fig. R51C). Permanova with adonis2 produced R^2 values of 0.285 for amino acid pathway orthologs, 0.291 for carbohydrate pathway orthologs, both slightly lower than 0.303 for all orthologs, and 0.319 for lipid pathway orthologs.

Taken together, the results of HUMAnN2-identified potential metabolic functional analyses demonstrate that there are characteristic differences between calculus samples from *Homo, Pan, Gorilla*, and *Alouatta*. Two different approaches to analyzing the potential metabolic functional profiles - pathway analysis and KEGG ortholog analysis - both cluster the samples by host genus in PCA. Exploration of the signals that drive host genus separation revealed that clustering is not strongly biased by the presence of present-day modern human samples, and that in certain cases, the orthologs that most strongly drive separation are derived from specific microbial species. In particular, *Streptococcus* contributes to many of the orthologs in *Homo* but not the remaining host genera, even when excluding present-day modern human samples, and suggests that *Streptococcus* characterises and fulfils a distinct role in *Homo*.

S7.2 AADDER

We additionally investigated functional assignments of our metagenomes using SEED categories (522), to test for consistency between the different methods of functional profiling. This was performed using the AADDER program, as bundled with MEGAN6 CE (v6.12.6 (209)), which infers functional information from reads aligning within gene regions of annotated reference genomes. This is beneficial for aDNA as fragmented reads are typically too short to translate into informative protein sequences, which are often used in modern contexts. First, we used the custom RefSeq database that we had created for our MALT analyses (External Data Repository Section R12.5), and re-ran malt-run but with the additional flags of --samSoftClip --format SAM, to create SAM files for the taxonomic alignments. Using the corresponding GFF files associated with the FASTA files downloaded from RefSeq, an AADDER annotation index was constructed and annotation information added to the taxonomic assignments using aadder-run on default settings. The resulting SAM files were then converted to RMA6 files using the blast2rma tool also bundled with MEGAN6 CE, and in addition to NCBI taxonomy information, SEED categories were also added to alignments using the corresponding acc2seed file.

S7.2.1 SEED profile

The SEED classifications consist of 3 levels, starting with broad categories (such as Carbohydrates and Amino Acids), then individual pathways within these broad categories, and finally proteins within each pathway. Assignments at each level are summed from the levels below, and assignments may be made at the pathway or broad category level that are not present at lower levels. All analyses for this section are provided in the External Data Repository Section R12.6.

To look at the efficiency of assigning functions to reads with SEED we calculated the percent of reads assigned in each sample type. *Pan* had the lowest percentage of reads assigned, consistent with the HUMAnN2 results, while *Alouatta* samples had the highest percentage reads assigned (External Data Repository Section R12.6 and Fig. R52A). Control samples had high read assignments, but also the fewest reads per sample. The percentage of reads assigned to Neanderthal and modern human samples overlapped. We looked at the percentage of assigned reads that were in the Carbohydrates category to see if changes in dietary carbohydrate consumption throughout human evolution was reflected in the amount of carbohydrate-processing proteins by this classification method, although this was not the case in HUMAnN2. A higher percentage of reads were assigned to Carbohydrates for all host genus calculus samples than for plaque (External Data Repository Section R12.6 and Fig. R52B), but the average for *Homo* was only slightly higher than *Pan*, *Gorilla* and *Alouatta.* The percentage of reads assigned to specific proteins in the Carbohydrates category was highest in *Alouatta* and lowest in *Pan*, but all groups had a wide spread (External Data Repository Section R12.6 and Fig. R52C). The number of Carbohydrate-category proteins identified in each sample is shown in External Data Repository Section R12.6 and Fig. R52D. *Alouatta* and *Gorilla* had more Carbohydrate-category proteins than *Pan* or *Homo*, which was in cases significant (External Data Repository Section R12.6 and Fig. R52D). The abundance

(number of assigned reads) of Carbohydrate-category proteins was higher in *Homo* than in *Pan, Gorilla,* and *Alouatta*, which was in cases significant (External Data Repository Section R12.6 and Fig. R52E), while *Pan* had the lowest abundance. Plaque had the lowest number of Carbohydrate-category proteins identified, but had the highest average abundance of Carbohydrate-category proteins. These trends are similar to those of HUMAnN2 (External Data Repository Section R12.3.2 and Fig. R42), where *Pan* has the lowest prevalence and abundance of identified enzymes and *Homo* has a wide spread, although the number of proteins identified involved in carbohydrate processing is much higher using AADDER (range 200-800) than HUMAnN2 (range 100-500).

We assessed differences in the SEED-identified protein composition of our samples by PCA to see if the presence and abundance of proteins across all SEED categories are sufficiently different between host genera to separate the samples by host genus. This analysis used only the assignments at the protein level. We followed similar steps for PCA (External Data Repository Section R12.6 and Fig. R53) as we used with the HUMAnN2 pathway (External Data Repository Section R12.3.2 and Fig. R41) and KEGG ortholog (External Data Repository Section R12.3.2 and Fig. R43) data. First we plotted all samples and controls (External Data Repository Section R12.6 and Fig. R53A), then removed outlier samples to ensure we did not include samples with signatures of poor preservation, and removed proteins that had been identified as possible contaminants by decontam (235) , which may derive from contamination (External Data Repository Section R12.6 and Fig. R53B). Outliers were those samples that overlapped in PC1 and PC2 with controls (excluding plaque) (External Data Repository Section R12.6 and File R37). Decontam was run on a table of only the protein-level assignments from all SEED categories, using the combined method with a threshold of 0.6, which identified 116 proteins as possible contaminants. Our controls included extraction blanks and library build blanks to identify putative laboratory contamination, as well as archaeological bone samples to identify possible contamination from burial soil, excavation, and handling. Testing the threshold showed that increasing it identified more potential protein contaminants, which were often low abundance. As we also included a filtering step later to remove all proteins present at less than 0.05% abundance, and this filter removed many of the same proteins identified by decontam at higher threshold, we selected a threshold of 0.6 as our cut-off.

We then removed the controls except plaque and filtered out all proteins present at <0.05% abundance and saw that the host genera form clusters with a little overlap (External Data Repository Section R12.6 and Fig. R53C). We filtered the proteins based on abundance to ensure that the data that went into the PCA was not a long tail of low-abundance proteins, in the same manner that we filtered low-abundance taxa. Finally we removed the plaque samples from the filtered data (External Data Repository Section R12.6 and Fig. R53D), which increased separation of samples by host genus. Permanova with adonis2 for this PCA gave an R^2 value of 0.345. The sample cluster pattern is similar to the clustering in PCAs of HUMAnN2-identified pathway abundance and KEGG ortholog profiles, where *Pan* and *Gorilla/Alouatta* separate along PC1 and *Homo* separates from *Pan/Gorilla/Alouatta* along PC2.

We next wanted to know which proteins drive separation of the samples for each host genus, and created biplots on the final filtered enzyme PCA with the top 10 enzymes with strongest positive and negative loadings in PC1 and PC2. We did this for a PCA that included all *Homo* samples (Fig. S13E; External Data Repository Section R12.6 and Fig. R54A,C), and again for a PCA that excluded the present-day modern human samples (Fig. S13G; External Data Repository Section R12.6 and Fig. R54B,D), to ensure that the signals we saw were not driven predominantly by present-day modern human samples. There was no overlap between the top SEED-identified proteins that drive sample separation and the top KEGG orthologs that drive sample separation, and even the higher levels of classification for the proteins or orthologs (SEED pathways or categories and KEGG pathways, respectively) were not concordant (External Data Repository Section R12.7 and Table R9). This stark difference in the proteins/pathways that separate host genera between SEED and KEGG orthologs suggests that the results are highly program-specific, and conclusions regarding functional differences should be drawn with caution.

The top 10 proteins with strongest negative loadings in PC1, predominantly separating *Gorilla/Alouatta* from *Pan*, belong to 8 categories, with 2 proteins from Amino Acids and Derivatives, and the remaining 8 from individual categories (External Data Repository Section R12.6 and Fig. R54A). When present-day modern human samples were removed, only 2 of the proteins with strongest loading were different from

those in the PCA including present-day modern human samples, and were part of the Carbohydrates category and the DNA Metabolism category, so that there were 2 proteins each in the categories Amino Acids and Derivatives, Carbohydrates, and DNA Metabolism. The 8 proteins that overlapped are indicated in bold in the tables of External Data Repository Section R12.6 and Fig. R54A,B. The top 10 proteins with the strongest positive loadings in PC1 (predominantly separating *Pan* from *Gorilla/Alouatta*) belong to a variety of categories, with 2 from Amino Acids and Derivatives and 2 from Phages, Prophages, Transposable elements, Plasmids, and the remaining 6 from individual categories (External Data Repository Section R12.6 and Fig. R54A). When present-day modern human samples were removed, only 2 of the proteins with strongest loadings were different from those in the PCA including present-day modern human samples, and they were part of the Respiration category. The 8 that overlapped are indicated in bold in the tables of External Data Repository Section R12.6 and Fig. R54A,B.

The PCA of proteins excluding the present-day modern human samples reversed the coordinates of samples along PC2 relative to the PCAs that include present-day modern human samples. To aid visual comparisons of the PCAs including and excluding present-day modern human samples, we reversed the y-axis (PC2) in the PCA that lack present-day modern human samples (Fig. S13G; External Data Repository Section R12.6 and Fig. R54B,D), just as we did for the PCAs with KEGG orthologs (Fig. S13C; External Data Repository Section R12.3.2, Fig. R42C,D). This also leads to some confusion when comparing the proteins with strongest loadings in PC2. In the PCA including present-day modern human samples, *Homo* plots predominantly in negative coordinates, but in the PCA excluding present-day modern human samples, *Homo* plots predominantly in positive coordinates. Therefore, we compare the strongest negative PC2 loadings of the PCA with present-day modern human samples to the strongest positive PC2 loadings of the PCA without present-day modern human samples. And the reverse, we compare the strongest positive PC2 loadings of the PCA with present-day modern human samples to the strongest negative PC2 loadings of the PCA without present-day modern human samples.

Keeping this in mind, 8 of the 10 proteins with strongest PC2 loadings in the direction characterizing *Homo* samples were shared both with and without present-day modern human samples (Fig. S13E,G; External Data Repository Section R12.6 and Fig. R54C,D, andTable R9). Five of the 10 proteins in the PCA including present-day modern human samples are in the Carbohydrates category, while a 6th protein categorised in Cell Wall and Capsule, is a sialidase that also processes carbohydrates, suggesting a strong representation of carbohydrate processing in calculus communities from *Homo*. The remaining 4 proteins that strongly characterise *Homo* are in individual categories, and two of these, choline-binding protein A and serine endopeptidase ScpC are *Streptococcus*-specific proteins. Four of the 10 proteins in the PCA excluding present-day modern human samples are in the Carbohydrates category, and only one is different from the Carbohydrate proteins in the PCA including present-day modern human samples. The strong signal of Carbohydrate proteins characterizing *Homo* samples remains after removing the present-day modern human samples, and is not an artefact of modern carbohydrate-rich diets.

The top 10 proteins with the strongest PC2 loadings in the direction characterizing *Pan/Gorilla/Alouatta* fall in a variety of categories, and this is true both when present-day modern humans are included and excluded in the PCA (Fig. S13G; External Data Repository Section R12.6 and Fig. R54B,D, and Table R9). Five of the 10 proteins with the strongest PC2 loadings in the direction characterizing *Pan/Gorilla/Alouatta* were shared in the PCAs both with and without present-day modern human samples (External Data Repository Section R12.6 and Fig. R54C,D, and Table R9). Only 1 protein falls in the Carbohydrates category when including or excluding present-day modern human samples, but it is a different protein in each PCA (External Data Repository Section R12.7 and Table R9). The four proteins that are unique to each PCA are in different categories, including DNA Metabolism, Nucleosides and Nucleotides, and Potassium Metabolism when including present-day modern human samples, and Cell Wall and Capsule, Iron Acquisition and Metabolism, and Phages, Prophages, Transposable elements, Plasmids when excluding present-day modern human samples.

We wanted to visualise the distribution of the 10 proteins with the strongest loadings in PC1 and PC2 across host genera to see how they differ between hosts. We generated heat maps with the CLR-

transformed total number of reads assigned to each protein in each sample for the PCA including present-day modern humans (External Data Repository Section R12.6 and Fig. R55) and for the PCA excluding present-day modern humans (External Data Repository Section R12.6 and Fig. R56). There are many more reads assigned to present-day modern human samples than to the historic modern humans, Neanderthals, *Pan*, *Gorilla,* or *Alouatta* in PC1 negative (External Data Repository Section R12.6 and Fig. R55A), PC1 positive (External Data Repository Section R12.6 and Fig. R55B), PC2 negative (External Data Repository Section R12.6 and Fig. R55C) and PC2 positive (External Data Repository Section R12.6 and Fig. R55D) space; however, there are nevertheless host genus-specific patterns. For example, the top 10 negative PC1 protein loadings in *Pan* were all present at very low abundance (External Data Repository Section R12.6 and Fig. R55A); by contrast the top 10 negative PC1 protein loadings in *Gorilla* and *Alouatta* were similar in abundance, which is expected given that *Gorilla* and *Alouatta* plot together in negative PC1 space, separate from *Pan* (External Data Repository Section R12.6 and Fig. R55). In contrast, the positive PC1 protein loadings are more abundant in *Pan* than *Gorilla/Alouatta*, while abundance in *Homo* is variable (External Data Repository Section R12.6 and Fig. R55B). The PC2 negative proteins are more abundant in *Homo* than *Pan/Gorilla/Alouatta* (Fig. S13E; External Data Repository Section R12.6 and Fig. R55C), which is expected since *Homo* separates from the other host genera in negative PC2 space (External Data Repository Section R12.6 and Fig. R54). The PC2 positive proteins are least abundant in *Pan*, which makes sense since all 10 proteins have negative PC1 values and therefore also drive the plotting of *Gorilla/Alouatta* (Fig. S13G; External Data Repository Section R12.6 and Figures R55D, R54).

The host genus-based patterns of the abundance of proteins with the strongest loadings from the PCA excluding present-day modern humans are similar to those including present-day modern humans (External Data Repository Section R12.6 and Fig. R56). Pan has a very low abundance of proteins in the top 10 PC1 negative loadings, while *Gorilla*/*Alouatta* are similar to each other and *Homo* is variable (External Data Repository Section R12.6 and Fig. R56A). In contrast, the proteins that have the top 10 loadings in PC1 positive values are more abundant in *Pan* than in *Gorilla/Alouatta* or *Homo* (External Data Repository Section R12.6 and Fig. R56B). Two proteins in the top 10 PC2 negative loadings (pc2n2 - conjugative transposon TraG; pc2n3 - Na(+)-translocating NADH quinone reductase subunit F [E.C. 1.6.5.-]) are much more abundant in *Pan* than the other proteins, and are much lower in several *Gorilla/Alouatta/Homo* samples than the other proteins (External Data Repository Section R12.6 and Fig. R56C). Generally, however, the PC2 negative top 10 proteins are slightly more abundant in *Gorilla/Alouatta* than in *Pan/Homo* (External Data Repository Section R12.6 and Fig. R56C). The proteins with the top loading scores in PC2 positive, which drive separation of *Homo* from *Pan/Gorilla/Alouatta*, are generally more abundant in *Homo* than in the other host genera, with the exception of several modern human samples in which the proteins are very low abundance.

S7.2.2 Species contributions to SEED proteins

To understand if there were specific species contributing to the proteins in the top 10 loadings for PC1 and PC2 in negative and positive directions, we examined the species assignments of the reads assigned to each of these proteins. Using MEGAN6, we created a comparison of the AADDER output .rma6 files for each host genus that included all samples in the final PCA (outliers excluded). Within the SEED viewer in MEGAN6, the node for each protein of interest was selected and individually exported to a new document. This document contains the species assignments for each read, but for the entire host genus comparison, not separated by individual. The LCA parameters were changed to Top Percent 0.001, and then all nodes at the species level were selected. The summed assignments of each species were exported as a .tsv file for import into R for data processing and visualization. In R, the species that contributed to less than 15% of the reads of each protein were grouped together as “Other”, and then those species that individually contributed >15% of the reads were plotted by protein and by host genus.

Three genera appear to be the main contributors to the proteins with the top negative loadings in PC1, irrespective of whether present-day modern human samples are included (External Data Repository Section R12.6 and Fig. R57). These are *Aggregatibacter*, *Neisseria*, and *Streptococcus*. There is a wide range in the number of reads assigned to each bacterial genus among host genera, so although it may appear from Fig. R57 that one bacterial genus is the dominant contributor to a protein across all host

genera, the number of reads of that protein may be orders of magnitude different between hosts. For example, for pc1n4 (glutathione biosynthesis bifunctional protein GshF), *Aggregatibacter* and *Streptococcus* are the two highest contributors across all host genera, but the read count break down for *Aggregatibacter* is *Homo*: 24,119 *Pan*: 2, *Gorilla*: 1,981, *Alouatta*: 2,069, and for *Streptococcus* is *Homo*: 51,257, *Pan*: 62, *Gorilla*: 449, *Alouatta*: 5,468 (External Data Repository Section R12.6 and Fig. R57A,B). The difference is even more stark for pc1n8 (serine endopeptidase ScpC), which is contributed entirely by *Streptococcus* across all host genera, but with a read break-down of *Homo*: 110,423, *Pan*: 12, *Gorilla*: 436, *Alouatta*: 1 (External Data Repository Section R12.6 and Fig. R57A,B). We see a similar trend when excluding present-day modern humans (External Data Repository Section R12.6 and Fig. R57C,D), although then the number of reads assigned to *Homo* is much lower; for example the number of GshF read counts drops for *Aggregatibacter to* 236 and for *Streptococcus* to 462. Pan has many more species that individually contribute less than 15% of the reads assigned to each protein, seen with the much larger “Other” bar across all proteins than in the other host genera, which is likely related to a database bias, where we lack species representative of the chimpanzee oral cavity.

There are many more species that contribute to the 10 proteins with top positive loadings in PC1 than to the top negative loadings. This is seen in the higher percent of “Other” genera contributing to positive loadings (External Data Repository Section R12.6 and Fig. R58). A variety of bacterial genera contribute at least 15% of the reads in each host genus, with none of the bacterial genera appearing to be characteristic of any protein (External Data Repository Section R12.6 and Fig. R58A) or host genus (External Data Repository Section R12.6 and Fig. R58B). When removing present-day modern human samples from the analysis, there are changes in the species that contribute proteins in the host genera (External Data Repository Section R12.6 and Fig. R58C,D), but still no species predominate.

The proteins with strongest PC2 loadings in the direction that characterised *Alouatta*/*Gorilla*/*Pan* are shown in External Data Repository Section R12.6 and Fig. R59. There are several species that contribute to the SEED proteins with the strongest PC2 loadings in the direction characterizing *Pan/Gorilla/Alouatta* that are shared across multiple host genera. These include *Aggregatibacter* in *Homo*, *Gorilla*, and *Alouatta*, *Neisseria* in *Homo* and *Gorilla,* and *Capnocytophaga* in *Homo,* and this pattern holds both with (External Data Repository Section R12.6 and Fig. R5A,B) and without (External Data Repository Section R12.6 and Fig. R59C,D) present-day modern human samples. For the proteins that remain in the top 10 strongest loadings when excluding present-day modern human samples, the microbial species that contribute reads to these proteins remain the same irrespective of whether present-day modern humans are included.

The proteins with the strongest PC2 loadings in the direction that characterises *Homo* in the PCAs (both with and without present-day modern humans) are shown in External Data Repository Section R12.6 and Fig. R60. There is clearly a strong bias of *Streptococcus* contributing a high percentage of the reads to several proteins across multiple host genera (External Data Repository Section R12.6 and Fig. R60A-D). There are stark differences in the number of reads assigned to each protein for each host genus and this demonstrates that *Streptococcus* is dominant only in *Homo* samples. For example, pc2n2 (serine endopeptidase ScpC, External Data Repository Section R12.6 and Fig. R54C,D) comes from only *Streptococcus* reads across all host genera, but the read breakdown when including present-day *Homo* is *Homo*: 110,423 (including present-day *Homo*)/7,930 (excluding present-day *Homo*), *Pan*: 12, *Gorilla*: 436, *Alouatta*: 1. Similarly, pc2n3 (Pullulanase, External Data Repository Section R12.6 and Fig. R54C,D) comes from mostly *Streptococcus* reads, with a breakdown of *Homo*: 196,683 (including present-day *Homo*)/15,104 (excluding present-day *Homo*), *Pan*: 92, *Gorilla*: 993, *Alouatta*: 58. Although removing the present-day modern human samples substantially reduces the number of reads assigned to *Streptococcus* for each protein, the pattern remains similar to when including present-day samples, suggesting that the present-day modern human samples are not biasing species profiles of *Homo* samples.

S7.2.3 Metabolic category PCAs

Finally, we wanted to know if the proteins in specific metabolic pathways involved in processing major biomolecule classes (carbohydrates, amino acids, and lipids) were distinctive among the host genera, as

was found for HUMAnN2-identified KEGG orthologs (Fig. S10, External Data Repository Section R12.6 and Fig. R61). The proteins in all pathways of the SEED categories carbohydrates, amino acids, and fatty acids, lipids, and isoprenoids were subsetted from all of the proteins identified in the samples, and used to run PCAs (External Data Repository Section R12.6 and Fig. R61). Carbohydrate processing proteins (External Data Repository Section R12.6 and Fig. R61A) clustered the host genera in a pattern similar to that of all proteins (External Data Repository Section R12.6 and Fig. R53D) yet with a slightly higher R^2 value (0.423 for carbohydrate pathway proteins vs 0.345 for all proteins), demonstrating distinctly different host profiles of carbohydrate processing proteins. Higher levels of sample overlap in PCA and lower permanova R^2 values were observed for amino acid processing proteins (R^2 = 0.289, External Data Repository Section R12.6 and Fig. R61B) and fatty acid processing proteins (R^2 = 0.258, External Data Repository Section R12.6 and Fig. R61C).

This is different from the results from HUMAnN2, in which the KEGG orthologs that process each of these major biomolecule classes separate the samples by host genus (External Data Repository Section R12.3, Fig. R51). To investigate the degree of overlap between the proteins/orthologs identified by HUMAnN2 and AADDER, we compared the KEGG orthologs and SEED-identified proteins in each of the three categories (carbohydrates, amino acids, and fatty acids/lipids) using their EC numbers. Details of this analysis are provided in External Data Repository Section R12.7. Only a small number of SEED-identified proteins did not have EC numbers, and we manually checked for these in the KEGG ortholog lists; however, only a single additional overlap in lipids was found. More KEGG orthologs were identified in the samples compared to SEED proteins for each metabolic category: amino acids, 262 orthologs vs 107 proteins; carbohydrates, 297 orthologs vs 79 proteins; and lipids, 94 orthologs vs 11 proteins. There was considerable overlap in the EC numbers of the orthologs/proteins identified in each category, with 87 shared in amino acids, 52 shared in carbohydrates, and 7 shared in lipids/fatty acids. The differences in PCA sample clustering between KEGG orthologs and SEED proteins may be due to the fact that the different classification systems do not include the same proteins/orthologs. This observation further supports our finding that differences in the observed functional profile are partly database-dependent, and therefore interpretation should be made with care.

Supplementary Figures and Tables

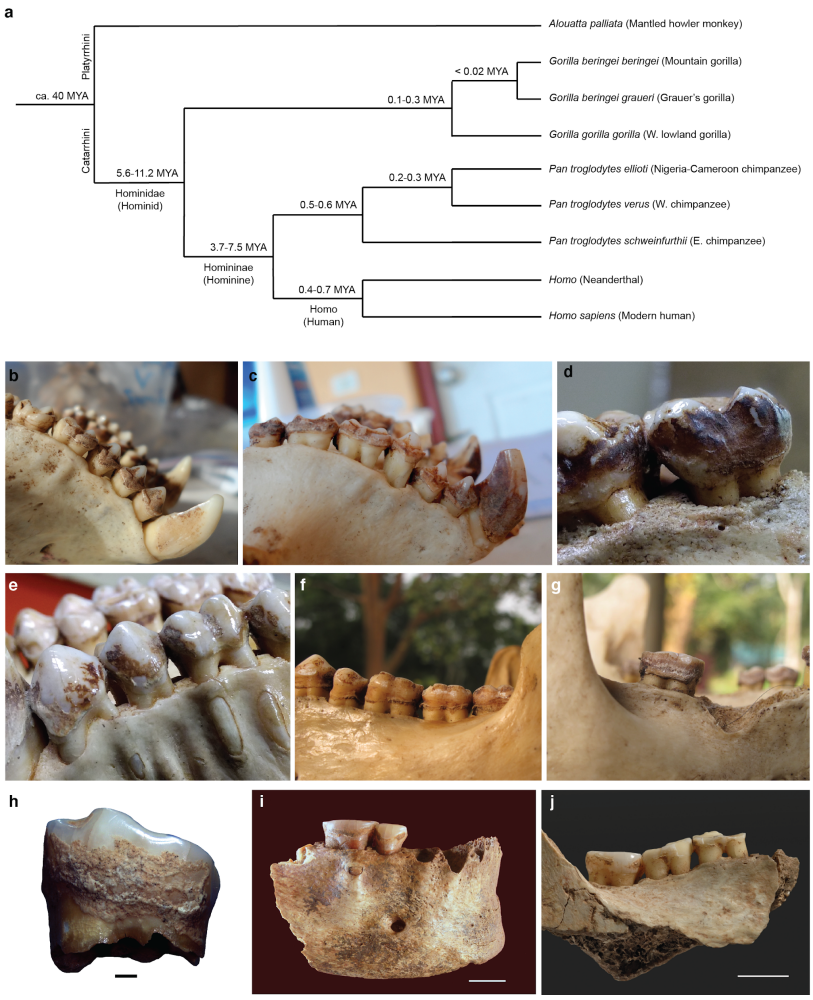

Fig. S1. Primate species analysed in this study and their dental calculus. (*A*), Cladogram of evolutionary relationships among host species analysed in this study. Estimated clade divergence times are reported from (3, 4, 12, 14, 37, 50). The catarrhine-platyrrhine split that separates New World monkeys from Old World monkeys and apes is indicated at ca. 40 million years ago, and within the human lineage, the taxonomic ranks of family (Hominidae), sub-family (Homininae), and genus (*Homo*) are indicated, together with their common names: hominid, hominine, and human respectively (528). We do not report a species designation for Neanderthals because it is disputed (49). (*B-G)*, Representative images of dental calculus. (*B-C*), Mantled howler monkey dental calculus; buccal deposits of supragingival calculus show particular buildup on the premolars and molars. (*D-E)*, Gorilla dental calculus. (*D*), 'Black' plaque and calculus deposits typically present on gorilla dentitions (ABM006). (*E*), Large dental calculus deposits on the premolars and molars of EBO001. Note the presence of a synthetic adhesive used by museum staff to stabilise the dentition. (*F-G*), Chimpanzee dental calculus. f, Minor calculus deposits typically form along the gingival margin (KNP005). (*G*), Heavier calculus deposits and antemortem tooth loss observed in KNP009. (*H*), Dental calculus deposit buildup on the deciduous first molar of a 70 ka Neanderthal (GDN001); photo courtesy of Julie Arnaud, University of Ferrara. (*I*), Dental calculus deposit along the gingival margin of the first molar of the Upper Palaeolithic individual EMN001. (*J*), Dental calculus deposits on the molars of a Magdalenian individual (RIG001); photo courtesy of Wolfgang Gerber, University of Tübingen. Scale bars represent 1 mm in (*H*) and 1 cm in (*I*) and (*J*).

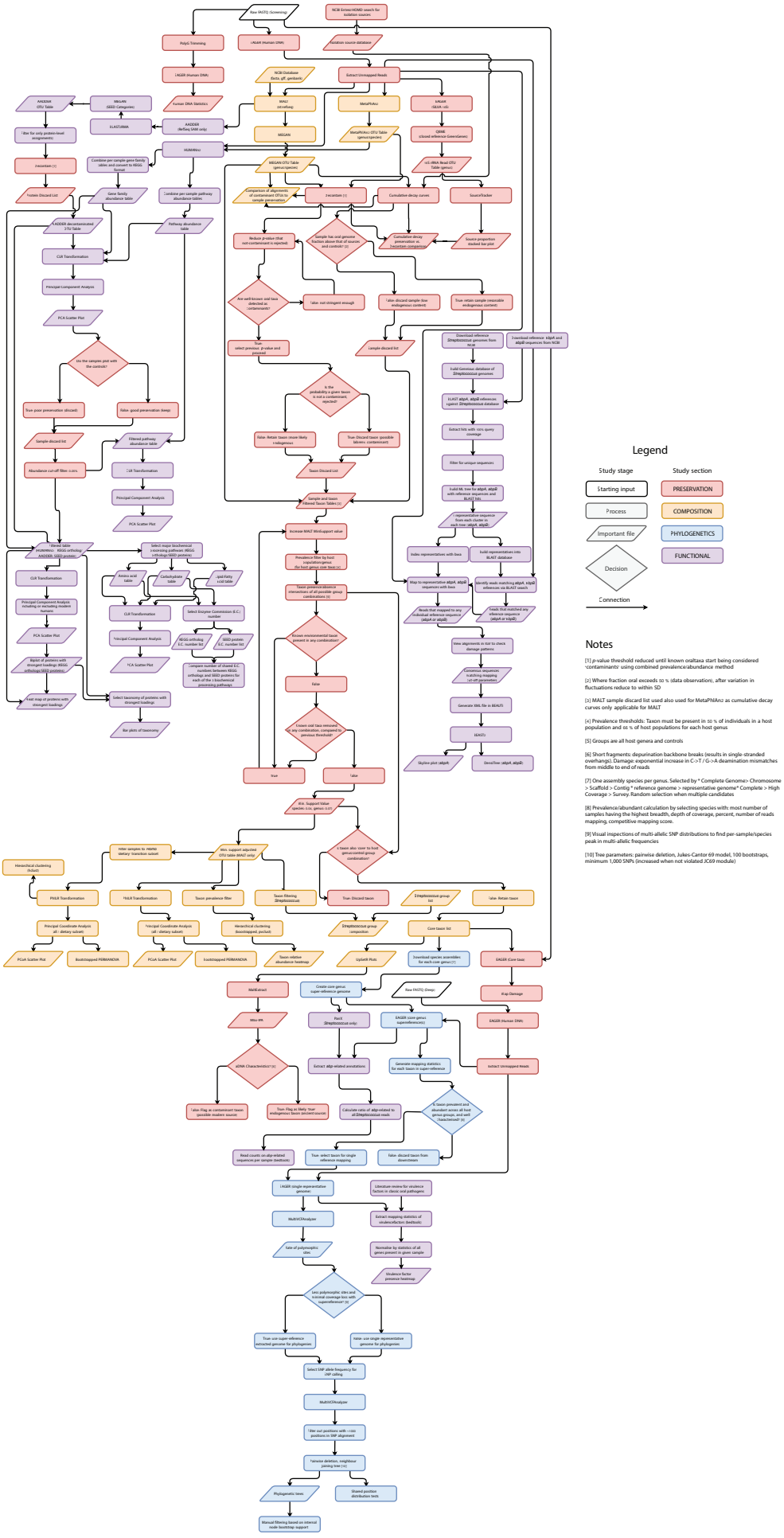

Fig. S2. Flowchart of study processing, authentication and analysis procedures.

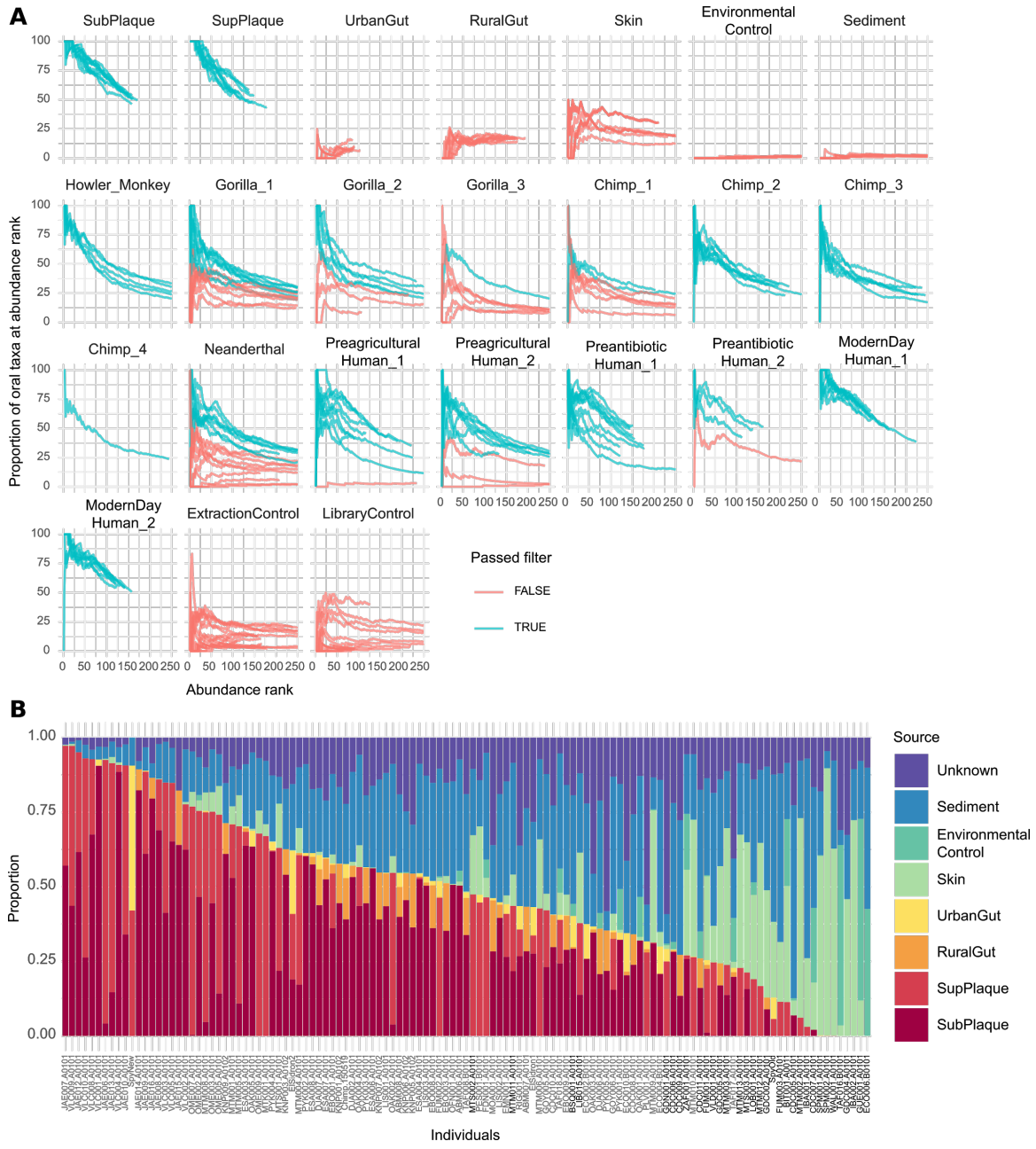

Fig. S3. Data quality and authentication filters applied to dental calculus samples in this study. (*A*), Cumulative percent decay plots showing the fraction of oral taxa relative to all taxa, as ordered by abundance rank (most to least number of alignments). X-axis restricted to 250 positions for visualization purposes. Modern reference datasets are shown in the first row for comparison. See methods for filter description. Species designations for each host group can be seen in Data S1. (*B*), Estimated proportions of a sample resembling a given source, as estimated by SourceTracker (223) from 16S reads taxon identification across all calculus samples. Sample labels signify whether the sample passed (grey) or failed (black) in (*A*).

Fig. S4. Representative ancient authenticity metrics for Neanderthals from Pesturina (PES001) and de Nadale (GDN001) caves. Damage mismatch profiles, read length, edit distance, and sequence identity are shown for four selected oral bacterial species: *Fretibacterium fastidiosum*, *Fusobacterium nucleatum*, *Tannerella forsythia*, *Treponema denticola*. Total mapping reads are shown in green (default), and mapping reads with damage are shown in purple. Plots were produced using the reporting function of the MaltExtract Interactive Plotting Application (MEx-IPA) from the MaltExtract tool of the HOPS pipeline (205). See SI Appendix section S5.3 for location of HOPS output of the entire dataset.

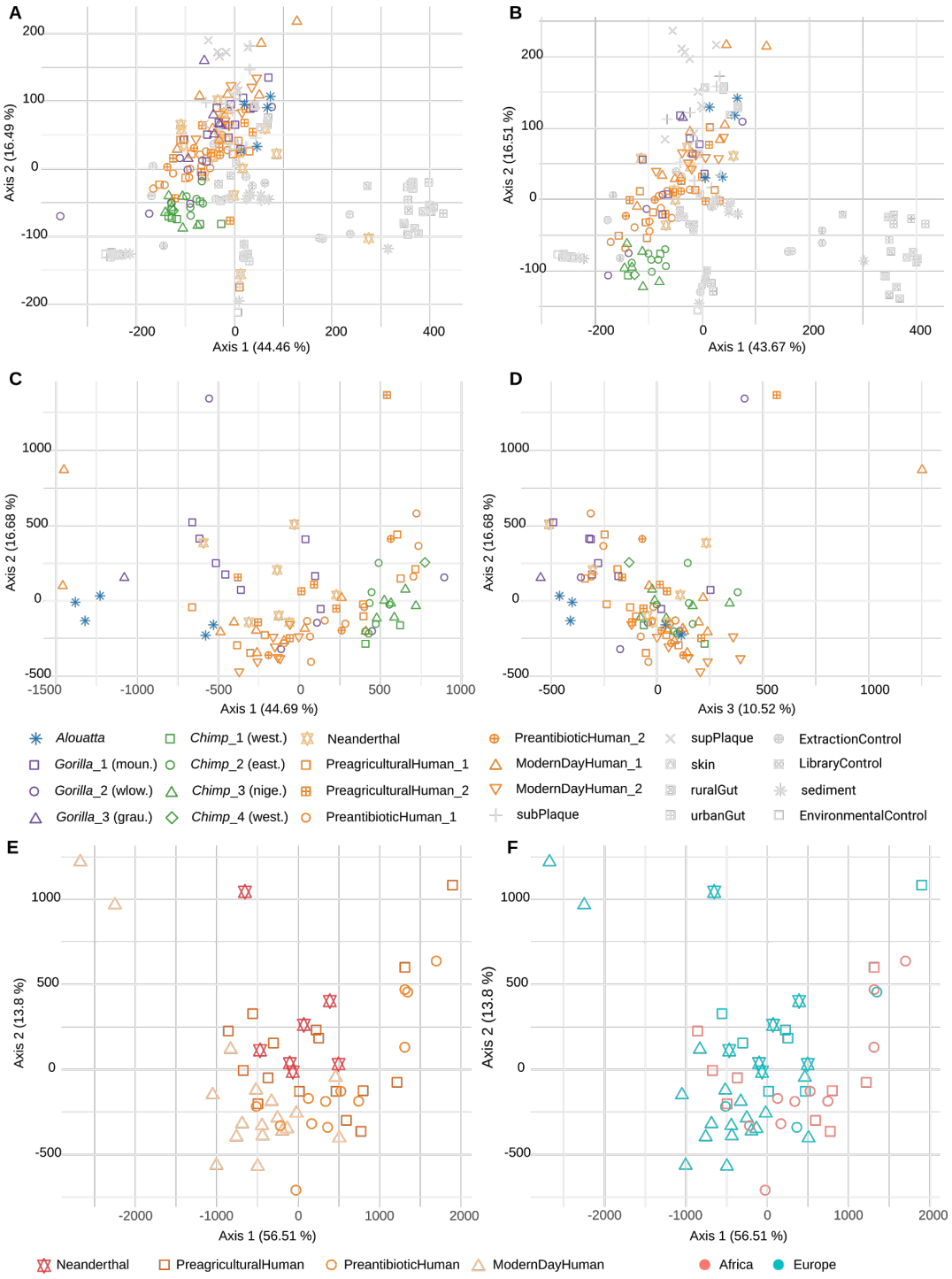

Fig. S5. Principal Coordinates Analysis (PCoA) of microbial communities in dental calculus and sources. (*A-B*), PCA of microbial communities within dental calculus and sources for (*A*), all dental calculus samples and (*B*), with poorly preserved samples removed. (*C-D*), PCA of microbial communities without sources and including only well-preserved dental calculus; species designations for each group can be seen in Data S1. Abbreviations for *Pan* and *Gorilla* groups refer to sub-species common names. (*E-F*), PCA of *Homo* dental calculus, categorised by (*E*), time period and region (*F*), of origin. There is a high degree of overlap among dental calculus microbial communities of all *Homo* groups. See SI Appendix section S4.1 for data input and transformation details.

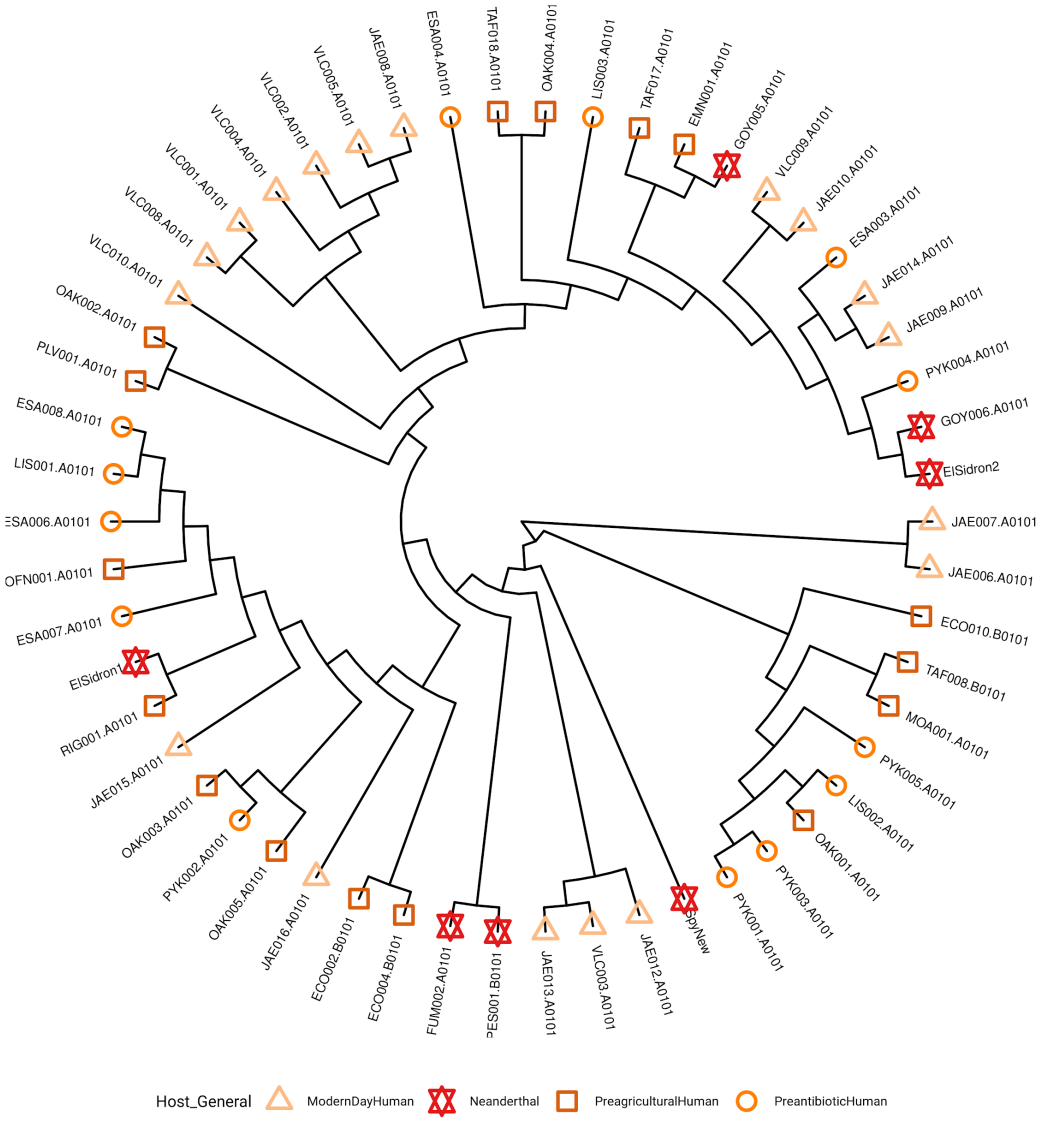

Fig. S6. Hierarchical clustering of different Homo calculus microbiomes, comparing different lifestyles and regions. We do not observe clustering of calculus microbiomes of individuals from Homo by broad dietary differences. Input is a genus level PhILR transformed OTU table without sources and controls, and low preservation samples removed. Clustering was performed with the average-linkage algorithm. Low abundant taxa are removed if under 0.01% of overall alignments ('min. support'), and putative laboratory contaminants removed as per the decontam R package.

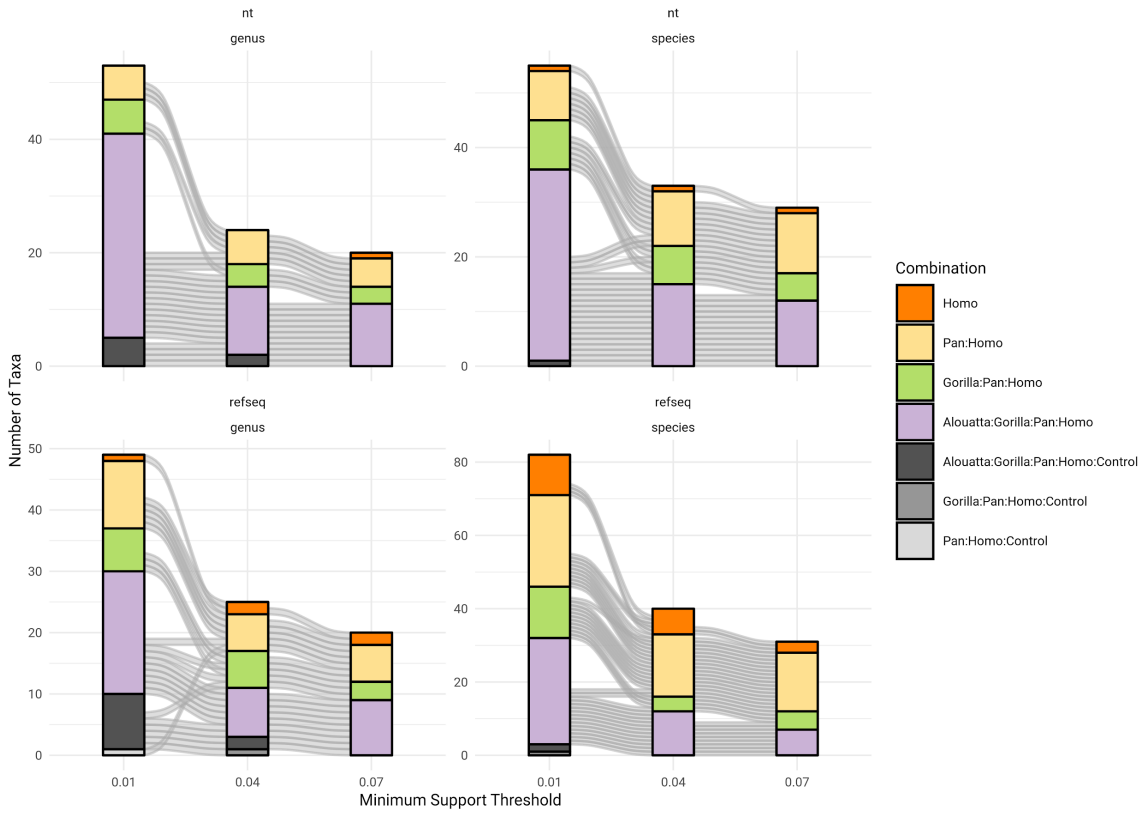

Fig S7. Alluvial diagram showing effects of increasing the minimum abundance threshold to the MALT OTU table-based core microbiome calculations. Increasing from 0.04% to 0.07% shows minimal changes in combination assignment. Comparisons are between the nt (top) and RefSeq (bottom) databases, and at genus (left) and species (right) taxonomic levels. Stacked bars represent the number of taxa to each combination, and alluviums represent the assignment of a given taxon between each minimum support threshold. Plots created using the ggalluvial R package 273, with input data as MALT aligned and MEGAN exported OTU tables excluding putative laboratory contaminants, badly preserved samples, and taxa with minimum support values < 0.07% (genus level) and < 0.04% (species level).

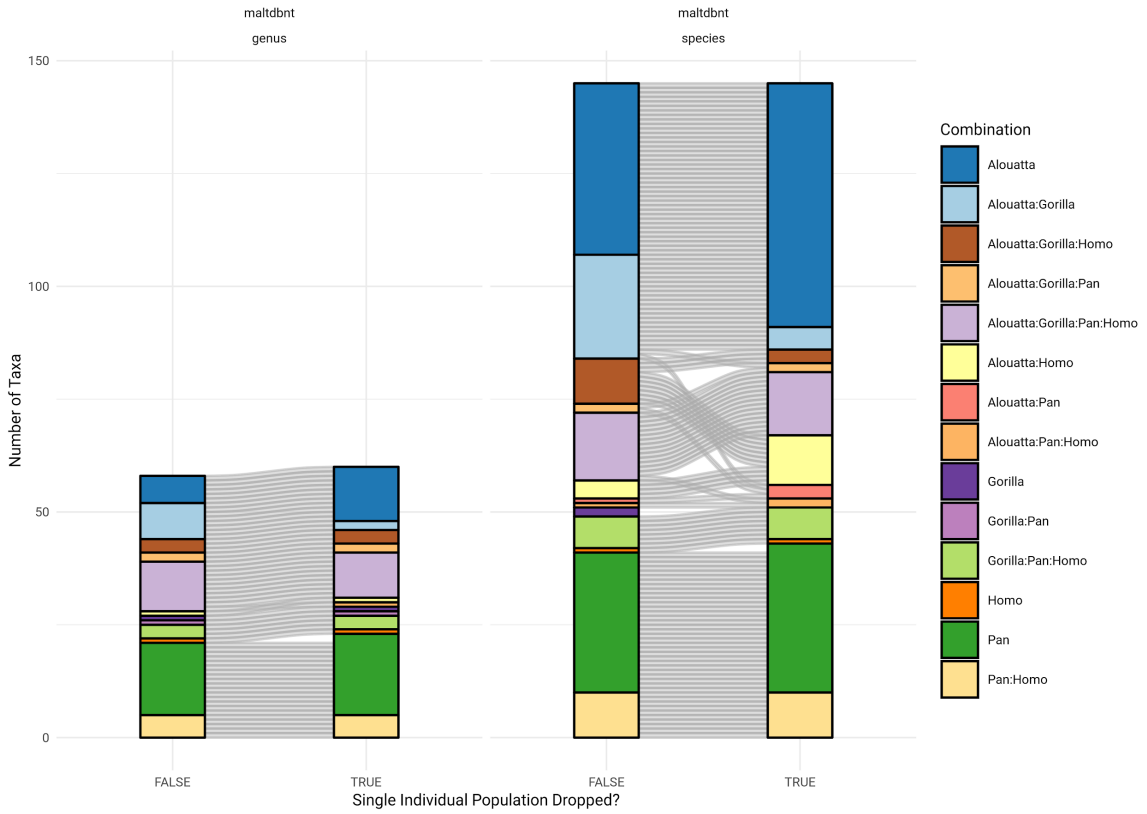

Fig S8. Alluvial diagram showing effects of dropping and retaining a single-individual Gorilla population in core microbiome calculations at genus and species taxonomic levels. Dropping the single-individual population results in minor combination assignments, mostly taxa being assigned to being core to the compositionally similar Alouatta combinations. Stacked bars represent the number of taxa to each combination, and alluviums represent the assignment of a given taxon between dataset. Plots created using the ggalluvial R package, with input as MALT NCBI nt aligned and MEGAN exported OTU tables with putative laboratory contaminants, badly preserved samples, and taxa with minimum support values < 0.07% (genus level) and < 0.04% (species level) removed.

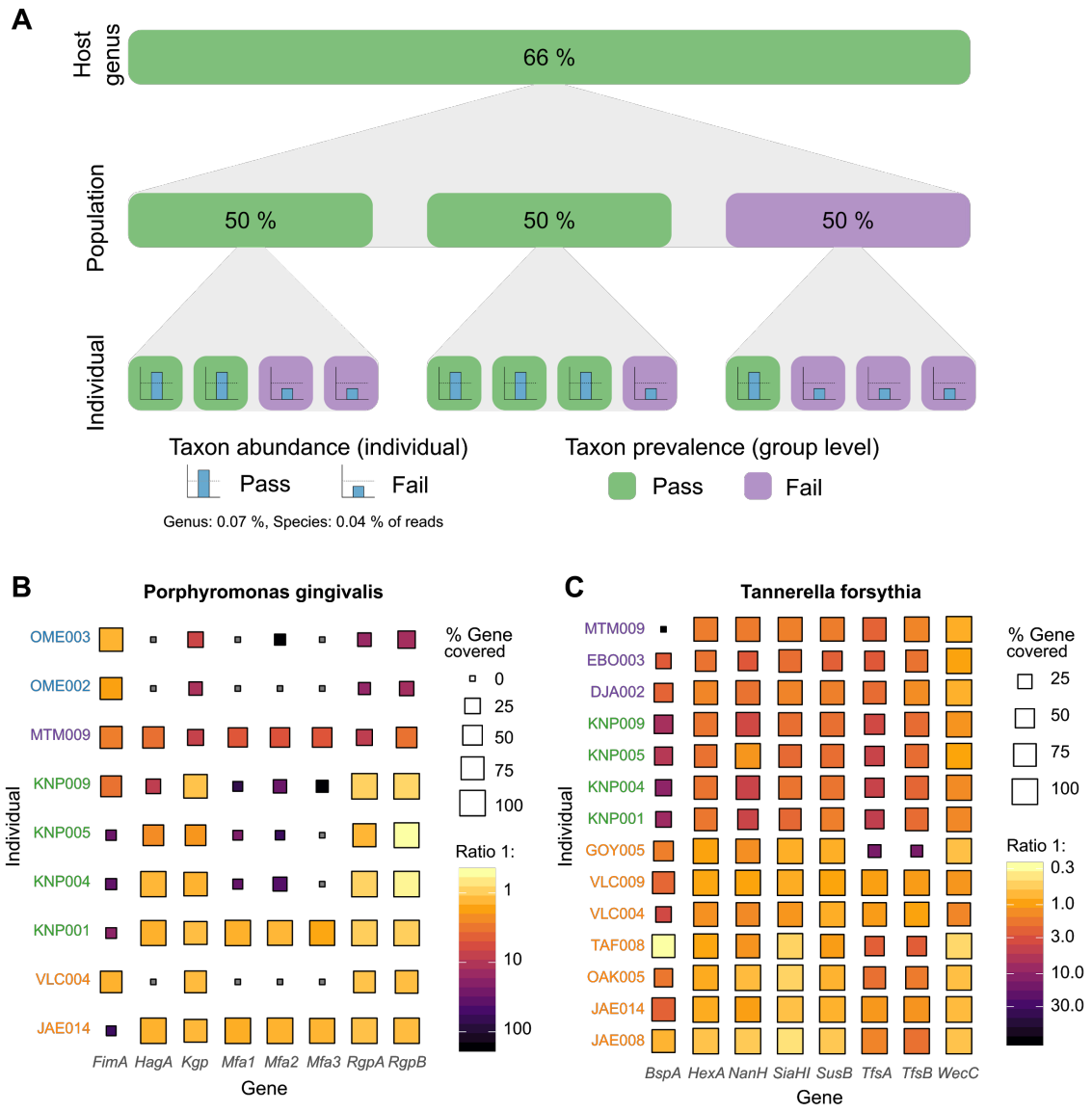

Fig. S9. Core microbiome strategy and virulence genes. (*A*), Selection criteria for core microbiome taxa. The outgroup howler monkeys (Alouatta) was exempted from the population level parameter because only one population was sampled. (*B-C*), Heatmaps of virulence gene coverage and completeness for two 'red complex' bacteria reference genomes: (*B*), P. gingivalis and (*C*), T. forsythia.Label colours correspond to host genus as in previous figures - Blue: Alouatta, Purple: Gorilla, Green: Pan, Orange: Homo. All virulence genes for both P. gingivalis and T. forsythia were observed in members of all great ape species, indicating a deep evolutionary history of these genes.

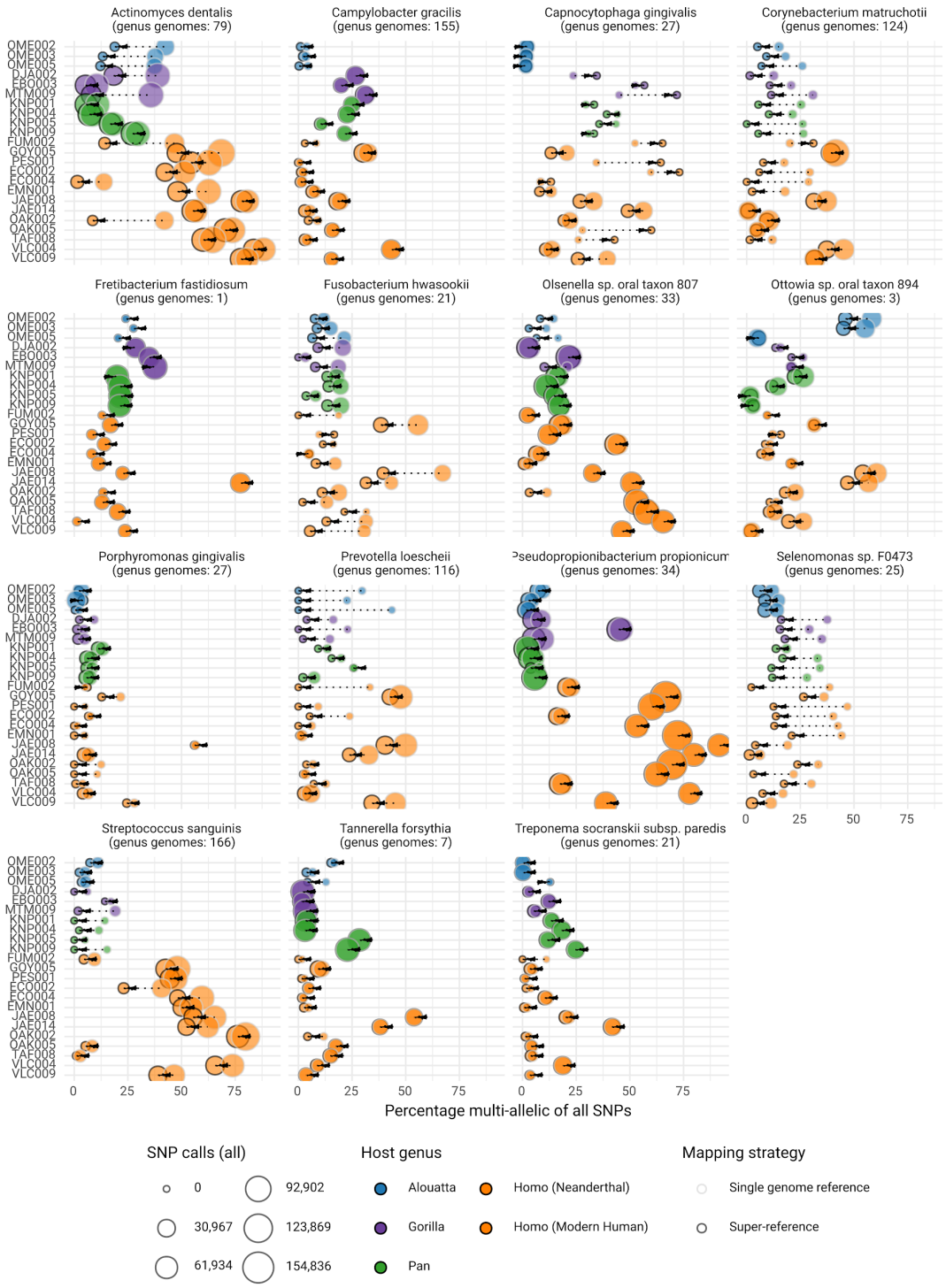

Fig. S10. Comparison of the number of multi-allelic single nucleotide polymorphisms (SNPs) identified when using a single representative genome mapping strategy versus a multi-reference genome (super-reference) mapping approach. Arrows indicate direction of change in the percentage of multi-allelic sites between single genome (grey circle outline) and super-reference (black circle outline) mapping strategies, where the expected direction is from right to left, indicating a reduction in multi-allelic sites. Colours refer to the host genus. The expected reduction in the percentage of multi-allelic SNPs (from right to left on the x-axis) between the two strategies is not consistently observed.

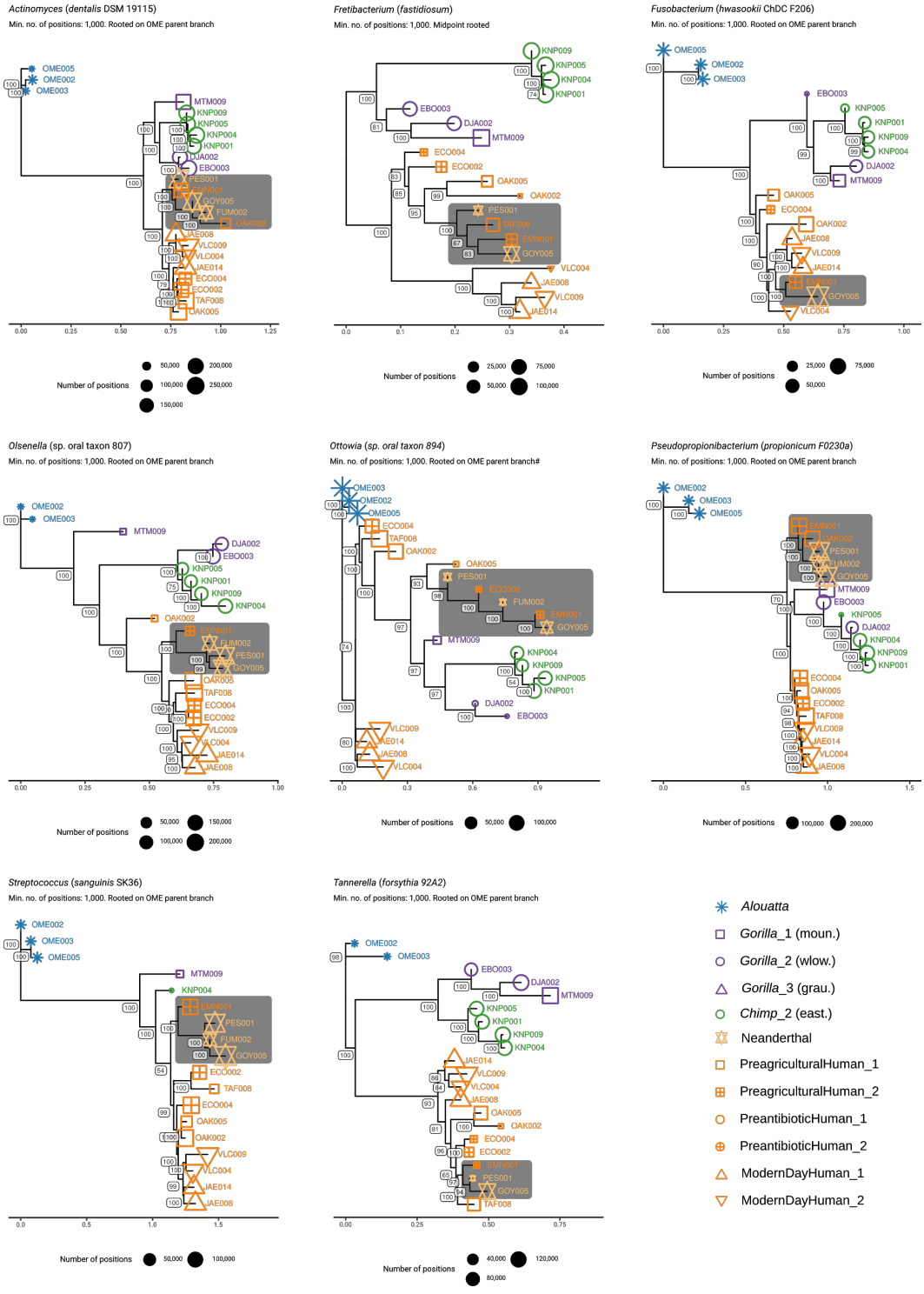

Fig. S11. Neighbour joining trees of eight well-supported core calculus taxa from single representative mappings of the deep sequenced dataset. Microbial genomic diversity generally clusters individuals with those of the same host genus. Pre-14 ka European individuals consistently group together to the exclusion of most post-14 ka European individuals. Grey boxes highlight European pre-14 ka BP individuals, as represented by El Mirón and Neanderthals. Titles refer to the *representative* genome of the selected species used for the reference genome, and the alignments are mixtures of strains/species as indicated by high levels of multi-allelic sites.

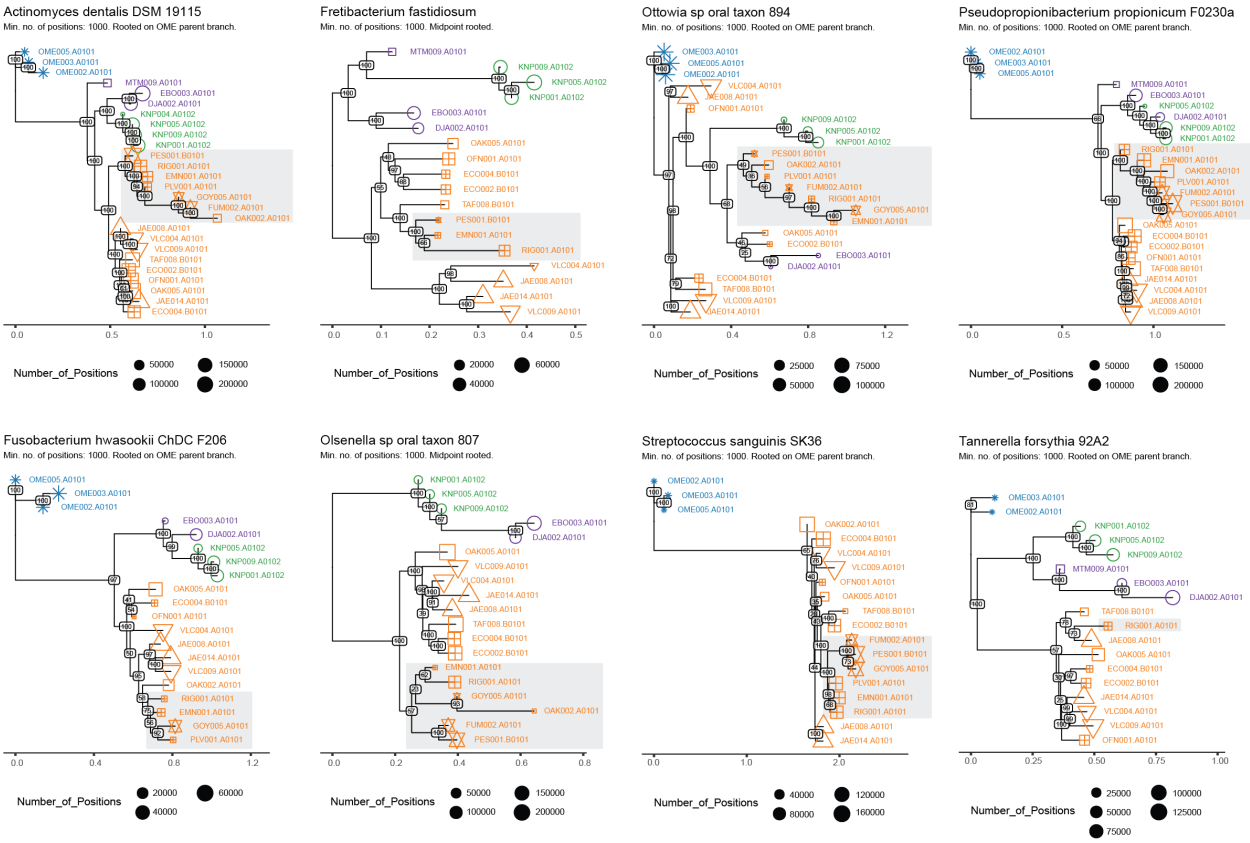

Fig. S12. Replication of production dataset phylogenies with low-coverage and damage-containing screening dataset with additional European individuals. The observed pattern of pre-14k BP Europeans and post-14k BP humans clustering separately in the production dataset phylogenies is replicated when including additional pre- and post-14k BP individuals when using screening dataset equivalents. Grey boxes indicate European 'pre-14k BP' of the Red Lady of El Mirón and Neanderthals. Representative genomes were selected based on abundance and prevalence across all individuals in the production dataset. SNPs were called using MultiVCFAnalyzer with a minimum coverage threshold of 2, and the majority allele threshold of 0.7. Alignments with <1000 called SNPs were removed. Genetic distance calculated using the ape R package, with the Jukes-Cantor 69 model and pairwise deletion strategy for missing data. Bootstraps are out of 100 bootstraps.

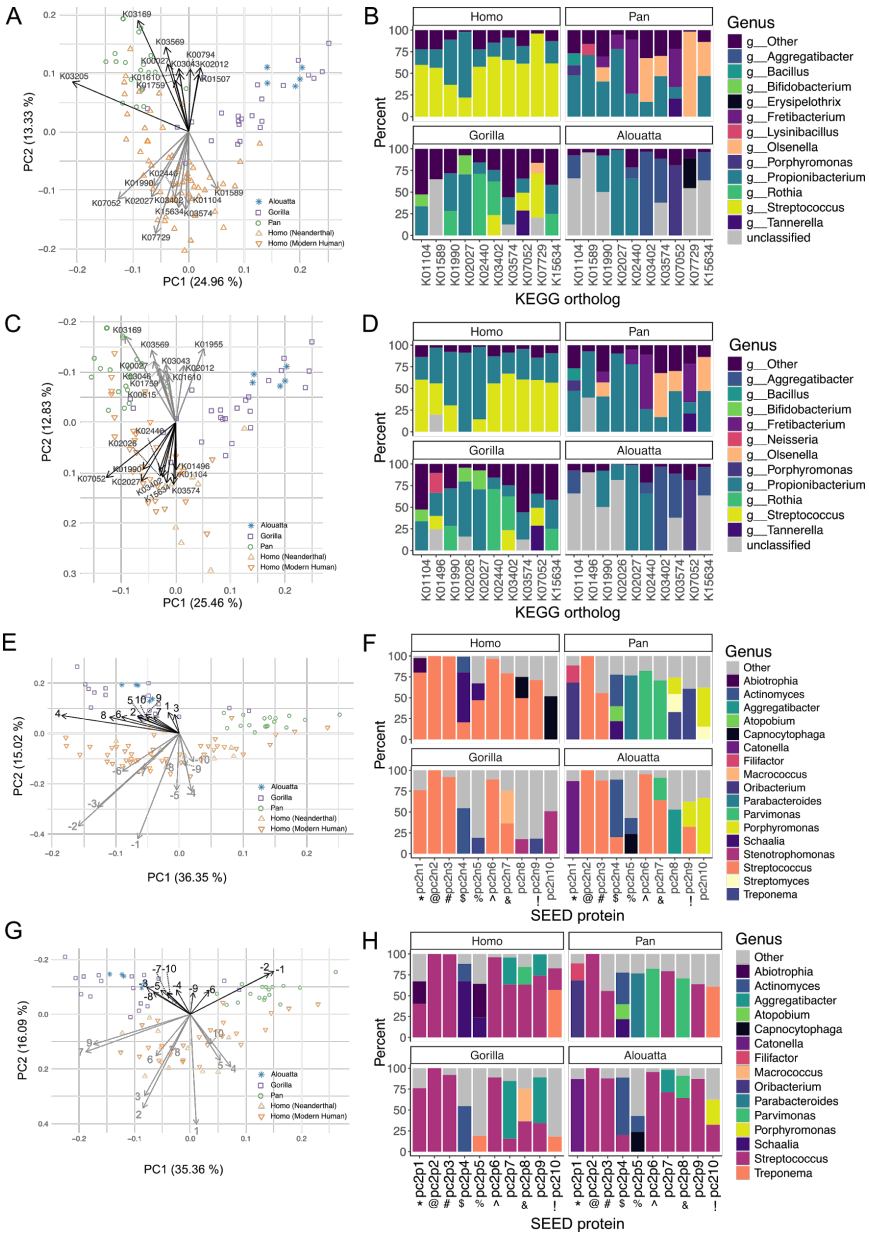

Fig. S13. Principal Components Analysis (PCA) biplots of functional annotations with the top 10 positive and negative loadings separating Homo from the other hosts. Samples were screened to gain a gene content profile by both HUMAnN2, which annotated genes using the KEGG ortholog database, and by AADDER, which annotated the genes using the SEED database. PCA biplot showing top-loading KEGG orthologs (*A,C*), or SEED-identified proteins (*E,G*) in PC2 including (*A,E*), or excluding (*C,G*) present-day modern humans. Bar graphs of the genera, summed from species, (*B,D*), that contribute to the KEGG orthologs or (*F,H*) the SEED-identified proteins, with strongest negative loadings in PC2, including present-day modern humans (*B,F*), or excluding present-day modern humans (*D,H*). All genera that individually contributed <12% of KEGG orthologs or <15% of SEED-identified proteins are grouped together as 'Other'. Protein names correspond to the numbers in the section S7, where for SEED proteins 'pc1/pc2' indicate the component, 'p/n' indicate positive/negative direction, and the final number indicates the protein indicated on the biplot. SEED-identified proteins that are the same between the PCAs including and excluding present-day modern humans are indicated with symbols (i.e., the '&' in panel (*F*) pc2n7 and panel (*H*) pc2p8 indicates these are the same protein).

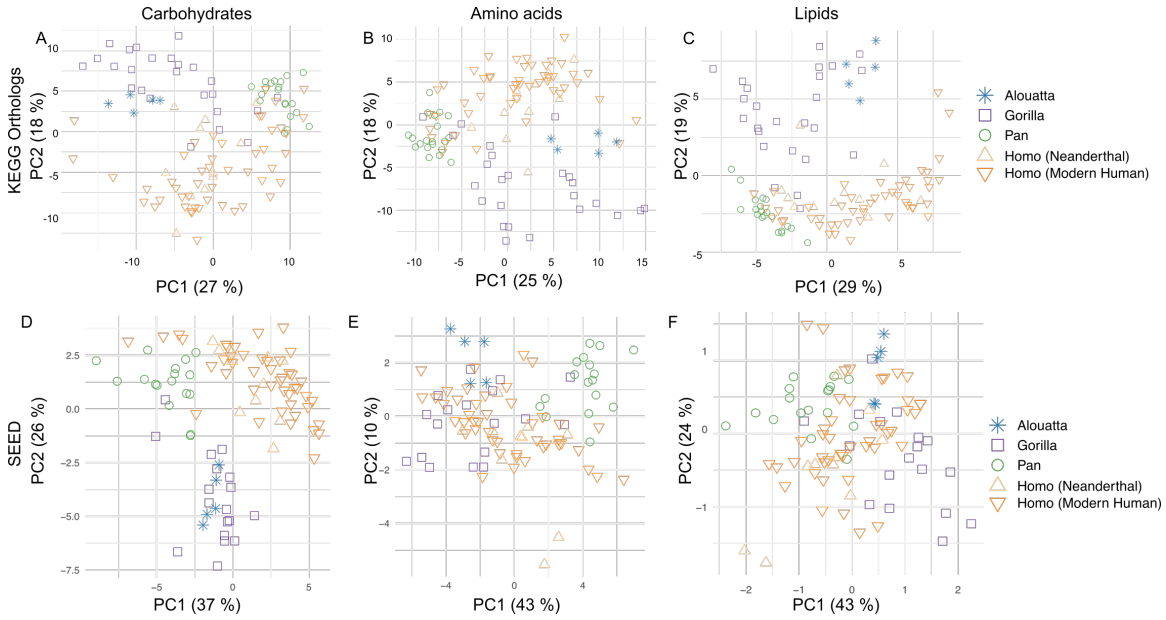

Fig. S14. Principal Components Analysis (PCA) using KEGG orthologs or SEED-classified proteins belonging to specific metabolic pathway categories. KEGG orthologs (*A*-*C*), and SEED-classified proteins (*D*-*F*), in pathways that process major biomolecules. KEGG orthologs in only the (*A*) Carbohydrate metabolism pathways, (*B*) Amino acid metabolism pathways, and the (*C*) Lipid metabolism pathways (permanova R^2 = 0.291, 0.285, 0.319, respectively). Only SEED proteins in the (*D*) Carbohydrates category separate the samples by host genera in a pattern similar to that seen for taxonomy and all proteins (carbohydrate pathway proteins permanova R^2 = 0.423 vs all proteins permanova R^2 = 0.345. SEED proteins in the (*E*) Amino acid and (*F*) Fatty acids, Lipids, and Isoprenoids categories do not separate samples by host genera as distinctly (permanova R^2 = 0.289 and R^2 = 0.258, respectively).

Table S1. Summary of host species and dental calculus samples included in study

Sample source	Origin	Period	Code	Total assessed individuals	Pass quality filters	Individuals passing quality filters[a]
Howler Monkey				**5**	**5**	
Alouatta palliata (Mantled howler monkey)						
Ometepe Island	Nicaragua	20th-21st century	OME	5	5	**OME002**; **OME003**; **OME005**; OME007; OME009
Gorilla				**29**	**14**	
Gorilla beringei beringei (Mountain gorilla)						
Mount Sabyinyo	DR Congo	20th century	MTS	3	1	MTS001
Mount Karisimbi	Rwanda	20th century	MTK	1	1	MTK001
Mount Mikeno	DR Congo	20th century	MTM	11	6	MTM001; MTM004; MTM005; MTM006; MTM008; **MTM009**
Gorilla beringei graueri (Grauer's gorilla)						
Lubero	DR Congo	20th century	LOB	1	0	-
Bitule	DR Congo	20th century	BIT	1	0	-
Foret de Maniema	DR Congo	20th century	FDM	1	1	FDM001
Ibatsero	DR Congo	20th century	IBA	2	0	-
Walikale	DR Congo	20th century	WAL	1	0	-
Gorilla gorilla gorilla (Western lowland gorilla)						
Abong Mbong	Cameroon	20th century	ABM	1	1	ABM006
Djaposten	Cameroon	20th century	DJA	1	1	**DJA002**
Ebolwa	Cameroon	20th century	EBO	2	2	EBO001; **EBO003**
Mount Mikeno[b]	DR Congo	20th century	MTM	1	1	MTM010
Not specified	French Congo	20th century	GDC	3	0	-
Chimpanzee				**21**	**16**	
Pan troglodytes schweinfurthii (Eastern chimpanzee)						
Kibale National Forest	Uganda	20th-21st century	KNP	7	7	**KNP001**; KNP003; **KNP004**; **KNP005**; KNP006; KNP007; **KNP009**
Pan troglodytes ellioti (Nigeria-Cameroon chimpanzee)						
Abong Mbong	Cameroon	20th century	ABM	2	2	ABM007; ABM008
Djaposten	Cameroon	20th century	DJA	2	2	DJA006; DJA007
Ebolwa	Cameroon	20th century	EBO	2	2	EBO008; EBO009
Pan troglodytes verus (Western chimpanzee)						
Gola Forest[c]	Sierra Leone	Historic	-	1	1	Chimp.150519
Not specified	Liberia	20th century	CDC	7	2	CDC008; CDC010
Neanderthal				**17**	**7**	
Homo (Neanderthal)[d]						
Banyoles	Spain	Middle Palaeolithic	BAN	1	0	-
La Güelga	Spain	Middle Palaeolithic	GUE	1	0	-
Cueva de Valdegoba	Spain	Middle Palaeolithic	VLD	1	0	-
El Sidrón[c]	Spain	Middle Palaeolithic	ElSidron	2	2	ElSidron1; ElSidron2
Cueva del Boquete de Zafarraya	Spain	Middle Palaeolithic	ZAF	1	0	-
Sima de las Palomas del Cabezo Gordo	Spain	Middle Palaeolithic	SPM	2	0	-
Pĕsturina Cave	Serbia	Middle Palaeolithic	PES	1	1	**PES001**
Fumane	Italy	Middle Palaeolithic	FUM	3	1	**FUM002**
de Nadale	Italy	Middle Palaeolithic	GDN	1	0	-
Troisième caverne of Goyet	Belgium	Mid./Upp. Palaeolithic	GOY	2	2	**GOY005**; GOY006
Spy[c]	Belgium	Mid./Upp. Palaeolithic	Spy	2	1	SpyNew

Table S1 (continued)

Sample	Origin	Period	Code	Total assessed individuals	Pass quality filters	Individuals passing quality filters
Modern human				**52**	**47**	
Homo sapiens (modern human, pre-agricultural)						
Oakhurst	South Africa	Later Stone Age	OAK	5	5	OAK001; **OAK002**; OAK003; OAK004; **OAK005**
Taforalt	Morocco	Later Stone Age	TAF	4	3	**TAF008**; TAF017; TAF018
Mota	Ethiopia	Later Stone Age	MOA	1	1	MOA001
Pavlov	Czech Republic	Upper Palaeolithic	PLV	1	1	PLV001
Dolní Věstonice	Czech Republic	Upper Palaeolithic	DLV	2	0	-
El Mirón	Spain	Upper Palaeolithic	EMN	1	1	**EMN001**
Rigney	France	Upper Palaeolithic	RIG	1	1	RIG001
El Collado	Spain	Mesolithic	ECO	4	3	**ECO002**; **ECO004**; ECO010
Grosse Ofnet Höhle	Germany	Mesolithic	OFN	1	1	OFN001
Homo sapiens (modern human, pre-antibiotic)						
Essaouira (Mogador)	Morocco	18th century	ESA	5	5	ESA003; ESA004; ESA006; ESA007; ESA008
Polyoak	South Africa	19th century	PYK	5	5	PYK001; PYK002; PYK003; PYK004; PYK005
Lisbon	Portugal	18th century	LIS	3	3	LIS001; LIS002; LIS003
Basque country (modern day region of)	Spain	1st-2nd century	BSQ	1	0	BSQ001
Homo sapiens (modern human, present day)						
Valencia	Spain	21th century	VLC	8	8	VLC001; VLC002; VLC003; **VLC004**; VLC005; VLC008; **VLC009**; VLC010
Jaen[e]	Spain	21th century	JAE	10	10	JAE006; JAE007; **JAE008**; JAE009; JAE010; JAE012; JAE013; **JAE014**; JAE015; JAE016

Notes:
[a]Samples also sequenced for the production dataset are indicated in **bold**.
[b]Although initially recorded as *Gorilla beringei*, this individual was genetically confirmed to be *Gorilla gorilla gorilla* (van der Valk, T. and Guschanski, K.).
[c]Previously published in: L. S. Weyrich et al., Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature*. 544, 357–361 (2017). [doi: 10.1038/nature21674]
[d]Because the species assignment of Neanderthals is disputed (*Homo neanderthalensis* vs. *Homo sapiens neanderthalensis*), we do not provide a species designation. See: C. Stringer, Evolution: What makes a modern human. *Nature*. 485, 33–35 (2012) [doi: 10.1038/485033a].
[e]Previously published in: I. M. Velsko, et al., Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. Microbiome. 7, 102 (2019). [doi: 10.1186/s40168-019-0717-3].

SI References

1. L. S. Weyrich, *et al.*, Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature* **544**, 357–361 (2017).

2. I. M. Velsko, *et al.*, Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. *Microbiome* **7**, 102 (2019).

3. C. G. Schrago, On the time scale of New World primate diversification. *Am. J. Phys. Anthropol.* **132**, 344–354 (2007).

4. M. Bond, *et al.*, Eocene primates of South America and the African origins of New World monkeys. *Nature* **525**, 552 (2015).

5. A. Estrada, S. Juan-Solano, T. Ortíz Martínez, R. Coates-Estrada, Feeding and general activity patterns of a howler monkey (Alouatta palliata) troop living in a forest fragment at Los Tuxtlas, Mexico. *Am. J. Primatol.* **48**, 167–183 (1999).

6. R. S. Scott, M. F. Teaford, P. S. Ungar, Dental microwear texture and anthropoid diets. *Am. J. Phys. Anthropol.* **147**, 551–579 (2012).

7. F. C. Espinosa-Gómez, *et al.*, Salivary tannin-binding proteins are a pervasive strategy used by the folivorous/frugivorous black howler monkey. *Am. J. Primatol.* **80**, e22737 (2018).

8. F. Espinosa Gómez, *et al.*, Howler Monkeys (Alouatta palliata mexicana) Produce Tannin-Binding Salivary Proteins. *Int. J. Primatol.* **36**, 1086–1100 (2015).

9. K. Milton, Physiological Ecology of Howlers (Alouatta): Energetic and Digestive Considerations and Comparison with the Colobinae. *Int. J. Primatol.* **19**, 513–548 (1998).

10. M. C. Janiak, Digestive enzymes of human and nonhuman primates. *Evol. Anthropol.* **25**, 253–266 (2016).

11. W. B. Hall, H. E. Grupe, C. K. Claycomb, The periodontium and periodontal pathology in the howler monkey. *Arch. Oral Biol.* **12**, 359–366 (1967).

12. J. Prado-Martinez, *et al.*, Great ape genetic diversity and population history. *Nature* **499**, 471–475 (2013).

13. IUCN, Gorilla beringei: Plumptre, A., Robbins, M.M. & Williamson, E.A.: The IUCN Red List of Threatened Species 2019: e.T39994A115576640 (2018) https:/doi.org/10.2305/IUCN.UK.2019-1.RLTS.T39994A115576640.en.

14. M. W. Tocheri, *et al.*, The evolutionary origin and population history of the grauer gorilla. *Am. J. Phys. Anthropol.* **159**, S4–S18 (2016).

15. A. Elgart-Berry, Fracture toughness of mountain gorilla (Gorilla gorilla beringei) food plants. *Journal of Primatology: Official Journal of the …* (2004).

16. K. Guschanski, *et al.*, Counting elusive animals: Comparing field and genetic census of the entire mountain gorilla population of Bwindi Impenetrable National Park, Uganda. *Biol.*

Conserv. **142**, 290–300 (2009).

17. M. Gray, *et al.*, Genetic census reveals increased but uneven growth of a critically endangered mountain gorilla population. *Biol. Conserv.* **158**, 230–238 (2013).

18. T. van der Valk, *et al.*, Significant loss of mitochondrial diversity within the last century due to extinction of peripheral populations in eastern gorillas. *Sci. Rep.* **8**, 6551 (2018).

19. Hickey, JR Uzabaho, E Akantorana, M Arinaitwe, J Bakebwa, I Bitariho, R Eckardt, W Gilardi, K Katutu, J Kayijamahe, C Kierepka, EM Mugabukomeye, B Musema, A Mutabaazi, H Robbins, MM Sacks, BN Zikusoka, GK, "Bwindi-Sarambwe 2018 Surveys Monitoring Mountain Gorillas, Other Select Mammals, and Human Activities" (GVTC, IGCP & Partners, 2019).

20. A. Elgart-Berry, Fracture toughness of mountain gorilla (Gorilla gorilla beringei) food plants. *Am. J. Primatol.* **62**, 275–285 (2004).

21. J. Yamagiwa, N. Mwanza, T. Yumoto, T. Maruhashi, Seasonal change in the composition of the diet of eastern lowland gorillas. *Primates* **35**, 1–14 (1994).

22. N. C. Lovell, *Patterns of injury and illness in great apes: a skeletal analysis* (Smithsonian Institution Press, 1990).

23. J. C. Brealey, *et al.*, A roadmap to mammalian oral microbiome evolution with dental calculus. *bioRxiv*, 596791 (2019).

24. A. A. Elgart, Dental wear, wear rate, and dental disease in the African apes. *Am. J. Primatol.* **72**, 481–491 (2010).

25. T. Smiley, *et al.*, Noninvasive saliva collection techniques for free-ranging mountain gorillas and captive eastern gorillas. *J. Zoo Wildl. Med.* **41**, 201–209 (2010).

26. T. van der Valk, D. Díez-Del-Molino, T. Marques-Bonet, K. Guschanski, L. Dalén, Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas. *Curr. Biol.* **29**, 165–170.e6 (2019).

27. IUCN, Gorilla gorilla: Maisels, F., Bergl, R.A. & Williamson, E.A.: The IUCN Red List of Threatened Species 2018: e.T9404A136250858 (2016) https:/doi.org/10.2305/IUCN.UK.2018-2.RLTS.T9404A136250858.en.

28. M. E. Rogers, *et al.*, Western gorilla diet: a synthesis from six sites. *Am. J. Primatol.* **64**, 173–192 (2004).

29. D. Doran-Sheehy, P. Mongo, J. Lodwick, N. L. Conklin-Brittain, Male and female western gorilla diet: Preferred foods, use of fallback resources, and implications for ape versus old world monkey foraging strategies. *American Journal of Physical Anthropology* **140**, 727–738 (2009).

30. M. J. Remis, Western lowland gorillas (Gorilla gorilla gorilla) as seasonal frugivores: use of variable resources. *Am. J. Primatol.* **43**, 87–109 (1997).

31. S. Thamadilok, *et al.*, Human and Nonhuman Primate Lineage-Specific Footprints in the Salivary Proteome. *Mol. Biol. Evol.* (2019) https:/doi.org/10.1093/molbev/msz223.

32. R. E. McDonald, *et al.*, Latherin: a surfactant protein of horse sweat and saliva. *PLoS One* **4**, e5726 (2009).

33. G. M. Wilson, *et al.*, Identification by full-coverage array CGH of human DNA copy number increases relative to chimpanzee and gorilla. *Genome Res.* **16**, 173–181 (2006).

34. P. Pajic, *et al.*, Independent amylase gene copy number bursts correlate with dietary preferences in mammals. *Elife* **8** (2019).

35. V. Behringer, *et al.*, Measurements of salivary alpha amylase and salivary cortisol in hominoid primates reveal within-species consistency and between-species differences. *PLoS One* **8**, e60773 (2013).

36. G. da Costa, *et al.*, Salivary Amylase Induction by Tannin-Enriched Diets as a Possible Countermeasure Against Tannins. *J. Chem. Ecol.* **34**, 376 (2008).

37. M. de Manuel, *et al.*, Chimpanzee genomic diversity reveals ancient admixture with bonobos. *Science* **354**, 477–481 (2016).

38. G. Hohmann, *et al.*, Plant foods consumed by Pan: exploring the variation of nutritional ecology across Africa. *Am. J. Phys. Anthropol.* **141**, 476–485 (2010).

39. R. W. Wrangham, N. L. Conklin-Brittain, K. D. Hunt, Dietary Response of Chimpanzees and Cercopithecines to Seasonal Variation in Fruit Abundance. I. Antifeedants. *Int. J. Primatol.* **19**, 949–970 (1998).

40. J. K. Matthews, A. Ridley, P. Niyigaba, B. A. Kaplin, C. C. Grueter, Chimpanzee feeding ecology and fallback food use in the montane forest of Nyungwe National Park, Rwanda. *Am. J. Primatol.* **81**, e22971 (2019).

41. V. Reynolds, A. J. Plumptre, J. Greenham, J. Harborne, Condensed tannins and sugars in the diet of chimpanzees (Pan troglodytes schweinfurthii) in the Budongo Forest, Uganda. *Oecologia* **115**, 331–336 (1998).

42. N. J. Dominy, E. R. Vogel, J. D. Yeakel, P. Constantino, P. W. Lucas, Mechanical Properties of Plant Underground Storage Organs and Implications for Dietary Models of Early Hominins. *Evol. Biol.* **35**, 159–175 (2008).

43. M. J. Schoeninger, H. T. Bunn, S. S. Murray, J. A. Marlett, Composition of Tubers Used by Hadza Foragers of Tanzania. *J. Food Compost. Anal.* **14**, 15–25 (2001).

44. M. C. Dean, M. E. Jones, J. R. Pilley, The natural history of tooth wear, continuous eruption and periodontal disease in wild shot great apes. *J. Hum. Evol.* **22**, 23–39 (1992).

45. L. Arnold, P. Baram, Periodontal disease in chimpanzees. *J. Periodontol.* **44**, 437–442 (1973).

46. G. H. Perry, *et al.*, Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* **39**, 1256–1260 (2007).

47. M. Mau, K.-H. Südekum, A. Johann, A. Sliwa, T. M. Kaiser, Indication of higher salivary alpha-amylase expression in hamadryas baboons and geladas compared to chimpanzees and humans. *J. Med. Primatol.* **39**, 187–190 (2010).

48. R. B. Eckhardt, R. Protsch von Zieten, Enamel hypoplasias as indicators of developmental stress in pongids and hominids. *Hum. Evol.* **8**, 93–99 (1993).

49. C. Stringer, Evolution: What makes a modern human. *Nature* **485**, 33–35 (2012).

50. P. Endicott, S. Y. W. Ho, C. Stringer, Using genetic evidence to evaluate four palaeoanthropological hypotheses for the timing of Neanderthal and modern human origins. *J. Hum. Evol.* **59**, 87–95 (2010).

51. A. R. Rogers, R. J. Bohlender, C. D. Huff, Early history of Neanderthals and Denisovans. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 9859–9863 (2017).

52. T. Higham, *et al.*, The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* **512**, 306–309 (2014).

53. P. Radović, J. Lindal, D. Mihailović, M. Roksandic, The first Neanderthal specimen from Serbia: Maxillary first molar from the Late Pleistocene of Pešturina Cave. *J. Hum. Evol.* **131**, 139–151 (2019).

54. J. Arnaud, *et al.*, A Neanderthal deciduous human molar with incipient carious infection from the Middle Palaeolithic De Nadale cave, Italy. *Am. J. Phys. Anthropol.* **162**, 370–376 (2017).

55. R. Grün, *et al.*, ESR and U-series analyses of enamel and dentine fragments of the Banyoles mandible. *J. Hum. Evol.* **50**, 347–358 (2006).

56. M. Peresani, *et al.*, Age of the final Middle Palaeolithic and Uluzzian levels at Fumane Cave, Northern Italy, using 14C, ESR, 234U/230Th and thermoluminescence methods. *J. Archaeol. Sci.* **35**, 2986–2996 (2008).

57. M. Peresani, M. Vanhaeren, E. Quaggiotto, A. Queffelec, F. d'Errico, An ochered fossil marine shell from the mousterian of fumane cave, Italy. *PLoS One* **8**, e68572 (2013).

58. T. Higham, *et al.*, Problems with radiocarbon dating the Middle to Upper Palaeolithic transition in Italy. *Quat. Sci. Rev.* **28**, 1257–1267 (2009).

59. H. Rougier, *et al.*, Neandertal cannibalism and Neandertal bones used as tools in Northern Europe. *Sci. Rep.* **6**, 29005 (2016).

60. R. E. Wood, *et al.*, Radiocarbon dating casts doubt on the late chronology of the Middle to Upper Palaeolithic transition in southern Iberia. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2781–2786 (2013).

61. T. Higham, European Middle and Upper Palaeolithic radiocarbon dates are often older than they look: problems with previous dates and some remedies. *Antiquity* **85**, 235–249 (2011).

62. E. Discamps, B. Gravina, N. Teyssandier, In the eye of the beholder: contextual issues for Bayesian modelling at the Middle-to-Upper Palaeolithic transition. *World Archaeol.* **47**, 601–621 (2015).

63. T. F. G. Higham, G. S. Heep, Reply to: "In the eye of the beholder: contextual issues for Bayesian modelling at the Middle-to-Upper Palaeolithic transition", by Discamps, Gravina and Teyssandier (2015). *World Archaeology* **51**, 126–133 (2019).

64. E. Discamps, B. Gravina, N. Teyssandier, Comments on Higham and Heep (2017): “Reply to: ‘In the eye of the beholder: contextual issues for Bayesian modelling at the middle-to-upper Palaeolithic transition’, by Discamps, Gravina and Teyssandier (2015).” *World Archaeology* **51**, 134–139 (2019).

65. P. Pettitt, J. Zilhão, Problematizing Bayesian approaches to prehistoric chronologies. *World Archaeol.* **47**, 525–542 (2015).

66. C. Ramsey, Bayesian Approaches to the Building of Archaeological Chronologies. *Mathematics and Archaeology*, 272–292 (2015).

67. K. Prüfer, *et al.*, A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017).

68. V. Slon, *et al.*, The genome of the offspring of a Neanderthal mother and a Denisovan father. *Nature* **561**, 113 (2018).

69. K. Prüfer, *et al.*, The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).

70. A. B. Wolf, J. M. Akey, Outstanding questions in the study of archaic hominin admixture. *PLoS Genet.* **14**, e1007349 (2018).

71. F. A. Villanea, J. G. Schraiber, Multiple episodes of interbreeding between Neanderthal and modern humans. *Nat Ecol Evol* **3**, 39–44 (2019).

72. P. Skoglund, I. Mathieson, Ancient Genomics of Modern Humans: The First Decade. *Annu. Rev. Genomics Hum. Genet.* **19**, 381–404 (2018).

73. M. C. Stiner, “Middle Paleolithic Subsistence Ecology in the Mediterranean Region” in *Transitions Before the Transition: Evolution and Stability in the Middle Paleolithic and Middle Stone Age*, E. Hovers, S. L. Kuhn, Eds. (Springer US, 2006), pp. 213–231.

74. M. P. Richards, E. Trinkaus, Isotopic evidence for the diets of European Neanderthals and early modern humans. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16034–16039 (2009).

75. L. Fiorenza, *et al.*, To meat or not to meat? New perspectives on Neanderthal ecology. *Am. J. Phys. Anthropol.* **156 Suppl 59**, 43–71 (2015).

76. I. Fiore, *et al.*, From feathers to food: Reconstructing the complete exploitation of avifaunal resources by Neanderthals at Fumane cave, unit A9. *Quat. Int.* **421**, 134–153 (2016).

77. E. Morin, *et al.*, New evidence of broader diets for archaic Homo populations in the northwestern Mediterranean. *Sci Adv* **5**, eaav9106 (2019).

78. C. B. Stringer, *et al.*, Neanderthal exploitation of marine mammals in Gibraltar. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14319–14324 (2008).

79. J. Zilhão, *et al.*, Last Interglacial Iberian Neandertals as fisher-hunter-gatherers. *Science* **367** (2020).

80. S. El Zaatari, F. E. Grine, P. S. Ungar, J.-J. Hublin, Ecogeographic variation in Neandertal dietary habits: Evidence from occlusal molar microwear texture analysis. *Journal of Human*

Evolution **61**, 411–424 (2011).

81. L. Fiorenza, *et al.*, Molar macrowear reveals Neanderthal eco-geographic dietary variation. *PLoS One* **6**, e14769 (2011).

82. K. L. Krueger, *et al.*, Anterior dental microwear textures show habitat-driven variability in Neandertal behavior. *J. Hum. Evol.* **105**, 13–23 (2017).

83. E. Lev, M. E. Kislev, O. Bar-Yosef, Mousterian vegetal food in Kebara Cave, Mt. Carmel. *J. Archaeol. Sci.* **32**, 475–484 (2005).

84. M. Madella, M. K. Jones, P. Goldberg, Y. Goren, E. Hovers, The exploitation of plant resources by Neanderthals in Amud Cave (Israel): the evidence from phytolith studies. *J. Archaeol. Sci.* **29**, 703–719 (2002).

85. A. G. Henry, A. S. Brooks, D. R. Piperno, Plant foods and the dietary ecology of Neanderthals and early modern humans. *J. Hum. Evol.* **69**, 44–54 (2014).

86. R. C. Power, *et al.*, Dental calculus indicates widespread plant use within the stable Neanderthal dietary niche. *J. Hum. Evol.* **119**, 27–41 (2018).

87. B. L. Hardy, Climatic variability and plant food distribution in Pleistocene Europe: Implications for Neanderthal diet and subsistence. *Quat. Sci. Rev.* **29**, 662–679 (2010).

88. D. C. Salazar-García, *et al.*, Neanderthal diets in central and southeastern Mediterranean Iberia. *Quat. Int.* **318**, 3–18 (2013).

89. C. G. Warinner, “From the ground up: Advances in stable isotope-based paleodietary inference” in *Method and Theory in Paleoethnobotany*, (Univ. Press of Colorado, 2014), pp. 275–291.

90. Y. I. Naito, *et al.*, Ecological niche of Neanderthals from Spy Cave revealed by nitrogen isotopes of individual amino acids in collagen. *J. Hum. Evol.* **93**, 82–90 (2016).

91. T. C. O’Connell, M. J. Collins, Comment on “Ecological niche of Neanderthals from Spy Cave revealed by nitrogen isotopes of individual amino acids in collagen” [J. Hum. Evol. 93 (2016) 82-90]. *J. Hum. Evol.* **117**, 53–55 (2018).

92. Y. I. Naito, *et al.*, Reply to “Comment on ”Ecological niche of Neanderthals from Spy Cave revealed by nitrogen isotopes of individual amino acids in collagen.“ [J. Hum. Evol. 93 (2016) 82-90]” [J. Hum. Evol. 117 (2018) 53-55]. *J. Hum. Evol.* **117**, 56–60 (2018).

93. J. H. Dickson, K. Oeggl, D. Stanton, “Forest Moss”: no part of the European Neanderthal diet. *Antiquity* **91** (2017).

94. C. E. Inchley, *et al.*, Selective sweep on human amylase genes postdates the split with Neanderthals. *Sci. Rep.* **6**, 37198 (2016).

95. D. Xu, *et al.*, Archaic Hominin Introgression in Africa Contributes to Functional Salivary MUC7 Genetic Variation. *Mol. Biol. Evol.* **34**, 2704–2715 (2017).

96. M. Menéndez, *et al.*, The Middle to Upper Paleolithic transition in La Güelga cave (Asturias, Northern Spain). *Quat. Int.* **474**, 71–84 (2018).

97. L. Dalén, *et al.*, Partial genetic turnover in neandertals: continuity in the East and population replacement in the West. *Mol. Biol. Evol.* **29**, 1893–1897 (2012).

98. F. Sanchez, Comparative biometrical study of the Mousterian mandible from Cueva del Boquete de Zafarraya (Málaga, Spain). *Hum. Evol.* **14**, 125–138 (1999).

99. M. J. Walker, *et al.*, Late neandertals in southeastern Iberia: Sima de las Palomas del Cabezo Gordo, Murcia, Spain. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 20631–20636 (2008).

100. M. J. Walker, A. V. Lombardi, J. Zapata, E. Trinkaus, Neandertal mandibles from the Sima de las Palomas del Cabezo Gordo, Murcia, southeastern Spain. *Am. J. Phys. Anthropol.* **142**, 261–272 (2010).

101. E. Trinkaus, M. J. (eds) Walker, *The People of Palomas: Neandertals from the Sima de las Palomas del Cabezo Gordo, Southeastern Spain*, E. Trinkaus, M. J. Walker, Eds. (Texas A&M University Press, 2017).

102. D. Mihailović, *Palaeolithic and Mesolithic Research in the Central Balkans*, D. Mihailović, Ed. (Serbian Archaeological Society, 2014).

103. S. Benazzi, *et al.*, Middle Paleolithic and Uluzzian human remains from Fumane Cave, Italy. *J. Hum. Evol.* **70**, 61–68 (2014).

104. H. Rougier, *et al.*, The Troisième caverne of Goyet (Belgium): An exceptional site with both Neandertal and Upper Paleolithic human remains (2016).

105. D. Richter, *et al.*, The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. *Nature* **546**, 293–296 (2017).

106. J.-J. Hublin, *et al.*, New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature* **546**, 289–292 (2017).

107. H. S. Groucutt, *et al.*, Rethinking the dispersal of Homo sapiens out of Africa. *Evol. Anthropol.* **24**, 149–164 (2015).

108. A. H. Osborne, *et al.*, A humid corridor across the Sahara for the migration of early modern humans out of Africa 120,000 years ago. *Proceedings of the National Academy of Sciences* **105**, 16444–16447 (2008).

109. I. Hershkovitz, *et al.*, The earliest modern humans outside Africa. *Science* **359**, 456–459 (2018).

110. K. Akhilesh, *et al.*, Early Middle Palaeolithic culture in India around 385–172 ka reframes Out of Africa models. *Nature* **554**, 97–101 (2018).

111. D. Langgut, A. Almogi-Labin, M. Bar-Matthews, N. Pickarski, M. Weinstein-Evron, Evidence for a humid interval at 56--44 ka in the Levant and its potential link to modern humans dispersal out of Africa. *J. Hum. Evol.* **124**, 75–90 (2018).

112. L. Wadley, L. Backwell, F. d'Errico, C. Sievers, Cooked starchy rhizomes in Africa 170 thousand years ago. *Science* **367**, 87–91 (2020).

113. C. Larbey, S. M. Mentzer, B. Ligouis, S. Wurz, M. K. Jones, Cooked starchy food in

hearths ca. 120 kya and 65 kya (MIS 5e and MIS 4) from Klasies River Cave, South Africa. *J. Hum. Evol.* **131**, 210–227 (2019).

114. R. Grün, *et al.*, U-series and ESR analyses of bones and teeth relating to the human burials from Skhul. *J. Hum. Evol.* **49**, 316–334 (2005).

115. C. Clarkson, *et al.*, Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306–310 (2017).

116. A. J. E. Pryor, M. Steele, M. K. Jones, J. Svoboda, D. G. Beresford-Jones, Plant foods in the Upper Palaeolithic at Dolní Věstonice? Parenchyma redux. *Antiquity* **87**, 971–984 (2013).

117. A. Revedin, *et al.*, Thirty thousand-year-old evidence of plant food processing. *Proceedings of the National Academy of Sciences* **107**, 18815–18819 (2010).

118. R. C. Power, D. C. Salazar-García, L. G. Straus, M. R. González Morales, A. G. Henry, Microremains from El Mirón Cave human dental calculus suggest a mixed plant–animal subsistence economy during the Magdalenian in Northern Iberia. *J. Archaeol. Sci.* **60**, 39–46 (2015).

119. A. Snir, *et al.*, The Origin of Cultivation and Proto-Weeds, Long Before Neolithic Farming. *PLoS One* **10**, e0131422 (2015).

120. L. A. Maher, T. Richter, J. T. Stock, The pre-Natufian Epipaleolithic: long-term behavioral trends in the Levant. *Evol. Anthropol.* **21**, 69–81 (2012).

121. M. Walker, *et al.*, Formal definition and dating of the GSSP (Global Stratotype Section and Point) for the base of the Holocene using the Greenland NGRIP ice core, and selected auxiliary records. *Journal of Quaternary Science* **24**, 3–17 (2009).

122. H. Renssen, *et al.*, Multiple causes of the Younger Dryas cold period. *Nature Geoscience* **8**, 946–949 (2015).

123. X. Yang, *et al.*, Early millet use in northern China. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3726–3730 (2012).

124. D. J. Cohen, The Beginnings of Agriculture in China: A Multiregional View. *Curr. Anthropol.* **52**, S273–S293 (2011).

125. G. Larson, *et al.*, Current perspectives and the future of domestication studies. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 6139–6146 (2014).

126. S. H. Ambrose, Chronology of the Later Stone Age and Food Production in East Africa. *J. Archaeol. Sci.* **25**, 377–392 (1998).

127. E. Loftus, J. Sealy, J. Lee-Thorp, New Radiocarbon Dates and Bayesian Models for Nelson Bay Cave and Byneskranskop 1: Implications for the South African Later Stone Age Sequence. *Radiocarbon* **58**, 365–381 (2016).

128. Q. Fu, *et al.*, An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).

129. M. Hajdinjak, *et al.*, Reconstructing the genetic history of late Neanderthals. *Nature* **555**, 652–656 (2018).

130. Q. Fu, *et al.*, The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).

131. C. Posth, *et al.*, Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr. Biol.* **26**, 827–833 (2016).

132. T. Günther, *et al.*, Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).

133. M. Sorensen, H. Lübke, D. Groß, "The Early Mesolithic in Southern Scandinavia and Northern Germany" in *Star Carr Volume 1. A Persistent Place in a Changing World*, (York: White Rose University Press, 2018), pp. 305–330.

134. C. Meiklejohn, G. Bosset, F. Valentin, Radiocarbon dating of Mesolithic human remains in France. *Mesolithic Miscellany* **21**, 10–56 (2010).

135. J. F. Gibaja, *et al.*, The emergence of mesolithic cemeteries in SW Europe: insights from the El Collado (Oliva, Valencia, Spain) radiocarbon record. *PLoS One* **10**, e0115505 (2015).

136. C. Meiklejohn, M. Roksandic, M. Jackes, D. Lubell, Radiocarbon dating of Mesolithic human remains in Portugal. *Mesolithic Miscellany* **20**, 4–16 (2009).

137. C. Meiklejohn, A. T. Chamberlain, R. J. Schulting, Radiocarbon dating of Mesolithic human remains in Great Britain. *Mesolithic Miscellany* **21**, 20–58 (2011).

138. N. Piotrowska, J. Tomczyk, S. Pawełczyk, Ł. M. Stanaszek, Radiocarbon AMS Dating of Mesolithic Human Remains from Poland. *Radiocarbon* **61**, 991–1007 (2019).

139. M. Gkiasta, T. Russell, S. Shennan, J. Steele, Neolithic transition in Europe: the radiocarbon record revisited. *Antiquity* **77**, 45–62 (2003).

140. D. Gronenborn, Migration, acculturation and culture change in western temperate Eurasia,. *Doc. praeh.* **30**, 79–91 (2003).

141. O. García-Puchol, A. A. Diez Castillo, S. Pardo-Gordó, "Timing the Western Mediterranean Last Hunter-Gatherers and First Farmers" in *Times of Neolithic Transition along the Western Mediterranean*, Fundamental Issues in Archaeology., O. García-Puchol, D. C. Salazar-García, Eds. (Springer International Publishing, 2017), pp. 69–99.

142. W. Haak, *et al.*, Ancient DNA from European early neolithic farmers reveals their near eastern affinities. *PLoS Biol.* **8**, e1000536 (2010).

143. I. Mathieson, *et al.*, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).

144. F. Marshall, E. Hildebrand, Cattle Before Crops: The Beginnings of Food Production in Africa. *Journal of World Prehistory* **16**, 99–143 (2002).

145. K. Wang, *et al.*, Ancient genomes reveal complex patterns of population movement,

interaction, and replacement in sub-Saharan Africa. *Sci Adv* **6**, eaaz0183 (2020).

146. D. Gifford-Gonzalez, Pastoralism in sub-Saharan Africa. *Oxford Handbooks Online* (2017) https:/doi.org/10.1093/oxfordhb/9780199686476.013.27.

147. D. Q. Fuller, E. Hildebrand, “Domesticating Plants in Africa” in *The Oxford Handbook of African Archaeology*, P. Mitchell, P. J. Lane, Eds. (Oxford University Press, 2013), pp. 1–25.

148. A. M. Mercuri, A. Catherine D’Andrea, R. Fornaciari, A. Höhn, *Plants and People in the African Past: Progress in African Archaeobotany* (Springer, 2018).

149. F. E. Dewhirst, *et al.*, The human oral microbiome. *J. Bacteriol.* **192**, 5002–5017 (2010).

150. A. R. Lieverse, Diet and the aetiology of dental calculus. *International Journal of Osteoarchaeology* **9**, 219–232 (1999).

151. N. Grassl, *et al.*, Ultra-deep and quantitative saliva proteome reveals dynamics of the oral microbiome. *Genome Med.* **8**, 44 (2016).

152. B. Manconi, *et al.*, The intriguing heterogeneity of human salivary proline-rich proteins: Short title: Salivary proline-rich protein species. *J. Proteomics* **134**, 47–56 (2016).

153. W. L. Siqueira, W. Zhang, E. J. Helmerhorst, S. P. Gygi, F. G. Oppenheim, Identification of Protein Components in in vivo Human Acquired Enamel Pellicle Using LC−ESI−MS/MS. *J. Proteome Res.* **6**, 2152–2160 (2007).

154. A. Bennick, *et al.*, The role of human salivary acidic proline-rich proteins in the formation of acquired dental pellicle in vivo and their fate after adsorption to the human enamel surface. *Arch. Oral Biol.* **28**, 19–27 (1983).

155. C. I. Fernández, A. S. Wiley, Rethinking the starch digestion hypothesis for AMY1 copy number variation in humans. *Am. J. Phys. Anthropol.* **163**, 645–657 (2017).

156. A. C. Poole, *et al.*, Human Salivary Amylase Gene Copy Number Impacts Oral and Gut Microbiomes. *Cell Host Microbe* **25**, 553–564.e7 (2019).

157. G. Laden, R. Wrangham, The rise of the hominids as an adaptive shift in fallback foods: plant underground storage organs (USOs) and australopith origins. *J. Hum. Evol.* **49**, 482–498 (2005).

158. I. Lazaridis, *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).

159. R. C. Power, D. C. Salazar-García, A. G. Henry, “Dental calculus evidence of Gravettian diet and behaviour at Dolní Věstonice and Pavlov” in *Dolní Věstonice II: Chronostratigraphy, Paleothnology, Paleoanthropology*, The Dolní Věstonice Studies., J. A. Svoboda, Ed. (Academy of Sciences of the Czech Republic, 2016), pp. 345–352.

160. Straus, L.G. Gonzalez Morales, M. Carretero, J.M., “The Red Lady of El Miron Cave”: Lower Magdalenian Human Burial in Cantabrian Spain. *J. Anthropol. Res.* **60**, 1–138 (2015).

161. J. Dabney, *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15758–15763 (2013).

162. K. Aagaard, *et al.*, The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. *FASEB J.* **27**, 1012–1022 (2013).

163. C. Marotz, *et al.*, DNA extraction for streamlined metagenomics of diverse environmental samples. *Biotechniques* **62**, 290–293 (2017).

164. L. R. Thompson, *et al.*, A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).

165. R. W. Hagan, *et al.*, Comparison of extraction methods for recovering ancient microbial DNA from paleofeces. *American Journal of Physical Anthropology* (2019) https:/doi.org/10.1002/ajpa.23978.

166. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, db.prot5448 (2010).

167. M. Kircher, S. Sawyer, M. Meyer, Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2012).

168. I. M. Velsko, C. G. Warinner, Bioarchaeology of the human microbiome. *Bioarchaeology International* **1**, 86–99 (2017).

169. A. T. Ozga, *et al.*, Oral microbiome diversity among cheyenne and arapaho individuals from Oklahoma. *Am. J. Phys. Anthropol.* (2016) https:/doi.org/10.1002/ajpa.23033.

170. C. Jeong, *et al.*, Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe. *Proc. Natl. Acad. Sci. U. S. A.*, 201813608 (2018).

171. A. T. Ozga, *et al.*, Successful enrichment and recovery of whole mitochondrial genomes from ancient human dental calculus. *Am. J. Phys. Anthropol.* (2016) https:/doi.org/10.1002/ajpa.22960.

172. A. W. Briggs, *et al.*, Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Res.* **38**, e87 (2010).

173. R Core Team, *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, 2015).

174. H. Wickham, *et al.*, Welcome to the Tidyverse. *JOSS* **4**, 1686 (2019).

175. H. Wickham, J. Hester, R. Francois, "readr: Read Rectangular Text Data" (2017).

176. H. Wickham, L. Henry, "tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions" (2018).

177. H. Wickham, R. François, L. Henry, K. Müller, "dplyr: A Grammar of Data Manipulation" (2018).

178. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York,

2016).

179. L. Henry, H. Wickham, “purrr: Functional Programming Tools” (2018).

180. S. M. Bache, H. Wickham, “magrittr: A Forward-Pipe Operator for R” (2014).

181. K. Müller, H. Wickham, “tibble: Simple Data Frames” (2018).

182. H. Wickham, “stringr: Simple, Consistent Wrappers for Common String Operations” (2018).

183. M. Dowle, A. Srinivasan, “data.table: Extension of `data.frame`” (2018).

184. E. Neuwirth, “RColorBrewer: ColorBrewer palettes” (2014).

185. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T.-Y. Lam, ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).

186. T. L. Pedersen, Patchwork: The Composer of Ggplots (2017).

187. V. Slon, *et al.*, Neandertal and Denisovan DNA from Pleistocene sediments. *Science* **356**, 605–608 (2017).

188. A. J. Obregon-Tito, *et al.*, Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* **6**, 6505 (2015).

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

190. J. Oh, *et al.*, Temporal Stability of the Human Skin Microbiome. *Cell* **165**, 854–866 (2016).

191. Consortium, The Human Microbiome Project, Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).

192. K. Sankaranarayanan, *et al.*, Gut Microbiome Diversity among Cheyenne and Arapaho Individuals from Western Oklahoma. *Curr. Biol.* **25**, 3161–3169 (2015).

193. R. Leinonen, H. Sugawara, M. Shumway, International Nucleotide Sequence Database Collaboration, The sequence read archive. *Nucleic Acids Res.* **39**, D19–21 (2011).

194. A. E. Mann, *et al.*, Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Sci. Rep.* **8**, 9822 (2018).

195. A. Peltzer, *et al.*, EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 1–14 (2016).

196. M. S. Longo, M. J. O’Neill, R. J. O’Neill, Abundant human DNA contamination identified in non-primate genome databases. *PLoS One* **6**, e16410 (2011).

197. K. Kryukov, T. Imanishi, Human Contamination in Public Genome Assemblies. *PLoS One* **11**, e0162424 (2016).

198. F. P. Breitwieser, M. Pertea, A. Zimin, S. L. Salzberg, Human contamination in bacterial genomes has created thousands of spurious proteins. *Genome Res.* **29**, 954–960 (2019).

199. E. S. Lander, *et al.*, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

200. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

201. M. Kircher, Analysis of high-throughput ancient DNA sequencing data. *Methods Mol. Biol.* **840**, 197–228 (2012).

202. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

203. C. Warinner, *et al.*, A Robust Framework for Microbial Archaeology. *Annu. Rev. Genomics Hum. Genet.* **18**, 321–356 (2017).

204. I. M. Velsko, L. A. F. Frantz, A. Herbig, G. Larson, C. Warinner, Selection of Appropriate Metagenome Taxonomic Classifiers for Ancient Microbiome Research. *mSystems* **3** (2018).

205. R. Hübler, *et al.*, HOPS: automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biology* **20**, 280 (2019).

206. A. Herbig, *et al.*, MALT: Fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. *bioRxiv*, 050559 (2016).

207. Å. J. Vågene, *et al.*, Salmonella enterica genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat Ecol Evol* (2018) https:/doi.org/10.1038/s41559-017-0446-6.

208. N. A. O'Leary, *et al.*, Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–45 (2016).

209. D. H. Huson, *et al.*, MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput. Biol.* **12**, e1004957 (2016).

210. J. Ho, T. Tumkaya, S. Aryal, H. Choi, A. Claridge-Chang, Moving beyond P values: data analysis with estimation graphics. *Nat. Methods* (2019) https:/doi.org/10.1038/s41592-019-0470-3.

211. E. Pasolli, *et al.*, Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* **0**, 649–662.E20 (2019).

212. M. G. Jones, The first filamentous fungal genome sequences: Aspergillus leads the way for essential everyday resources or dusty museum specimens? *Microbiology* **153**, 1–6 (2007).

213. E. T. Saitta, *et al.*, Cretaceous dinosaur bone contains recent organic material and provides an environment conducive to microbial communities. *Elife* **8** (2019).

214. M. E. Benbow, A. J. Lewis, J. K. Tomberlin, J. L. Pechal, Seasonal necrophagous insect

community assembly during vertebrate carrion decomposition. *J. Med. Entomol.* **50**, 440–450 (2013).

215. J. L. Pechal, C. J. Schmidt, H. R. Jordan, M. E. Benbow, A large-scale survey of the postmortem human microbiome, and its potential to provide insight into the living health condition. *Sci. Rep.* **8**, 5724 (2018).

216. J. L. Metcalf, *et al.*, Microbial community assembly and metabolic function during mammalian corpse decomposition. *Science* (2015) https:/doi.org/10.1126/science.aad2646.

217. S. Sawyer, J. Krause, K. Guschanski, V. Savolainen, S. Pääbo, Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131 (2012).

218. C. I. Smith, *et al.*, Neanderthal DNA. Not just old but old and cold? *Nature* **410**, 771–772 (2001).

219. M. E. Allentoft, *et al.*, The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc. Biol. Sci.* **279**, 4724–4733 (2012).

220. T. Chen, *et al.*, The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database* **2010**, baq013 (2010).

221. C. Warinner, *et al.*, Pathogens and host immunity in the ancient human oral cavity. *Nat. Genet.* **46**, 336–344 (2014).

222. M. Delgado-Baquerizo, *et al.*, A global atlas of the dominant bacteria found in soil. *Science* **359**, 320–325 (2018).

223. D. Knights, *et al.*, Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **8**, 761–763 (2011).

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

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

226. C. Quast, *et al.*, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–6 (2013).

227. J. G. Caporaso, *et al.*, QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).

228. T. Větrovský, P. Baldrian, The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One* **8**, e57923 (2013).

229. T. Z. DeSantis, *et al.*, Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).

230. D. McDonald, *et al.*, The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Gigascience* **1**, 7 (2012).

231. P. J. McMurdie, S. Holmes, Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol.* **10**, e1003531 (2014).

232. A. W. Briggs, *et al.*, Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14616–14621 (2007).

233. W. Chang, J. Cheng, J. J. Allaire, Y. Xie, J. McPherson, “shiny: Web Application Framework for R” (2019).

234. J. Neukamm, “DamageProfiler” (Github, 2019) (September 30, 2019).

235. N. M. Davis, D. M. Proctor, S. P. Holmes, D. A. Relman, B. J. Callahan, Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **6**, 226 (2018).

236. A. Kazarina, *et al.*, Analysis of the bacterial communities in ancient human bones and burial soil samples: Tracing the impact of environmental bacteria. *J. Archaeol. Sci.* **109**, 104989 (2019).

237. M. D. J. Lynch, J. D. Neufeld, Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* **13**, 217–229 (2015).

238. G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, J. J. Egozcue, Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **8**, 2224 (2017).

239. P. J. McMurdie, S. Holmes, phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).

240. J. D. Silverman, A. D. Washburne, S. Mukherjee, L. A. David, A phylogenetic transform enhances analysis of compositional microbiota data. *Elife* **6** (2017).

241. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

242. J. A. Martín-Fernández, C. Barceló-Vidal, V. Pawlowsky-Glahn, Zero Replacement in Compositional Data Sets in *Data Analysis, Classification, and Related Methods*, (Springer Berlin Heidelberg, 2000), pp. 155–160.

243. J. Palarea-Albaladejo, J. A. Martín-Fernández, zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics Intellig. Lab. Syst.* **143**, 85–96 (2015).

244. M. J. Anderson, A new method for non-parametric multivariate analysis of variance: NON-PARAMETRIC MANOVA FOR ECOLOGY. *Austral Ecol.* **26**, 32–46 (2001).

245. J. Oksanen, *et al.*, vegan: Community Ecology Package. *https://CRAN.R-project.org/package=vegan* (2018).

246. M. J. Anderson, D. C. I. Walsh, PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? *Ecol. Monogr.* **83**, 557–574 (2013).

247. M. J. Anderson, Distance-based tests for homogeneity of multivariate dispersions.

Biometrics **62**, 245–253 (2006).

248. M. J. Anderson, D. C. I. Walsh, K. Robert Clarke, R. N. Gorley, E. Guerra-Castro, Some solutions to the multivariate Behrens-Fisher problem for dissimilarity-based analyses. *Aust. N. Z. J. Stat.* **59**, 57–79 (2017).

249. J. J. Egozcue, V. Pawlowsky-Glahn, Isometric logratio transformations for compositional data analysis. *Mathematical Geology* **35**, 279–300 (2003).

250. R. Suzuki, H. Shimodaira, Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).

251. G. Tonkin-Hill, J. A. Lees, S. D. Bentley, S. D. W. Frost, J. Corander, Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Res.* **47**, 5539–5549 (2019).

252. L. C. Reimer, *et al.*, BacDive in 2019: bacterial phenotypic data for High-throughput biodiversity analysis. *Nucleic Acids Res.* **47**, D631–D636 (2019).

253. P. E. Kolenbrander, R. J. Palmer Jr, S. Periasamy, N. S. Jakubovics, Oral multispecies biofilm development and the key role of cell-cell distance. *Nat. Rev. Microbiol.* **8**, 471–480 (2010).

254. D. J. Bradshaw, P. D. Marsh, G. K. Watson, C. Allison, Role of Fusobacterium nucleatum and coaggregation in anaerobe survival in planktonic and biofilm oral microbial communities during aeration. *Infect. Immun.* **66**, 4729–4732 (1998).

255. J. L. Mark Welch, B. J. Rossetti, C. W. Rieken, F. E. Dewhirst, G. G. Borisy, Biogeography of a human oral microbiome at the micron scale. *Proceedings of the National Academy of Sciences* (2016) https:/doi.org/10.1073/pnas.1522149113.

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

257. S. Nayfach, Z. J. Shi, R. Seshadri, K. S. Pollard, N. C. Kyrpides, New insights from uncultivated genomes of the global human gut microbiome. *Nature* **568**, 505–510 (2019).

258. D. J. Nasko, S. Koren, A. M. Phillippy, T. J. Treangen, RefSeq database growth influences the accuracy of k-mer-based lowest common ancestor species identification. *Genome Biol.* **19**, 165 (2018).

259. C. J. Beall, *et al.*, Single cell genomics of uncultured, health-associated Tannerella BU063 (Oral Taxon 286) and comparison to the closely related pathogen Tannerella forsythia. *PLoS One* **9**, e89398 (2014).

260. M. Dufrêne, P. Legendre, SPECIES ASSEMBLAGES AND INDICATOR SPECIES:THE NEED FOR A FLEXIBLE ASYMMETRICAL APPROACH. *Ecol. Monogr.* **67**, 345–366 (1997).

261. M. De Cáceres, P. Legendre, M. Moretti, Improving indicator species analysis by combining groups of sites. *Oikos* **119**, 1674–1684 (2010).

262. G. C. Armitage, Comparison of the microbiological features of chronic and aggressive periodontitis. *Periodontol. 2000* **53**, 70–88 (2010).

263. P. I. Diaz, Microbial diversity and interactions in subgingival biofilm communities. *Front. Oral Biol.* **15**, 17–40 (2012).

264. P. W. Lepp, *et al.*, Methanogenic Archaea and human periodontal disease. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 6176–6181 (2004).

265. C. J. Adler, *et al.*, Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat. Genet.*, 1–7 (2013).

266. D. Belstrøm, *et al.*, Bacterial profiles of saliva in relation to diet, lifestyle factors, and socioeconomic status. *J. Oral Microbiol.* **6** (2014).

267. F. De Filippis, *et al.*, The same microbiota and a potentially discriminant metabolome in the saliva of omnivore, ovo-lacto-vegetarian and Vegan individuals. *PLoS One* **9**, e112373 (2014).

268. M. K. Keller, C. A. Kressirer, D. Belstrøm, S. Twetman, A. C. R. Tanner, Oral microbial profiles of individuals with different levels of sugar intake. *J. Oral Microbiol.* **9**, 1355207 (2017).

269. C. Signoretto, G. Burlacchini, F. Bianchi, G. Cavalleri, P. Canepari, Differences in microbiological composition of saliva and dental plaque in subjects with different drinking habits. *New Microbiol.* **29**, 293–302 (2006).

270. C. Signoretto, *et al.*, Drinking habits are associated with changes in the dental plaque microbial community. *J. Clin. Microbiol.* **48**, 347–356 (2010).

271. E. Söderling, *et al.*, Effects of short-term xylitol gum chewing on the oral microbiome. *Clin. Oral Investig.* **19**, 237–244 (2015).

272. E. M. Prosdocimi, *et al.*, Effect of maltitol-containing chewing gum use on the composition of dental plaque microbiota in subjects with active dental caries. *J. Oral Microbiol.* **9**, 1374152 (2017).

273. M. A. Listgarten, The structure of dental plaque. *Periodontol. 2000* **5**, 52–65 (1994).

274. S. S. Socransky, A. D. Haffajee, Periodontal microbial ecology. *Periodontol. 2000* **38**, 135–187 (2005).

275. S. C. Holt, J. L. Ebersole, Porphyromonas gingivalis, Treponema denticola, and Tannerella forsythia: the “red complex”, a prototype polybacterial pathogenic consortium in periodontitis. *Periodontol. 2000* **38**, 72–122 (2005).

276. E. Zaura, B. J. F. Keijser, S. M. Huse, W. Crielaard, Defining the healthy “core microbiome” of oral microbial communities. *BMC Microbiol.* **9**, 259 (2009).

277. V. Zijnge, *et al.*, Oral biofilm architecture on natural teeth. *PLoS One* **5**, e9321 (2010).

278. A. Stacy, *et al.*, Bacterial fight-and-flight responses enhance virulence in a polymicrobial infection. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 7819–7824 (2014).

279. J. Slots, Periodontal herpesviruses: prevalence, pathogenicity, systemic risk.

Periodontol. 2000 **69**, 28–45 (2015).

280. M. Bonner, M. Fresno, N. Gironès, N. Guillén, J. Santi-Rocca, Reassessing the Role of Entamoeba gingivalis in Periodontitis. *Front. Cell. Infect. Microbiol.* **8**, 379 (2018).

281. S. S. Socransky, A. D. Manganiello, D. Propas, V. Oram, J. van Houte, Bacteriological studies of developing supragingival dental plaque. *J. Periodontal Res.* **12**, 90–106 (1977).

282. P. D. Marsh, D. J. Bradshaw, Dental plaque as a biofilm. *J. Ind. Microbiol.* **15**, 169–175 (1995).

283. W. L. Siqueira, W. Custodio, E. E. McDonald, New Insights into the Composition and Functions of the Acquired Enamel Pellicle. *J. Dent. Res.* **91**, 1110–1118 (2012).

284. D. M. Carlson, Salivary proline-rich proteins: biochemistry, molecular biology, and regulation of expression. *Crit. Rev. Oral Biol. Med.* **4**, 495–502 (1993).

285. A. E. Hagerman, L. G. Butler, The specificity of proanthocyanidin-protein interactions. *J. Biol. Chem.* **256**, 4494–4497 (1981).

286. R. J. Gibbons, D. I. Hay, Human salivary acidic proline-rich proteins and statherin promote the attachment of Actinomyces viscosus LY7 to apatitic surfaces. *Infect. Immun.* **56**, 439–445 (1988).

287. M. Mau, K.-H. Südekum, A. Johann, A. Sliwa, T. M. Kaiser, Saliva of the graminivorous Theropithecus gelada lacks proline-rich proteins and tannin-binding capacity. *Am. J. Primatol.* **71**, 663–669 (2009).

288. C. Hannig, M. Hannig, T. Attin, Enzymes in the acquired enamel pellicle. *Eur. J. Oral Sci.* **113**, 2–13 (2005).

289. H. L. Gibbins, G. B. Proctor, G. E. Yakubov, S. Wilson, G. H. Carpenter, Concentration of salivary protective proteins within the bound oral mucosal pellicle. *Oral Dis.* **20**, 707–713 (2014).

290. D. Deimling, *et al.*, Electron microscopic detection of salivary α-amylase in the pellicle formed in situ. *Eur. J. Oral Sci.* **112**, 503–509 (2004).

291. P. E. Kolenbrander, *et al.*, Communication among Oral Bacteria. *Microbiol. Mol. Biol. Rev.* **66**, 486–505 (2002).

292. U. Lendenmann, J. Grogan, F. G. Oppenheim, Saliva and Dental Pellicle-A Review. *Adv. Dent. Res.* **14**, 22–28 (2000).

293. H. Löe, E. Theilade, S. B. Jensen, Experimental Gingivitis in Man. *J. Periodontol.* **36**, 177–187 (1965).

294. E. Theilade, W. H. Wright, S. B. Jensen, H. Löe, Experimental gingivitis in man. II. A longitudinal clinical and bacteriological investigation. *J. Periodontal Res.* **1**, 1–13 (1966).

295. P. E. Kolenbrander, Oral microbial communities: biofilms, interactions, and genetic systems. *Annu. Rev. Microbiol.* **54**, 413–437 (2000).

296. N. Tinanoff, A. Gross, J. M. Brady, Development of plaque on enamel. Parallel investigations. *J. Periodontal Res.* **11**, 197–209 (1976).

297. H. L. Ritz, Microbial population shifts in developing human dental plaque. *Arch. Oral Biol.* **12**, 1561–1568 (1967).

298. B. Nyvad, M. Kilian, Microbiology of the early colonization of human enamel and root surfaces in vivo. *Scand. J. Dent. Res.* **95**, 369–380 (1987).

299. P. Ramberg, S. Sekino, N. G. Uzel, S. Socransky, J. Lindhe, Bacterial colonization during de novo plaque formation. *J. Clin. Periodontol.* **30**, 990–995 (2003).

300. P. I. Diaz, *et al.*, Molecular characterization of subject-specific oral microflora during initial colonization of enamel. *Appl. Environ. Microbiol.* **72**, 2837–2848 (2006).

301. I. H. Pratt-Terpstra, A. H. Weerkamp, H. J. Busscher, The effects of pellicle formation on streptococcal adhesion to human enamel and artificial substrata with various surface free-energies. *J. Dent. Res.* **68**, 463–467 (1989).

302. P. E. Kolenbrander, *et al.*, Bacterial interactions and successions during plaque development. *Periodontol. 2000* **42**, 47–79 (2006).

303. P. E. Kolenbrander, R. N. Andersen, L. V. Moore, Coaggregation of Fusobacterium nucleatum, Selenomonas flueggei, Selenomonas infelix, Selenomonas noxia, and Selenomonas sputigena with strains from 11 genera of oral bacteria. *Infect. Immun.* **57**, 3194–3203 (1989).

304. P. E. Kolenbrander, R. N. Andersen, Inhibition of coaggregation between Fusobacterium nucleatum and Porphyromonas (Bacteroides) gingivalis by lactose and related sugars. *Infect. Immun.* **57**, 3204–3209 (1989).

305. D. F. Mangan, M. J. Novak, S. A. Vora, J. Mourad, P. S. Kriger, Lectinlike interactions of Fusobacterium nucleatum with human neutrophils. *Infect. Immun.* **57**, 3601–3611 (1989).

306. D. Polak, L. Shapira, E. I. Weiss, Y. Houri-Haddad, The role of coaggregation between Porphyromonas gingivalis and Fusobacterium nucleatum on the host response to mixed infection. *J. Clin. Periodontol.* **39**, 617–625 (2012).

307. J. Highfield, Diagnosis and classification of periodontal disease. *Aust. Dent. J.* **54 Suppl 1**, S11–26 (2009).

308. C. W. Lamell, A. L. Griffen, D. L. McClellan, E. J. Leys, Acquisition and colonization stability of Actinobacillus actinomycetemcomitans and Porphyromonas gingivalis in children. *J. Clin. Microbiol.* **38**, 1196–1199 (2000).

309. L. Abusleme, *et al.*, The subgingival microbiome in health and periodontitis and its relationship with community biomass and inflammation. *ISME J.* **7**, 1016–1025 (2013).

310. P. Lancy Jr, J. M. Dirienzo, B. Appelbaum, B. Rosan, S. C. Holt, Corncob formation between Fusobacterium nucleatum and Streptococcus sanguis. *Infect. Immun.* **40**, 303–309 (1983).

311. A. Shrivastava, *et al.*, Cargo transport shapes the spatial organization of a microbial

community. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 8633–8638 (2018).

312. Y. Jin, H.-K. Yip, Supragingival calculus: formation and control. *Crit. Rev. Oral Biol. Med.* **13**, 426–441 (2002).

313. B. T. Rosier, M. De Jager, E. Zaura, B. P. Krom, Historical and contemporary hypotheses on the development of oral diseases: are we there yet? *Front. Cell. Infect. Microbiol.* **4**, 92 (2014).

314. P. D. Marsh, Microbial ecology of dental plaque and its significance in health and disease. *Adv. Dent. Res.* **8**, 263–271 (1994).

315. G. Hajishengallis, R. J. Lamont, Beyond the red complex and into more complexity: the polymicrobial synergy and dysbiosis (PSD) model of periodontal disease etiology. *Mol. Oral Microbiol.* **27**, 409–419 (2012).

316. A. L. Griffen, *et al.*, Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. *ISME J.* **6**, 1176–1185 (2012).

317. L. M. Shaddox, *et al.*, Microbiological characterization in children with aggressive periodontitis. *J. Dent. Res.* **91**, 927–933 (2012).

318. H. Chen, *et al.*, A Filifactor alocis-centered co-occurrence group associates with periodontitis across different oral habitats. *Sci. Rep.* **5**, 9053 (2015).

319. R. R. D. S. Oliveira, *et al.*, Levels of Candidate Periodontal Pathogens in Subgingival Biofilm. *J. Dent. Res.* **95**, 711–718 (2016).

320. D. Belstrøm, *et al.*, Metagenomic and metatranscriptomic analysis of saliva reveals disease-associated microbiota in patients with periodontitis and dental caries. *NPJ Biofilms Microbiomes* **3**, 23 (2017).

321. M. Naginyte, T. Do, J. Meade, D. A. Devine, P. D. Marsh, Enrichment of periodontal pathogens from the biofilms of healthy adults. *Sci. Rep.* **9**, 5491 (2019).

322. P. F. ter Steeg, J. S. van der Hoeven, Growth stimulation of Treponema denticola by periodontal microorganisms. *Antonie Van Leeuwenhoek* **57**, 63–70 (1990).

323. D. Grenier, Nutritional interactions between two suspected periodontopathogens, Treponema denticola and Porphyromonas gingivalis. *Infect. Immun.* **60**, 5298–5301 (1992).

324. M. Yoneda, *et al.*, Stimulation of growth of Porphyromonas gingivalis by cell extracts from Tannerella forsythia. *J. Periodontal Res.* **40**, 105–109 (2005).

325. Y. Zhu, *et al.*, Porphyromonas gingivalis and Treponema denticola synergistic polymicrobial biofilm development. *PLoS One* **8**, e71727 (2013).

326. K. H. Tan, *et al.*, Porphyromonas gingivalis and Treponema denticola exhibit metabolic symbioses. *PLoS Pathog.* **10**, e1003955 (2014).

327. J. Li, *et al.*, Identification of early microbial colonizers in human dental biofilm. *J. Appl. Microbiol.* **97**, 1311–1318 (2004).

328. S. Periasamy, P. E. Kolenbrander, Mutualistic biofilm communities develop with Porphyromonas gingivalis and initial, early, and late colonizers of enamel. *J. Bacteriol.* **191**, 6804–6811 (2009).

329. S. H. Lee, D. H. Baek, Characteristics of Porphyromonas gingivalis lipopolysaccharide in co-culture with Fusobacterium nucleatum. *Mol. Oral Microbiol.* **28**, 230–238 (2013).

330. M. Rangarajan, *et al.*, Identification of a second lipopolysaccharide in Porphyromonas gingivalis W50. *J. Bacteriol.* **190**, 2920–2932 (2008).

331. G. R. Riviere, *et al.*, Periodontal status and detection frequency of bacteria at sites of periodontal health and gingivitis. *J. Periodontol.* **67**, 109–115 (1996).

332. S. J. Byrne, *et al.*, Progression of chronic periodontitis can be predicted by the levels of Porphyromonas gingivalis and Treponema denticola in subgingival plaque. *Oral Microbiol. Immunol.* **24**, 469–477 (2009).

333. L. Chu, *et al.*, A 52-kDa leucyl aminopeptidase from treponema denticola is a cysteinylglycinase that mediates the second step of glutathione metabolism. *J. Biol. Chem.* **283**, 19351–19358 (2008).

334. S. G. Dashper, C. A. Seers, K. H. Tan, E. C. Reynolds, Virulence factors of the oral spirochete Treponema denticola. *J. Dent. Res.* **90**, 691–703 (2011).

335. J. C. Fenno, *et al.*, Cytopathic effects of the major surface protein and the chymotrypsinlike protease of Treponema denticola. *Infect. Immun.* **66**, 1869–1877 (1998).

336. J. C. Fenno, Treponema denticola interactions with host proteins. *J. Oral Microbiol.* **4**, 1–13 (2012).

337. A. Ikegami, K. Honma, A. Sharma, H. K. Kuramitsu, Multiple functions of the leucine-rich repeat protein LrrA of Treponema denticola. *Infect. Immun.* **72**, 4619–4627 (2004).

338. A. Sharma, Virulence mechanisms of Tannerella forsythia. *Periodontol. 2000* **54**, 106–116 (2010).

339. M. Bielecki, *et al.*, Tannerella forsythia Tfo belongs to Porphyromonas gingivalis HmuY-like family of proteins but differs in heme-binding properties. *Biosci. Rep.* **38** (2018).

340. S. Yost, A. E. Duran-Pinedo, The contribution of Tannerella forsythia dipeptidyl aminopeptidase IV in the breakdown of collagen. *Mol. Oral Microbiol.* **33**, 407–419 (2018).

341. A. Sharma, *et al.*, Cloning, expression, and sequencing of a cell surface antigen containing a leucine-rich repeat motif from Bacteroides forsythus ATCC 43037. *Infect. Immun.* **66**, 5703–5710 (1998).

342. T. Saito, K. Ishihara, T. Kato, K. Okuda, Cloning, expression, and sequencing of a protease gene from Bacteroides forsythus ATCC 43037 in Escherichia coli. *Infect. Immun.* **65**, 4888–4891 (1997).

343. K. Honma, E. Mishima, A. Sharma, Role of Tannerella forsythia NanH sialidase in epithelial cell attachment. *Infect. Immun.* **79**, 393–401 (2011).

344. A. M. Frey, K. Ansbro, N. S. Kamble, T. K. Pham, G. P. Stafford, Characterisation and pure culture of putative health-associated oral bacterium BU063 (Tannerella sp. HOT-286) reveals presence of a potentially novel glycosylated S-layer. *FEMS Microbiol. Lett.* **365** (2018).

345. S. A. Thompson, Campylobacter surface-layers (S-layers) and immune evasion. *Ann. Periodontol.* **7**, 43–53 (2002).

346. G. Sekot, *et al.*, Analysis of the cell surface layer ultrastructure of the oral pathogen Tannerella forsythia. *Arch. Microbiol.* **194**, 525–539 (2012).

347. N. Shimotahira, *et al.*, The Surface Layer of Tannerella forsythia Contributes to Serum Resistance and Oral Bacterial Coaggregation. *Infection and Immunity* **81**, 1198–1206 (2013).

348. M. B. Tomek, *et al.*, A General Protein O-Glycosylation Gene Cluster Encodes the Species-Specific Glycan of the Oral Pathogen Tannerella forsythia: O-Glycan Biosynthesis and Immunological Implications. *Front. Microbiol.* **9**, 2008 (2018).

349. S. Bloch, T. Thurnheer, Y. Murakami, G. N. Belibasakis, C. Schäffer, Behavior of two Tannerella forsythia strains and their cell surface mutants in multispecies oral biofilms. *Mol. Oral Microbiol.* **32**, 404–418 (2017).

350. S. R. Vartoukian, J. Downes, R. M. Palmer, W. G. Wade, Fretibacterium fastidiosum gen. nov., sp. nov., isolated from the human oral cavity. *Int. J. Syst. Evol. Microbiol.* **63**, 458–463 (2013).

351. S. R. Vartoukian, *et al.*, In Vitro Cultivation of “Unculturable” Oral Bacteria, Facilitated by Community Culture and Media Supplementation with Siderophores. *PLoS One* **11**, e0146926 (2016).

352. M. You, S. Mo, R. M. Watt, W. K. Leung, Prevalence and diversity of Synergistetes taxa in periodontal health and disease. *J. Periodontal Res.* **48**, 159–168 (2013).

353. O.-J. Park, *et al.*, Pyrosequencing Analysis of Subgingival Microbiota in Distinct Periodontal Conditions. *J. Dent. Res.* **94**, 921–927 (2015).

354. P. J. Pérez-Chaparro, *et al.*, Do different probing depths exhibit striking differences in microbial profiles? *J. Clin. Periodontol.* **45**, 26–37 (2018).

355. M. Hamady, R. Knight, Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res.* **19**, 1141–1152 (2009).

356. P. J. Turnbaugh, *et al.*, The human microbiome project. *Nature* **449**, 804–810 (2007).

357. K. Li, M. Bihan, B. A. Methé, Analyses of the stability and core taxonomic memberships of the human microbiome. *PLoS One* **8**, e63139 (2013).

358. C. Petersen, J. L. Round, Defining dysbiosis and its influence on host immunity and disease. *Cell. Microbiol.* **16**, 1024–1033 (2014).

359. Integrative HMP (iHMP) Research Network Consortium, The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of

human health and disease. *Cell Host Microbe* **16**, 276–289 (2014).

360. F. Bäckhed, *et al.*, Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. *Cell Host Microbe* **12**, 611–622 (2012).

361. D. R. Utter, J. L. Mark Welch, G. G. Borisy, Individuality, Stability, and Variability of the Plaque Microbiome. *Front. Microbiol.* **7**, 564 (2016).

362. A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, H. Pfister, UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992 (2014).

363. J. R. Conway, A. Lex, N. Gehlenborg, UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).

364. N. J. Palleroni, "Pseudomonas" in *Bergey's Manual of Systematics of Archaea and Bacteria*, W. B. Whitman, *et al.*, Eds. (John Wiley & Sons, Ltd, 2015), pp. 323–278.

365. P. Kämpfer, "Streptomyces" in *Bergey's Manual of Systematics of Archaea and Bacteria*, W. B. Whitman, *et al.*, Eds. (John Wiley & Sons, Ltd, 2015), pp. 1–414.

366. A. L. Byrd, Y. Belkaid, J. A. Segre, The human skin microbiome. *Nat. Rev. Microbiol.* **16**, 143–155 (2018).

367. C. M. Walsh, M. J. Gebert, M. Delgado-Baquerizo, F. T. Maestre, N. Fierer, A Global Survey of Mycobacterial Diversity in Soil. *Appl. Environ. Microbiol.* **85** (2019).

368. L. Macovei, *et al.*, The hidden "mycobacteriome" of the human healthy oral cavity and upper respiratory tract. *J. Oral Microbiol.* **7**, 26094 (2015).

369. A. K. Wilbur, *et al.*, Deficiencies and challenges in the study of ancient tuberculosis DNA. *J. Archaeol. Sci.* **36**, 1990–1997 (2009).

370. R. M. Thomson, *et al.*, Factors associated with the isolation of Nontuberculous mycobacteria (NTM) from a large municipal water system in Brisbane, Australia. *BMC Microbiol.* **13**, 89 (2013).

371. L. V. H. Moore, J. L. Johnson, W. E. C. Moore, Selenomonas noxia sp. nov., Selenomonas flueggei sp. nov., Selenomonas infelix sp. nov., Selenomonas dianae sp. nov., and Selenomonas artemidis sp. nov., from the Human Gingival Crevice. *Int. J. Syst. Bacteriol.* **37**, 271–280 (1987).

372. Y. Noiri, K. Ozaki, H. Nakae, T. Matsuo, S. Ebisu, An immunohistochemical study on the localization of Porphyromonas gingivalis, Campylobacter rectus and Actinomyces viscosus in human periodontal pockets. *J. Periodontal Res.* **32**, 598–607 (1997).

373. Y. Noiri, L. Li, S. Ebisu, The localization of periodontal-disease-associated bacteria in human periodontal pockets. *J. Dent. Res.* **80**, 1930–1934 (2001).

374. Y. Noiri, L. Li, F. Yoshimura, S. Ebisu, Localization of Porphyromonas gingivalis-carrying fimbriae in situ in human periodontal pockets. *J. Dent. Res.* **83**, 941–945 (2004).

375. J. Drescher, *et al.*, Molecular epidemiology and spatial distribution of Selenomonas spp. in subgingival biofilms. *Eur. J. Oral Sci.* **118**, 466–474 (2010).

376. M. D. Ferrer, A. Mira, Oral Biofilm Architecture at the Microbial Scale. *Trends Microbiol.* **24**, 246–248 (2016).

377. I. Olsen, Organization of supragingival plaque at the micron scale. *J. Oral Microbiol.* **10**, 1438722 (2018).

378. D. Kita, *et al.*, Involvement of the Type IX Secretion System in Capnocytophaga ochracea Gliding Motility and Biofilm Formation. *Appl. Environ. Microbiol.* **82**, 1756–1766 (2016).

379. C. A. Brennan, W. S. Garrett, Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. *Nature Reviews Microbiology* **17**, 156–166 (2019).

380. B. J. Tindall, Arachnia propionica (Buchanan and Pine 1962) Pine and Georg 1969 (Approved Lists 1980), Propionibacterium propionicum corrig. (Buchanan and Pine 1962) Charfreitag et al. 1988 and Pseudopropionibacterium propionicum (Buchanan and Pine 1962) Scholz and Kilian 2016 and the nomenclatural consequences of changes in the taxonomy of the genus Propionibacterium. *Int. J. Syst. Evol. Microbiol.* **69**, 2612–2615 (2019).

381. R. J. Lamont, H. Koo, G. Hajishengallis, The oral microbiota: dynamic communities and host interactions. *Nat. Rev. Microbiol.* **16**, 745–759 (2018).

382. P. G. Egland, R. J. Palmer Jr, P. E. Kolenbrander, Interspecies communication in Streptococcus gordonii-Veillonella atypica biofilms: signaling in flow conditions requires juxtaposition. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16917–16922 (2004).

383. L. Dethlefsen, M. McFall-Ngai, D. A. Relman, An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449**, 811–818 (2007).

384. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

385. R. D. Pathirana, N. M. O'Brien-Simpson, G. C. Brammar, N. Slakeski, E. C. Reynolds, Kgp and RgpB, but not RgpA, are important for Porphyromonas gingivalis virulence in the murine periodontitis model. *Infect. Immun.* **75**, 1436–1442 (2007).

386. A. T. Ozga, *et al.*, Oral microbiome diversity in chimpanzees from Gombe National Park. *Sci. Rep.* **9**, 17354 (2019).

387. R. A. Whiley, J. M. Hardie, "Streptococcus" in *Bergey's Manual of Systematics of Archaea and Bacteria*, Systematic Bacteriology., W. B. Whitman, *et al.*, Eds. (John Wiley & Sons, Ltd, 2015), pp. 1–86.

388. A. Jensen, C. F. P. Scholz, M. Kilian, Re-evaluation of the taxonomy of the Mitis group of the genus Streptococcus based on whole genome phylogenetic analyses, and proposed reclassification of Streptococcus dentisani as Streptococcus oralis subsp. dentisani comb. nov., Streptococcus tigurinus as Streptococcus oralis subsp. tigurinus comb. nov., and Streptococcus oligofermentans as a later synonym of Streptococcus cristatus. *Int. J. Syst. Evol. Microbiol.* **66**, 4803–4820 (2016).

389. J. Abranches, *et al.*, Biology of Oral Streptococci. *Microbiol Spectr* **6** (2018).

390. J. L. Espinoza, *et al.*, Supragingival Plaque Microbiome Ecology and Functional Potential in the Context of Health and Disease. *MBio* **9**, e01631–18 (2018).

391. V. P. Richards, *et al.*, Phylogenomics and the dynamic genome evolution of the genus Streptococcus. *Genome Biol. Evol.* **6**, 741–753 (2014).

392. E. M. Haase, *et al.*, Comparative genomics and evolution of the amylase-binding proteins of oral streptococci. *BMC Microbiol.* **17**, 94 (2017).

393. J. Täpp, M. Thollesson, B. Herrmann, Phylogenetic relationships and genotyping of the genus Streptococcus by sequence determination of the RNase P RNA gene, rnpB. *Int. J. Syst. Evol. Microbiol.* **53**, 1861–1871 (2003).

394. X.-Y. Gao, X.-Y. Zhi, H.-W. Li, H.-P. Klenk, W.-J. Li, Comparative genomics of the bacterial genus Streptococcus illuminates evolutionary implications of species groups. *PLoS One* **9**, e101229 (2014).

395. J. D. Rogers, R. J. Palmer Jr, P. E. Kolenbrander, F. A. Scannapieco, Role of Streptococcus gordonii amylase-binding protein A in adhesion to hydroxyapatite, starch metabolism, and biofilm formation. *Infect. Immun.* **69**, 7046–7056 (2001).

396. Hendy Jessica, *et al.*, Proteomic evidence of dietary sources in ancient dental calculus. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20180977 (2018).

397. J. C. Pronk, R. R. Frants, W. Jansen, A. W. Eriksson, G. J. Tonino, Evidence of duplication of the human salivary amylase gene. *Hum. Genet.* **60**, 32–35 (1982).

398. M. Emi, *et al.*, Overlapping two genes in human DNA: a salivary amylase gene overlaps with a gamma-actin pseudogene that carries an integrated human endogenous retroviral DNA. *Gene* **62**, 229–235 (1988).

399. L. C. Samuelson, K. Wiebauer, D. L. Gumucio, M. H. Meisler, Expression of the human amylase genes: recent origin of a salivary amylase promoter from an actin pseudogene. *Nucleic Acids Res.* **16**, 8261–8276 (1988).

400. L. C. Samuelson, K. Wiebauer, C. M. Snow, M. H. Meisler, Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution. *Mol. Cell. Biol.* **10**, 2513–2520 (1990).

401. L. C. Samuelson, R. S. Phillips, L. J. Swanberg, Amylase gene structures in primates: retroposon insertions and promoter evolution. *Mol. Biol. Evol.* **13**, 767–779 (1996).

402. G. H. Perry, L. Kistler, M. A. Kelaita, A. J. Sams, Insights into hominin phenotypic and dietary evolution from ancient DNA sequence data. *J. Hum. Evol.* **79**, 55–63 (2015).

403. R. L. McGeachin, J. R. Akin, Amylase levels in the tissues and body fluids of several primate species. *Comp. Biochem. Physiol. A Comp. Physiol.* **72**, 267–269 (1982).

404. V. Behringer, T. Deschner, E. Möstl, D. Selzer, G. Hohmann, Stress affects salivary alpha-Amylase activity in bonobos. *Physiol. Behav.* **105**, 476–482 (2012).

405. D. A. Froehlich, R. M. Pangborn, J. R. Whitaker, The effect of oral stimulation on human parotid salivary flow rate and alpha-amylase secretion. *Physiology & Behavior* **41**, 209–217

(1987).

406. C. Boehlke, O. Zierau, C. Hannig, Salivary amylase – The enzyme of unspecialized euryphagous animals. *Archives of Oral Biology* **60**, 1162–1176 (2015).

407. E. Axelsson, *et al.*, The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* **495**, 360–364 (2013).

408. M. Ollivier, *et al.*, Amy2B copy number variation reveals starch diet adaptations in ancient European dogs. *R Soc Open Sci* **3**, 160449 (2016).

409. C. C. Tseng, F. A. Scannapieco, M. J. Levine, Use of a replica-plate assay for the rapid assessment of salivary protein-bacteria interactions. *Oral Microbiol. Immunol.* **7**, 53–56 (1992).

410. F. A. Scannapieco, L. Solomon, R. O. Wadenya, Emergence in human dental plaque and host distribution of amylase-binding streptococci. *J. Dent. Res.* **73**, 1627–1635 (1994).

411. M. Kilian, B. Nyvad, Ability to bind salivary alpha-amylase discriminates certain viridans group streptococcal species. *J. Clin. Microbiol.* **28**, 2576–2577 (1990).

412. J. P. Gwynn, C. W. Douglas, Comparison of amylase-binding proteins in oral streptococci. *FEMS Microbiol. Lett.* **124**, 373–379 (1994).

413. J. D. Rogers, *et al.*, Identification and analysis of a gene (abpA) encoding a major amylase-binding protein in Streptococcus gordonii. *Microbiology* **144 (Pt 5)**, 1223–1233 (1998).

414. X. Liang, *et al.*, A distinct sortase SrtB anchors and processes a streptococcal adhesin AbpA with a novel structural property. *Sci. Rep.* **6**, 30966 (2016).

415. F. A. Scannapieco, G. G. Haraszthy, M. I. Cho, M. J. Levine, Characterization of an amylase-binding component of Streptococcus gordonii G9B. *Infect. Immun.* **60**, 4726–4733 (1992).

416. C. W. Douglas, Characterization of the alpha-amylase receptor of Streptococcus gordonii NCTC 7868. *J. Dent. Res.* **69**, 1746–1752 (1990).

417. E. M. Haase, X. Feng, J. Pan, J. C. Miecznikowski, F. A. Scannapieco, Dynamics of the Streptococcus gordonii Transcriptome in Response to Medium, Salivary α-Amylase, and Starch. *Appl. Environ. Microbiol.* **81**, 5363–5374 (2015).

418. A. E. Nikitkova, E. M. Haase, F. A. Scannapieco, Effect of starch and amylase on the expression of amylase-binding protein A in Streptococcus gordonii. *Mol. Oral Microbiol.* **27**, 284–294 (2012).

419. A. E. Nikitkova, E. M. Haase, F. A. Scannapieco, Taking the starch out of oral biofilm formation: molecular basis and functional significance of salivary α-amylase binding to oral streptococci. *Appl. Environ. Microbiol.* **79**, 416–423 (2013).

420. L. Li, J. M. Tanzer, F. A. Scannapieco, Identification and analysis of the amylase-binding protein B (AbpB) and gene (abpB) from Streptococcus gordonii. *FEMS Microbiol. Lett.* **212**, 151–157 (2002).

421. B. Chaudhuri, *et al.*, Amylase-binding protein B of Streptococcus gordonii is an extracellular dipeptidyl-peptidase. *Infect. Immun.* **76**, 4530–4537 (2008).

422. B. Chaudhuri, J. Rojek, M. M. Vickerman, J. M. Tanzer, F. A. Scannapieco, Interaction of salivary alpha-amylase and amylase-binding-protein A (AbpA) of Streptococcus gordonii with glucosyltransferase of S. gordonii and Streptococcus mutans. *BMC Microbiol.* **7**, 60 (2007).

423. J. Vorrasi, B. Chaudhuri, E. M. Haase, F. A. Scannapieco, Identification and characterization of amylase-binding protein C from Streptococcus mitis NS51. *Mol. Oral Microbiol.* **25**, 150–156 (2010).

424. D. Denapaite, *et al.*, Highly Variable Streptococcus oralis Strains Are Common among Viridans Streptococci Isolated from Primates. *mSphere* **1** (2016).

425. K. Hardy, J. Brand-Miller, K. D. Brown, M. G. Thomas, L. Copeland, The Importance of Dietary Carbohydrate in Human Evolution. *Q. Rev. Biol.* **90**, 251–268 (2015).

426. W. Ding, F. Baumdicker, R. A. Neher, panX: pan-genome analysis and exploration. *Nucleic Acids Res.* **46**, e5 (2018).

427. D. Charif, J. R. Lobry, "SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis" in *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations*, U. Bastolla, M. Porto, H. E. Roman, M. Vendruscolo, Eds. (Springer Berlin Heidelberg, 2007), pp. 207–232.

428. D. J. Winter, rentrez: an R package for the NCBI eUtils API. *The R Journal* **9**, 520–526 (2017).

429. S. Mangiafico, "rcompanion: Functions to Support Extension Education Program Evaluation" (2020).

430. L. C. Aiello, P. Wheeler, The Expensive-Tissue Hypothesis: The Brain and the Digestive System in Human and Primate Evolution. *Curr. Anthropol.* **36**, 199–221 (1995).

431. K. Fonseca-Azevedo, S. Herculano-Houzel, Metabolic constraint imposes tradeoff between body size and number of brain neurons in human evolution. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 18571–18576 (2012).

432. A. Navarrete, C. P. van Schaik, K. Isler, Energetics and the evolution of human brain size. *Nature* **480**, 91–93 (2011).

433. J. A. J. Gowlett, The discovery of fire by humans: a long and convoluted process. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371** (2016).

434. R. Potts, Evolution: Big brains explained. *Nature* **480**, 43–44 (2011).

435. K. Milton, A hypothesis to explain the role of meat-eating in human evolution. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews* **8**, 11–21 (1999).

436. E. Assaf, *et al.*, Shaped stone balls were used for bone marrow extraction at Lower

Paleolithic Qesem Cave, Israel. *PLOS ONE* **15**, e0230972 (2020).

437. R. Wrangham, N. Conklin-Brittain, Cooking as a biological trait. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **136**, 35–46 (2003).

438. S. P. McPherron, *et al.*, Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature* **466**, 857–860 (2010).

439. M. Domínguez-Rodrigo, T. R. Pickering, S. Semaw, M. J. Rogers, Cutmarked bones from Pliocene archaeological sites at Gona, Afar, Ethiopia: implications for the function of the world's oldest stone tools. *J. Hum. Evol.* **48**, 109–121 (2005).

440. J. de Heinzelin, *et al.*, Environment and behavior of 2.5-million-year-old Bouri hominids. *Science* **284**, 625–629 (1999).

441. A. B. Stahl, *et al.*, Hominid Dietary Selection Before Fire [and Comments and Reply]. *Curr. Anthropol.* **25**, 151–168 (1984).

442. R. N. Carmody, R. W. Wrangham, Cooking and the human commitment to a high-quality diet. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 427–434 (2009).

443. E. Singels, *et al.*, Foraging potential of underground storage organ plants in the southern Cape, South Africa. *J. Hum. Evol.* **101**, 79–89 (2016).

444. R. B. Lee, "What hunters do for a living, or, how to make out on scarce resources" in *Man the Hunter*, (Routledge, 2017), pp. 30–48.

445. K. Milton, Hunter-gatherer diets—a different perspective. *Am. J. Clin. Nutr.* **71**, 665–667 (2000).

446. A. N. Crittenden, Ethnobotany in evolutionary perspective: wild plants in diet composition and daily use among Hadza hunter-gatherers. *Wild harvest: plants in the hominin and pre-agrarian human worlds*, 319–339 (2016).

447. K. D. Zink, D. E. Lieberman, Impact of meat and Lower Palaeolithic food processing techniques on chewing in humans. *Nature* **531**, 500–503 (2016).

448. R. N. Carmody, G. S. Weintraub, R. W. Wrangham, Energetic consequences of thermal and nonthermal food processing. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19199–19203 (2011).

449. S. L. Schnorr, A. N. Crittenden, K. Venema, F. W. Marlowe, A. G. Henry, Assessing digestibility of Hadza tubers using a dynamic in-vitro model. *Am. J. Phys. Anthropol.* **158**, 371–385 (2015).

450. A. Stahl, 11 Plant-food processing: implications for dietary quality. *Foraging farming* **31**, 171 (2014).

451. A. Crowther, The differential survival of native starch during cooking and implications for archaeological analyses: a review. *Archaeological and Anthropological Sciences* **4**, 221–235 (2012).

452. R. N. Carmody, *et al.*, Cooking shapes the structure and function of the gut microbiome. *Nat Microbiol* **4**, 2052–2063 (2019).

453. R. Wrangham, R. Carmody, Human adaptation to the control of fire. *Evol. Anthropol.* **19**, 187–199 (2010).

454. J. A. J. Gowlett, R. W. Wrangham, Earliest fire in Africa: towards the convergence of archaeological evidence and the cooking hypothesis. *Azania: Archaeological Research in Africa* **48**, 5–30 (2013).

455. W. Roebroeks, P. Villa, On the earliest evidence for habitual use of fire in Europe. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5209–5214 (2011).

456. R. Shimelmitz, *et al.*, "Fire at will": The emergence of habitual fire use 350,000 years ago. *Journal of Human Evolution* **77**, 196–203 (2014).

457. J. A. J. Gowlett, J. W. K. Harris, D. Walton, B. A. Wood, Early archaeological sites, hominid remains and traces of fire from Chesowanja, Kenya. *Nature* **294**, 125–129 (1981).

458. F. Berna, *et al.*, Microstratigraphic evidence of in situ fire in the Acheulean strata of Wonderwerk Cave, Northern Cape province, South Africa. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E1215–20 (2012).

459. N. Goren-Inbar, *et al.*, Evidence of hominin control of fire at Gesher Benot Ya'aqov, Israel. *Science* **304**, 725–727 (2004).

460. Y. Melamed, M. E. Kislev, E. Geffen, S. Lev-Yadun, N. Goren-Inbar, The plant component of an Acheulian diet at Gesher Benot Ya 'aqov, Israel. *Proceedings of the National Academy of Sciences* **113**, 14674–14679 (2016).

461. V. Wobber, B. Hare, R. Wrangham, Great apes prefer cooked food. *J. Hum. Evol.* **55**, 340–348 (2008).

462. J. M. Marston, J. D. Guedes, C. Warinner, *Method and Theory in Paleoethnobotany* (University Press of Colorado, 2015).

463. J. M. Sept, Beyond bones: archaeological sites, early hominid subsistence, and the costs and benefits of exploiting wild plant foods in east African riverine landscapes. *J. Hum. Evol.* **27**, 295–320 (1994).

464. B. L. Hardy, Neanderthal behaviour and stone tool function at the Middle Palaeolithic site of La Quina, France. *Antiquity* **78**, 547–565 (2004).

465. R. N. E. Barton, *et al.*, Gibraltar Neanderthals and results of recent excavations in Gorham's, Vanguard and Ibex Caves. *Antiquity* **73**, 13–23 (1999).

466. B. L. Hardy, M. Kay, A. E. Marks, K. Monigal, Stone tool function at the paleolithic sites of Starosele and Buran Kaya III, Crimea: behavioral implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10972–10977 (2001).

467. A. G. Henry, A. S. Brooks, D. R. Piperno, Microfossils in calculus demonstrate consumption of plants and cooked foods in Neanderthal diets (Shanidar III, Iraq; Spy I and II, Belgium). *Proc. Natl. Acad. Sci. U. S. A.* **108**, 486–491 (2011).

468. A. G. Henry, H. F. Hudson, D. R. Piperno, Changes in starch grain morphologies from cooking. *J. Archaeol. Sci.* **36**, 915–922 (2009).

469. A. V. Thoms, A. R. Laurence, L. Short, M. Kamiya, Baking Geophytes and Tracking Microfossils: Taphonomic Implications for Earth-Oven and Paleodietary Research. *Journal of Archaeological Method and Theory* **22**, 1038–1070 (2015).

470. H. Barton, R. Torrence, Cooking up recipes for ancient starch: assessing current methodologies and looking to the future. *J. Archaeol. Sci.* **56**, 194–201 (2015).

471. D. G. Drucker, *et al.*, Isotopic analyses suggest mammoth and plant in the diet of the oldest anatomically modern humans from far southeast Europe. *Sci. Rep.* **7**, 6833 (2017).

472. C. Koebnick, C. Strassner, I. Hoffmann, C. Leitzmann, Consequences of a long-term raw food diet on body weight and menstruation: results of a questionnaire survey. *Ann. Nutr. Metab.* **43**, 69–79 (1999).

473. V. de Freitas, N. Mateus, Structural Features of Procyanidin Interactions with Salivary Proteins. *Journal of Agricultural and Food Chemistry* **49**, 940–945 (2001).

474. L. Kandra, G. Gyémánt, Á. Zajácz, G. Batta, Inhibitory effects of tannin on human salivary α-amylase. *Biochem. Biophys. Res. Commun.* **319**, 1265–1271 (2004).

475. D. A. Granger, K. T. Kivlighan, M. El-Sheikh, E. B. Gordis, L. R. Stroud, Salivary α-Amylase in Biobehavioral Research: Recent Developments and Applications. *Ann. N. Y. Acad. Sci.* **1098**, 122–144 (2007).

476. M. C. Janiak, Of starch and spit. *Elife* **8** (2019).

477. G. K. Aji, F. J. Warren, E. Roura, Salivary α-Amylase Activity and Starch-Related Sweet Taste Perception in Humans. *Chem. Senses* **44**, 249–256 (2019).

478. A. L. Mandel, C. Peyrot des Gachons, K. L. Plank, S. Alarcon, P. A. S. Breslin, Individual differences in AMY1 gene copy number, salivary α-amylase levels, and the perception of oral starch. *PLoS One* **5**, e13352 (2010).

479. T. J. Lapis, M. H. Penner, A. S. Balto, J. Lim, Oral Digestion and Perception of Starch: Effects of Cooking, Tasting Time, and Salivary α-Amylase Activity. *Chem. Senses* **42**, 635–645 (2017).

480. J. G. Tye, R. C. Karn, A. D. Merritt, Differential expression of salivary (Amy1) and pancreatic (Amy2) human amylase loci in prenatal and postnatal development. *J. Med. Genet.* **13**, 96–102 (1976).

481. S. McKeen, *et al.*, Infant Complementary Feeding of Prebiotics for the Microbiome and Immunity. *Nutrients* **11**, 364 (2019).

482. C. L. Meehan, J. W. Roulette, Early supplementary feeding among central African foragers and farmers: a biocultural approach. *Soc. Sci. Med.* **96**, 112–120 (2013).

483. G. E. Kennedy, From the ape's dilemma to the weanling's dilemma: early weaning and its evolutionary context. *J. Hum. Evol.* **48**, 123–145 (2005).

484. L. T. Humphrey, Weaning behaviour in human evolution. *Semin. Cell Dev. Biol.* **21**, 453–461 (2010).

485. R. N. Carmody, *et al.*, Genetic Evidence of Human Adaptation to a Cooked Diet. *Genome Biology and Evolution* **8**, 1091–1103 (2016).

486. T. D. Hubbard, *et al.*, Divergent Ah Receptor Ligand Selectivity during Hominin Evolution. *Mol. Biol. Evol.* **33**, 2648–2658 (2016).

487. S. D. M. Brown, W. Wurst, R. Kühn, J. M. Hancock, The functional annotation of mammalian genomes: the challenge of phenotyping. *Annu. Rev. Genet.* **43**, 305–333 (2009).

488. J. P. Demuth, T. De Bie, J. E. Stajich, N. Cristianini, M. W. Hahn, The evolution of mammalian gene families. *PLoS One* **1**, e85 (2006).

489. S. W. Cheetham, G. J. Faulkner, M. E. Dinger, Overcoming challenges and dogmas to understand the functions of pseudogenes. *Nat. Rev. Genet.* **21**, 191–201 (2020).

490. L. F. Franchini, K. S. Pollard, Human evolution: the non-coding revolution. *BMC Biol.* **15**, 89 (2017).

491. S. L. Schnorr, K. Sankaranarayanan, C. M. Lewis Jr, C. Warinner, Insights into human evolution from ancient and contemporary microbiome studies. *Curr. Opin. Genet. Dev.* **41**, 14–26 (2016).

492. C. Warinner, C. Speller, M. J. Collins, C. M. Lewis Jr, Ancient human microbiomes. *J. Hum. Evol.* **79**, 125–136 (2015).

493. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

494. C. Camacho, *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

495. H. Thorvaldsdóttir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).

496. R. Bouckaert, *et al.*, BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **10**, e1003537 (2014).

497. R. Bouckaert, *et al.*, BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **15**, e1006650 (2019).

498. S. Tavaré, “Some probabilistic and statistical problems in the analysis of DNA sequences” in *Lectures on Mathematics in the Life Sciences.*, R. M. Miura, Ed. (American Mathematical Society, 1986), pp. 57–86.

499. A. J. Drummond, A. Rambaut, B. Shapiro, O. G. Pybus, Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005).

500. D. Albanese, C. Donati, Strain profiling and epidemiology of bacterial species from metagenomic sequencing. *Nat. Commun.* **8**, 2260 (2017).

501. D. T. Truong, A. Tett, E. Pasolli, C. Huttenhower, N. Segata, Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* **27**, 626–638 (2017).

502. S. Nayfach, B. Rodriguez-Mueller, N. Garud, K. S. Pollard, An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res.* **26**, 1612–1625 (2016).

503. M. Scholz, *et al.*, Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat. Methods* **13**, 435–438 (2016).

504. T. Daley, A. D. Smith, Predicting the molecular complexity of sequencing libraries. *Nat. Methods* **10**, 325–327 (2013).

505. C. F. P. Scholz, M. Kilian, The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus Propionibacterium to the proposed novel genera Acidipropionibacterium gen. nov., Cutibacterium gen. nov. and Pseudopropionibacterium gen. nov. *Int. J. Syst. Evol. Microbiol.* **66**, 4422–4432 (2016).

506. A. Andrades Valtueña, *et al.*, The Stone Age Plague and Its Persistence in Eurasia. *Curr. Biol.* **27**, 3683–3691.e8 (2017).

507. K. I. Bos, *et al.*, Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature* **514**, 494–497 (2014).

508. M. Zolfo, *et al.*, Profiling microbial strains in urban environments using metagenomic sequencing data. *Biol. Direct* **13**, 9 (2018).

509. A. T. Duggan, *et al.*, 17th Century Variola Virus Reveals the Recent History of Smallpox. *Curr. Biol.* **26**, 3407–3412 (2016).

510. T. Jombart, adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).

511. T. Jombart, I. Ahmed, adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).

512. T. H. Jukes, C. R. Cantor, Evolution of protein molecules. In “Mammalian Protein Metabolism, Volume III”.(Ed. HN Munro.) pp. 21--132 (1969).

513. L. G. Straus, M. R. González Morales, T. Higham, M. Richards, S. Talamo, Radiocarbon Dating the Late Upper Paleolithic of Cantabrian Spain: El Mirón Cave Date List IV. *Radiocarbon* **57**, 183–188 (2015).

514. S. Sankararaman, *et al.*, The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).

515. I. Lazaridis, *et al.*, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).

516. W. Haak, *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).

517. V. Villalba-Mouco, *et al.*, Survival of Late Pleistocene Hunter-Gatherer Ancestry in the Iberian Peninsula. *Curr. Biol.* **29**, 1169–1177.e7 (2019).

518. M. van de Loosdrecht, *et al.*, Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. *Science* **360**, 548–552 (2018).

519. F. Maixner, *et al.*, The 5300-year-old Helicobacter pylori genome of the Iceman. *Science* **351**, 162–165 (2016).

520. E. A. Franzosa, *et al.*, Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* **15**, 962–968 (2018).

521. L. A. David, *et al.*, Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).

522. R. Overbeek, *et al.*, The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* **33**, 5691–5702 (2005).

523. D. T. Truong, *et al.*, MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015).

524. B. E. Suzek, *et al.*, UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31**, 926–932 (2015).

525. R. Wrangham, Control of Fire in the Paleolithic: Evaluating the Cooking Hypothesis. *Curr. Anthropol.* **58**, S303–S313 (2017).

526. C. S. Larsen, Bioarchaeology: Interpreting Behavior from the Human Skeleton. Cambridge Univ Press. *Cambridge, UK* (2015).

527. L. Cordain, *et al.*, Origins and evolution of the Western diet: health implications for the 21st century. *Am. J. Clin. Nutr.* **81**, 341–354 (2005).

528. B. Wood, M. Grabowski, “Macroevolution in and Around the Hominin Clade” in *Macroevolution: Explanation, Interpretation and Evidence*, E. Serrelli, N. Gontier, Eds. (Springer International Publishing, 2015), pp. 345–376.