LETTER

doi:10.1038/nature21674

Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus

Laura S. Weyrich[1], Sebastian Duchene[2], Julien Soubrier[1], Luis Arriola[1], Bastien Llamas[1], James Breen[1], Alan G. Morris[3], Kurt W. Alt[4,5,6,7], David Caramelli[8], Veit Dresely[5,6], Milly Farrell[9], Andrew G. Farrer[1], Michael Francken[10], Neville Gully[11], Wolfgang Haak[1], Karen Hardy[12,13], Katerina Harvati[10], Petra Held[14], Edward C. Holmes[2], John Kaidonis[11], Carles Lalueza-Fox[15], Marco de la Rasilla[16], Antonio Rosas[17], Patrick Semal[18], Arkadiusz Soltysiak[19], Grant Townsend[11], Donatella Usai[20], Joachim Wahl[21], Daniel H. Huson[22], Keith Dobney[23,24,25] & Alan Cooper[1]

Recent genomic data have revealed multiple interactions between Neanderthals and modern humans[1], but there is currently little genetic evidence regarding Neanderthal behaviour, diet, or disease. Here we describe the shotgun-sequencing of ancient DNA from five specimens of Neanderthal calcified dental plaque (calculus) and the characterization of regional differences in Neanderthal ecology. At Spy cave, Belgium, Neanderthal diet was heavily meat based and included woolly rhinoceros and wild sheep (mouflon), characteristic of a steppe environment. In contrast, no meat was detected in the diet of Neanderthals from El Sidrón cave, Spain, and dietary components of mushrooms, pine nuts, and moss reflected forest gathering[2,3]. Differences in diet were also linked to an overall shift in the oral bacterial community (microbiota) and suggested that meat consumption contributed to substantial variation within Neanderthal microbiota. Evidence for self-medication was detected in an El Sidrón Neanderthal with a dental abscess[4] and a chronic gastrointestinal pathogen (*Enterocytozoon bieneusi*). Metagenomic data from this individual also contained a nearly complete genome of the archaeal commensal *Methanobrevibacter oralis* (10.2× depth of coverage)—the oldest draft microbial genome generated to date, at around 48,000 years old. DNA preserved within dental calculus represents a notable source of information about the behaviour and health of ancient hominin specimens, as well as a unique system that is useful for the study of long-term microbial evolution.

Neanderthals remain our closest known, extinct, hominin relatives, who co-existed and occasionally interbred with anatomically modern humans across Eurasia in the Late Pleistocene epoch[1]. Neanderthals became extinct in Europe around 40,000 years ago (40 ka), although the extinction process across the rest of Eurasia is less clear[5]. Archaeological and isotopic data from the last glacial cycle (around 120–12 ka) suggest that Neanderthals were as carnivorous as polar bears or wolves[6] with a diet heavily based on large terrestrial herbivores, such as reindeer, woolly mammoth, and woolly rhinoceros[7]. By contrast, microwear analysis of tooth surfaces from Neanderthals in different ecological settings, such as wooded areas or open plains, suggests that diets were guided by local food availability[3]. Analyses of phytoliths, starch granules, and proteins from calcified dental calculus also indicate that Neanderthal diets included many plants, including some that were used for medicinal purposes[8]. As a result, Neanderthal diet remains a topic of considerable debate, with limited data on the specific animals and plants directly consumed or the potential effects on Neanderthal health and disease.

Although genomic studies continue to reveal evidence of interbreeding between anatomically modern humans and Neanderthals across Eurasia[9], little is known about the health consequences of these interactions. The genetic analysis of Neanderthal dental calculus represents an opportunity to examine this issue and to reconstruct Neanderthal diet, behaviour, and disease[10]. Here, we report the first genetic analysis of dental calculus from five Neanderthals (two individuals from El Sidrón cave in Spain; two individuals from Spy cave in Belgium; and one individual from Breuil Grotta in Italy) and compare these data to a historic wild-caught chimpanzee ($n = 1$) and modern human ($n = 1$), as well as to low coverage sequencing of calculus from a wide-range of ancient humans (Supplementary Table 1). To provide increased resolution of the diseases that may have affected Neanderthals, we also deeply sequenced (>147 million reads) dental calculus from the best-preserved Neanderthal, El Sidrón 1, which suffered from a dental abscess[4].

Size-based PCR-amplification biases can confound standard metabarcoding analyses (for example, sequencing of 16S ribosomal (r)RNA amplicons[11,12]) of ancient dental calculus[13]. Consequently, we compared metagenomic-shotgun sequencing and 16S rRNA amplicon (V4 region) analyses of the Neanderthal dental calculus specimens—by far the oldest examined to date. The 16S amplicon datasets were not representative of the biodiversity revealed by shotgun sequencing (Extended Data Figs 1, 2 and Supplementary Tables 2, 3, 7, 16), as samples clustered according to methodolgy (Fig. 1) and contained disproportionately large amounts of non-oral microorganisms that were environmental contaminants (Supplementary Tables 2, 7 and Extended Data Figs 3, 4a). As a result, the 16S amplicon datasets were excluded from downstream analysis, along with the Neanderthal sample from the Breuil Grotta, which failed to produce amplifiable sequences.

The shotgun datasets consisted of short DNA fragments (<70 bp), which complicated accurate identification of bacterial species using standard software, such as MG-RAST or DIAMOND[14] (Extended Data Fig. 4b, see Supplementary Information). To circumvent this problem, we benchmarked and used a new metagenomic alignment

[1]Australian Centre for Ancient DNA, School of Biological Sciences and The Environment Institute, University of Adelaide, Adelaide, South Australia, Australia. [2]Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Life and Environmental Sciences and Sydney Medical School, University of Sydney, Sydney, Australia. [3]Department of Human Biology, University of Cape Town, Cape Town, South Africa. [4]Danube Private University, Krems, Austria. [5]State Office for Heritage Management and Archaeology, Saxony-Anhalt, Germany. [6]Heritage Museum, Halle, Germany. [7]Institute for Prehistory and Archaeological Science, Basel University, Switzerland. [8]Department of Biology, University of Florence, Florence, Italy. [9]Human Origins and Palaeo Environments Group, Oxford Brookes University, Oxford, UK. [10]Paleoanthropology, Senckenberg Centre for Human Evolution and Paleoenvironments, Eberhard Karls University of Tübingen, Tübingen, Germany. [11]School of Dentistry, The University of Adelaide, Adelaide, Australia. [12]Catalan Institution for Research and Advanced Studies (ICREA), Pg Lluís Companys 23, 08010 Barcelona, Catalonia, Spain. [13]Departament de Prehistòria, Facultat de Filosofia i Lletres, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain. [14]Institute of Anthropology, University of Mainz, Mainz, Germany. [15]Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain. [16]Área de Prehistoria, Departamento de Historia, Universidad de Oviedo, Oviedo, Spain. [17]Paleoanthropology Group, Department of Paleobiology, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. [18]Scientific Service Heritage, Royal Belgian Institute of Natural Sciences, Brussels, Belgium. [19]Department of Bioarchaeology, Institute of Archaeology, University of Warsaw, Warsaw, Poland. [20]Istituto Italiano per l'Africa e l'Oriente (IsIAO), Rome, Italy. [21]State Office for Cultural Heritage Management Baden-Württemberg, Esslingen, Germany. [22]Department of Algorithms in Bioinformatics, University of Tübingen, Tübingen, Germany. [23]Department of Archaeology, Classics and Egyptology, School of Histories, Languages and Cultures, University of Liverpool, Liverpool, UK. [24]Department of Archaeology, University of Aberdeen, Aberdeen, UK. [25]Department of Archaeology, Simon Fraser University, Burnaby, British Columbia, Canada.

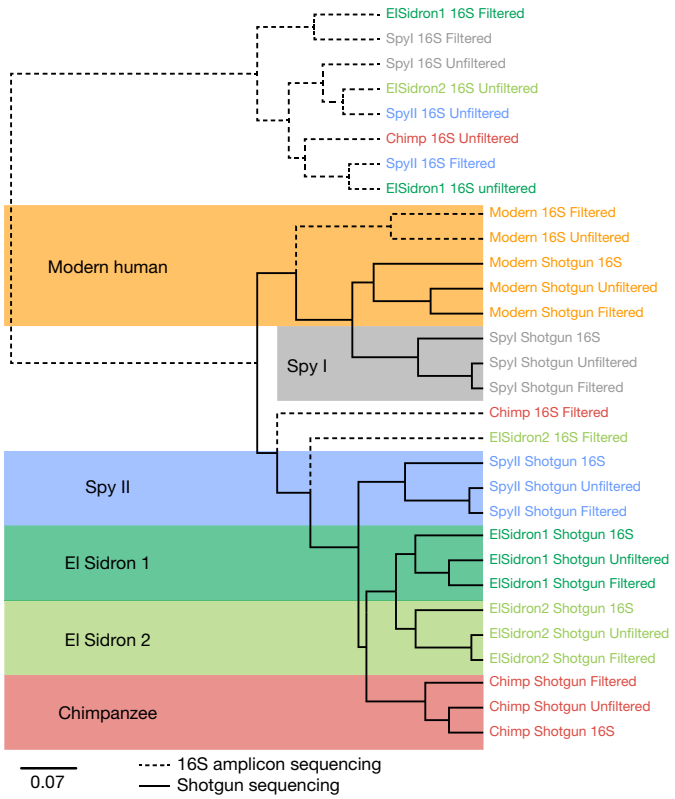


Figure 1 | Comparison of 16S amplicon and shotgun sequencing datasets obtained from ancient, historic, and modern dental calculus samples. Filtered and unfiltered 16S rRNA amplicon and shotgun sequencing datasets, as well as the 16S rRNA shotgun sequences identified using GraftM, were compared using UPGMA clustering of Bray–Curtis distances from a wild-caught chimpanzee (red), Neanderthals (El Sidrón 1 (dark green), El Sidrón 2 (light green), Spy I (grey), Spy II (blue), and a modern human (orange) ($n = 6$ total samples).

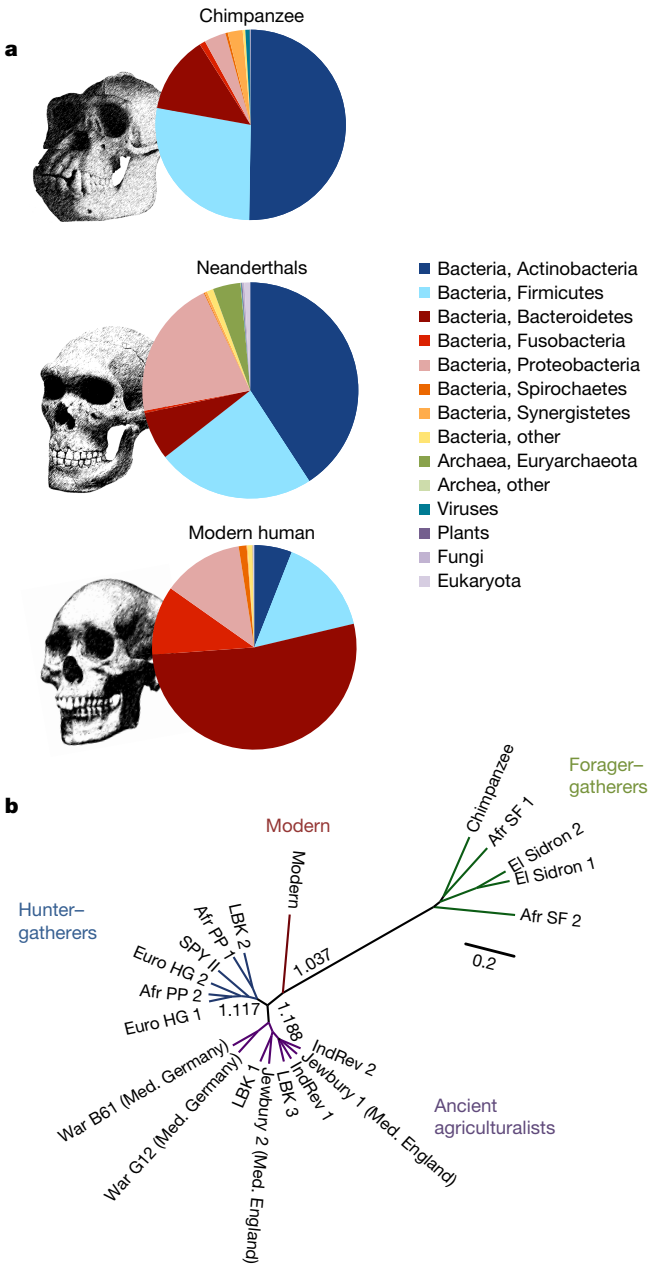


Figure 2 | Bacterial community composition at the phyla level of oral microbiota from chimpanzee, Neanderthal and modern human samples. a, Oral microbiota from shotgun sequencing datasets of a wild-caught chimpanzee (top), Neanderthals ($n = 3$; middle) and a modern human (bottom) are presented at the phyla level. Names of the phyla were simplified for clarity, and unidentified reads were excluded. Gram-positive (blue) and Gram-negative (red) phyla are differentiated by colour. **b**, UPGMA clustering of Bray–Curtis distances are displayed for 22 oral metagenomes, revealing a strong correlation with meat consumption. The scale bar represents differences in Bray–Curtis distances. The UPGMA clustering reveals four distinct groupings: Forager-Gatherers, Hunter-Gatherers, Ancient Agriculturalists, and the Modern human. Definitions for abbreviations can be found in the Supplementary Table 1: Spy and El Sidron, Neanderthals; Afr SF, African forager; LBK, Early European farmer; Afr PP, African Pastoralist period; Euro HG, European hunter–gatherer; Jewbury, UK Medieval; War, German (published German Medieval data sets from ref. 12) Medieval; and modern, modern human (C10).

tool, MALTX, which rapidly identifies species from shorter fragment lengths using a rapid BLASTX-like algorithm[15] (Extended Data Fig. 5 and Supplementary Tables 4, 5, 6, 7). Bioinformatic filtering showed that the Spy Neanderthals samples were more heavily affected by environmental contamination (Extended Data Fig. 6 and Supplementary Table 7). Indeed, shotgun sequences from Spy I had DNA damage patterns characteristic of contamination with modern DNA sequences (Extended Data Fig. 7), clustered more closely to the modern individual than to other Neanderthals (Fig. 1), and contained similar diversity to environmental samples (Extended Data Fig. 8). Therefore, this individual was also excluded from further analyses. The shotgun-sequencing datasets from the three remaining, robust Neanderthal samples (El Sidrón 1, El Sidrón 2 and Spy II) contained an average of 93.76% bacterial, 5.91% archaeal, 0.27% eukaryotic, and 0.06% viral identifiable sequences, similar to previously published ancient and modern human dental calculus[12] (Fig. 2a and Extended Data Fig. 6c). The six dominant bacterial phyla in the modern human mouth (Actinobacteria, Firmicutes, Bacteroidetes, Fusobacteria, Proteobacteria, and Spirochaetes) were also dominant in each of the Neanderthals with an average of 222 bacterial species per individual (Fig. 2a and Extended Data Fig. 6c).

We first examined Neanderthal diets using the eukaryotic diversity preserved within the dental calculus (see Supplementary Information). Calculus from the Spy II individual contained high numbers of reads mapping to rhinoceros (*Ceratotherium simum*) and sheep (*Ovis aries*), as well as the edible grey shag mushroom (*Coprinopsis cinerea*) (Table 1). Bones of woolly rhinoceros, reindeer, mammoth, and horses were present in Spy Cave[16], while wild mouflon sheep were broadly distributed in Europe throughout the Pleistocene[17]. Woolly rhinoceros has long been suspected to be part of the Spy Neanderthal diet[18], confirming the highly carnivorous lifestyle inferred from the isotope and dental microwear data obtained from the Spy individuals[3,6,19].

Table 1 | Dietary information preserved in calculus

Scientific name	Common name of probable source	Hominid pathogen or medicinal use	El Sidrón 1	El Sidrón 2	Spy II	Chimpanzee	Modern human	Laboratory control (EBC)	Spy I*
Zymoseptoria tritici	Plant (wheat) pathogen		4.13%	0	0	0	0	0	2.87%
Phaeosphaeria nodorum	Plant (wheat) pathogen		12.22%	0	0	3.98%	0	0	0.45%
Penicillium rubens	Food fungus	MU	3.97%	0	0	0	0	0	1.35%
Myceliophthora thermophila	Cellulose fungus		0	0	0.56%	0	0	0	0.13%
Coprinopsis cinerea	Edible mushroom (grey shag)		0	0	2.44%	0	0	0	0
Schizophyllum commune	Edible mushroom (split gill)		3.65%	0	0	0	0	0	0.10%
Malassezia globosa	Human fungal commensal		3.65%	8.89%	0	0	19.92%	0	5.49%
Enterocytozoon bieneusi	Intracellular parasite (microsporidia)	HP	8.10%	0	0	0	0	0	0
Ovis aries	Sheep (wild mouflon)		0	0	62.03%	0	0	0	1.17%
Ceratotherium simum	White rhinoceros (woolly rhinoceros)		0	0	34.40%	0	0	0	0.11%
*Ixodes scapularis**	Tick		0	0	0	0	2.15%	0	0.15%
Physcomitrella patens	Moss		2.06%	0	0	0	0	0	0.09%
Pinus koraiensis	Pine tree		13.49%	19.60%	0	4.45%	0	0	0.40%
Populus trichocarpa	Poplar tree	MU	2.86%	0	0	0	0	0	0.44%
Total eukaryotic reads			630	551	532	427	3,760	5	25,294

DNA sequences mapping to eukaryotic species are shown as a proportion of the total eukaryotic reads identified within each sample. Eukaryotic sequences were also identified in the extraction blank controls (EBCs) and the Spy I Neanderthal, which is heavily contaminated with modern DNA; these samples are shown to the right.
HP, human pathogen; MU, medicinal use.
*Samples or taxa that are probably the results of contamination, as they do not represent biological processes (see Supplementary Information).

These findings also support recent isotope evidence that suggests Spy Neanderthals were regularly consuming mushrooms[20].

The dietary profile in El Sidrón Neanderthals was markedly different from Spy, and contained no sequences matching large herbivores or suggesting high meat consumption. However, reads mapping to edible mushrooms (split gill; *Schizophyllum commune*), pine nuts (*Pinus koraiensis*), forest moss (*Physcomitrella patens*), and poplar (*Populus trichocarpa*) were identified (Table 1). Sequences mapping to plant fungal pathogens were also observed (*Zymoseptoria tritici, Phaeosphaeria nodorum, Penicillium rubens*, and *Myceliophthora thermophila*), suggesting that the El Sidrón Neanderthals may have consumed moulded herbaceous material. Limited zooarchaeological evidence exists for the El Sidrón individuals, and our first genetic description of their diet supports evidence that Neanderthal groups across Europe used multiple subsistence strategies according to location and food availability[2,3]. Additional approaches are needed to verify and extend these dietary reconstructions[17] (see Supplementary Information).

Our findings support previous suggestions that El Sidrón 1 may have been self-medicating a dental abscess[8]. This was the only individual whose calculus included sequences corresponding to poplar, which contains the natural pain-killer salicylic acid (the active ingredient in aspirin), and also notably contained sequences of the natural antibiotic producing *Penicillium* from the moulded herbaceous material. The sample from this individual also included sequences matching the intracellular eukaryotic pathogen microsporidia (*Enterocytozoon bieneusi*), which causes acute diarrhoea in humans[21], indicating another health issue that potentially required self-medication.

To examine how oral microorganisms (microbiota) in Neanderthals reflected dietary composition, we compared the filtered shotgun data to a wide range of ancient calculus specimens from humans with varying diets, including ancient African gatherers from the Later Stone Age; individuals from the African Pastoralist Period with high meat consumption[22]; European hunter–gatherers with a diet that included a wide range of protein sources; and early European farmers with diets largely based around high carbohydrate and milk consumption (see Supplementary Information). We clustered Bray–Curtis distances using the unweighted-pair group method with arithmetic means (UPGMA) and found four distinct groups: forager–gatherers with limited meat consumption (El Sidrón Neanderthals, chimpanzee, and African gatherers from the Later Stone Age); hunter–gatherers (or pastoralists) with a frequent meat diet (Spy Neanderthal, African pastoralists, and European hunter–gatherers); ancient agriculturalists (European farming individuals); and modern humans (Fig. 2b and Extended Data Fig. 9a). This analysis identifies a split between

Table 2 | Draft microbial genomes present in El Sidrón 1

					Mapped Reads							
Reference Genome	Sequence reference number	Length (Mb)	GC content (%)	Bases covered (Mb)	Unique hits	Depth (average coverage)	Average read length	5′-C-T	3′-G-A	ΔD	ΔS	λ
Methanobrevibacter oralis JMR01	NZ_CBWS00000000	2.107	27.8	0.941	370115	15.16	58.67	0.33	0.36	0.05	1	0.38
Candidatus Saccharibacteria oral TM7	NZ_CP007496.1	0.705	44.5	0.131	108919	5.83	52.46	0.37	0.41	0.01	1	0.38
Campylobacter gracilis ATCC 33236	NZ_CP012196.1	2.282	46.6	1.199	94472	2.40	51.7	0.38	0.41	0.01	1	0.36
Propionibacterium propionicum F0230a	NZ_018142.1	3.449	66.1	2.083	130748	1.89	48.85	0.37	0.43	0	1	0.43
Fretibacterium fastidiosum	gi 296110870	2.728	55.5	1.466	121822	2.43	48	0.39	0.43	0	1	0.41
Eubacterium infirmum F0142	NZ_AGWI00000000	1.9	40.1	0.176	52170	10.73	51.53	0.33	0.38	0.02	1	0.41
Peptostreptococcus stomatis DSM 17678	GCF_000147675.1	1.988	36.7	1.222	94743	2.90	54.62	0.36	0.4	0.02	1	0.38
Eubacterium sphenum ATCC 49989	NZ_GG688422.1	1.084	40.6	0.261	23124	3.46	52.87	0.37	0.41	0.03	1	0.36

Eight draft microbial genomes from Gram-positive bacteria, Gram-negative bacteria, eubacteria, and archaea were obtained from the deep-sequenced El Sidrón 1 specimen by read mapping. The sequence coverage, GC content, sequencing depth, and damage profile (average fragment length and base-pair modifications calculated using mapDamage 2.0) are shown for each genome. ΔD, ΔS, and λ are outputs from mapDamage 2.0 and represent aspects of ancient DNA damage identified in the samples.

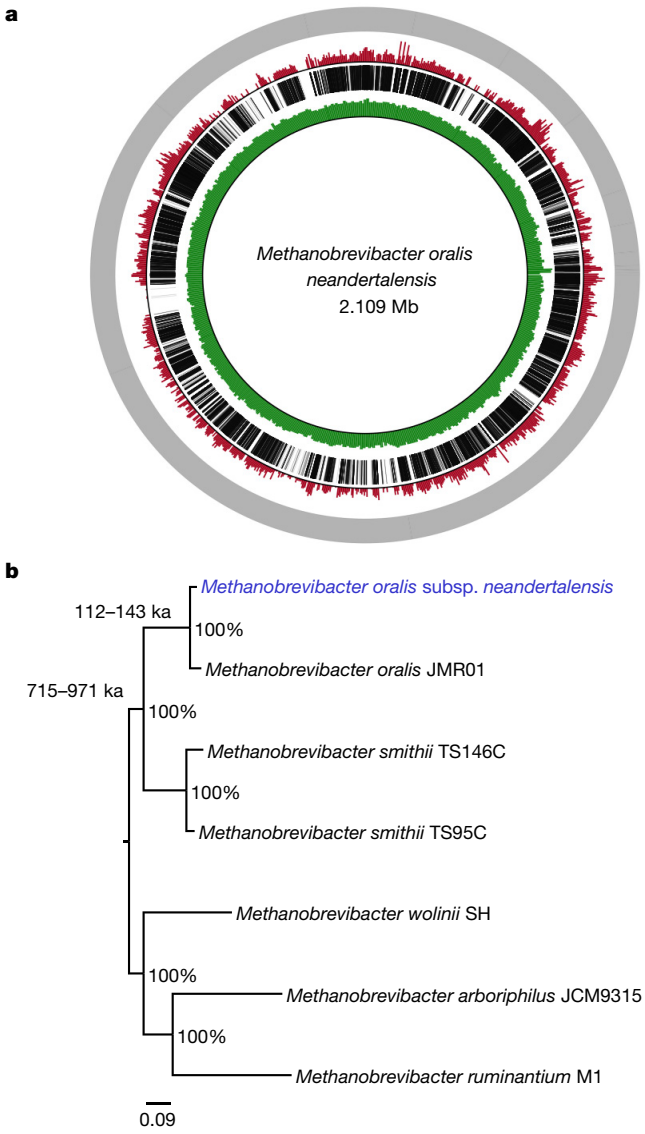


Figure 3 | Draft genome and phylogeny of a 48,000-year-old archaeon, *Methanobrevibacter oralis* subsp. *neandertalensis*. **a**, Ancient sequences mapping to *Methanobrevibacter oralis* JMR01 are displayed in a circos plot (black), alongside the depth of coverage obtained (red; scale, 0–2,757 sequences). The reference sequence is displayed (grey) with the GC content of the reference sequence calculated in 2,500 bp bins (green; 0–0.4852% GC). **b**, A *Methanobrevibacter* phylogeny was constructed from alignment of the whole genome in RAxML with 100 bootstrap replicates, with the per cent support shown in each node. The estimated dates were calculated from a whole genome phylogeny using a Bayesian methodology (in BEAST) assuming a strict clock model (see Supplementary Information).

hunter–gatherers and agriculturalists, as previously observed[11], but also shows two distinct gatherer groups (hunter–gatherers and forager–gatherers), which are apparently differentiated by the quantity of meat consumed in their diet. As such, meat consumption appears to have affected early hominin microbiota, analogous to carnivorous and herbivorous mammals[23]. This also suggests that microorganisms preserved in dental calculus can be used to record details of dietary behaviour in ancient hominins.

We also examined the microbial diversity of the Neanderthal samples for potential pathogens as a sign of disease. Neanderthal microbiota were more similar to the historic chimpanzee sample than the modern human sample and contained less potentially pathogenic Gram-negative species, which are associated with secondary enamel colonization, increased plaque formation, and periodontal disease[24] (18.9% Gram-negative bacteria in Neanderthals compared to 77.6% in the modern human; Extended Data Fig. 9b). All types of microbial taxa were equally damaged and fragmented, suggesting this difference is not simply owing to a preservation bias in Gram-negative microorganisms, as previously reported[11] (Table 2; see Supplementary Information). The low levels of immunostimulatory Gram-negative taxa in Neanderthals may be related to the reduced presence of Fusobacteria (Extended Data Fig. 6c), which can facilitate the binding of Gram-negative microorganisms to the Gram-positive primary colonizers that bind tooth enamel (*Streptococcus*, *Actinomyces*, and *Methanobrevibacter* species)[25]. Notably, the increased diversity of Gram-negative immunostimulatory taxa in modern humans is strongly linked to a wide-range of Western diseases[26].

Several oral pathogens could be identified within the shotgun sequencing data, although the short ancient sequences and diverse metagenomic background complicated authentication. We established a number of criteria to verify the presence of specific bacterial pathogens, including the assessment of ancient DNA damage, phylogenetic position, and bioinformatic comparisons to close relatives (see Supplementary Information). We identified the caries-associated species *Streptococcus mutans* (0.08%–0.18%) and the members of the 'red complex' associated with modern periodontal disease (*Porphyromonas gingivalis*, 0–0.52%; *Tannerella forsythia*, 0.05–2.4%; and *Treponema denticola*, 0–1.87%), consistent with evidence of dental

Table 3 | Purifying and positive selection in *M. oralis neandertalensis*

Gene number	CDS	Genebank	d_N/d_S ratio	Gene annotation	Coding protein function
Gene1211	1184	WP_042693702.1	0	NZ_HG796201.1	Preprotein translocase subunit SecG
Gene291	283	WP_042691749.1	0	NZ_HG796199.1	SAM-dependent methyltransferase
Gene303	295	WP_042691777.1	0	NZ_HG796199.1	Fibrillarin
Gene343	343	WP_042691868.1	0	NZ_HG796199.1	Sugar-fermentation stimulation protein SfsA
Gene394	394	WP_042691937.1	0	NZ_HG796199.1	30S ribosomal protein S2
Gene401	401	WP_042691950.1	0	NZ_HG796199.1	Transcriptional regulator
Gene745	745	WP_042692741.1	0	NZ_HG796200.1	50S ribosomal protein L37
Gene757	757	WP_042693268.1	0	NZ_HG796200.1	Acyltransferase
Gene766	766	WP_042692795.1	0	NZ_HG796200.1	DNA-directed RNA polymerase
Gene769	769	WP_042692805.1	0	NZ_HG796200.1	30S ribosomal protein S6
Gene772	772	WP_042692815.1	0	NZ_HG796200.1	50S ribosomal protein L24
Gene773	773	WP_042692817.1	0	NZ_HG796200.1	30S ribosomal protein
Gene810	810	WP_042692911.1	0	NZ_HG796200.1	Transcriptional regulator
Gene836	836	WP_042692956.1	0	NZ_HG796200.1	Endonuclease DDE
Gene880	880	WP_042693050.1	1.52	NZ_HG796200.1	Uracil transporter
Gene724	724	WP_042692699.1	2.67	NZ_HG796200.1	Acetylesterase
Gene269	269	WP_042691703.1	3.64	NZ_HG796199.1	Conjugal-transfer protein TraB
Gene1206	1206	WP_042693692.1	12	NZ_HG796201.1	DNA mismatch-repair protein MutT

The ratio of non-synonymous to synonymous substitutions per site (d_N/d_S) was calculated for coding regions with sufficient coverage and that were conserved between *M. oralis* and *M. oralis* subsp. *neandertalensis*. Genes that have undergone strong purifying ($d_N/d_S < 0.1$) or positive ($d_N/d_S > 1$) selection are displayed if the function of the gene was annotated. Hypothetical proteins and those not matching the *M. oralis* genome during BLAST searches are not shown. CDS, coding DNA sequence.

caries and periodontal disease in Neanderthals[27] (Supplementary Tables 9–11). A variety of other pathogens (*Bordetella parapertusssis*, *Pasteurella multocida*, *Neisseria gonorrhoeae*, *Streptococcus pyogenes*, and *Corynebacterium diptheriae*) were identified but could not be unambiguously distinguished from closely related commensal oral taxa (Extended Data Fig. 10a and Supplementary Tables 9, 14), highlighting the need for rigorous criteria when identifying pathogenic strains from ancient metagenomic data.

Lastly, we examined Neanderthal commensal microorganisms. Within the El Sidrón 1 specimen that was deeply sequenced, we were able to recover draft genomes (>1× depth of coverage) for the eight most prevalent microbial species (Table 2). Of particular note was a dominant archaeal species (14.7%; Extended Data Fig. 6c) in El Sidrón 1 that was present in lower proportions in other Neanderthals (1.4% and 1.2% in El Sidrón 2 and Spy II, respectively). The large differences in G/C content between bacteria and archaea facilitated efficient read mapping (Table 2) to the modern-human-associated *Methanobrevibacter oralis* JMR01 strain. At around 48 thousand years (kyr)[28], *Methanobrevibacter oralis* subsp. *neandertalensis* is the oldest draft microbial genome to date (44.7% of 2.1 Mb, with a 10.3× depth of coverage; Table 2 and Fig. 3).

Date estimates using a strict molecular clock place the divergence between the *M. oralis* strains of Neanderthals and modern humans between 112–143 ka (95% highest posterior density interval; mean date of 126 ka) (Fig. 3b; see Supplementary Information). As this is long after the genomic divergence of Neanderthals and modern humans (450–750 ka)[29], it appears that commensal microbial species were transferred between the two hosts during subsequent interactions, potentially in the Near East[30]. Further genome comparisons revealed 136 coding sequences in the modern human *M. oralis* that were putatively absent in *M. oralis* subsp. *neandertalensis* (Supplementary Table 15), including genes encoding antiseptic resistance (*qacE*), maltose metabolism regulation (*sfsA*), and bacterial immunity (CRISPR Cas2 and Cas6; Supplementary Table 15), which probably reflect dietary and hygiene differences between modern humans and Neanderthals. A comparison of 375 translatable protein-coding sequences from the *M. oralis* strains from Neanderthals and modern humans indicated that 58% were under strong purifying selection ($d_N/d_S < 0.1$) (Table 3, see Supplementary Information). Only 4% appeared to be under putative positive selection ($d_N/d_S > 1$) and include regions for conjugal gene transfer (that is, uptake of foreign or plasmid DNA; *traB*) and DNA-mismatch repair (*mutT*).

Preserved dental calculus represents a notable source of information about behaviour, diet, and health of ancient hominin specimens, as well as a unique long-term system that can be used to study how hundreds of different microbial species have evolved and spread among hominins.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 2 August 2016; accepted 30 January 2017.
Published online 8 March 2017.

1. Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328,** 710–722 (2010).
2. Fiorenza, L. *et al.* Molar macrowear reveals Neanderthal eco-geographic dietary variation. *PLoS One* **6,** e14769 (2011).
3. El Zaatari, S., Grine, F. E., Ungar, P. S. & Hublin, J.-J. Neandertal versus modern human dietary responses to climatic fluctuations. *PLoS One* **11,** e0153277 (2016).
4. Rosas, A. *et al.* Paleobiology and comparative morphology of a late Neandertal sample from El Sidron, Asturias, *Spain. Proc. Natl Acad. Sci. USA* **103,** 19266–19271 (2006).
5. Higham, T. *et al.* The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* **512,** 306–309 (2014).
6. Richards, M. P. & Trinkaus, E. Isotopic evidence for the diets of European Neanderthals and early modern humans. *Proc. Natl Acad. Sci. USA* **106,** 16034–16039 (2009).
7. Bocherens, H., Drucker, D. G., Billiou, D., Patou-Mathis, M. & Vandermeersch, B. Isotopic evidence for diet and subsistence pattern of the Saint-Césaire I Neanderthal: review and use of a multi-source mixing model. *J. Hum. Evol.* **49,** 71–87 (2005).
8. Hardy, K. *et al.* Neanderthal medics? Evidence for food, cooking, and medicinal plants entrapped in dental calculus. *Naturwissenschaften* **99,** 617–626 (2012).
9. Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524,** 216–219 (2015).
10. Weyrich, L. S., Dobney, K. & Cooper, A. Ancient DNA analysis of dental calculus. *J. Hum. Evol.* **79,** 119–124 (2015).
11. Adler, C. J. *et al.* Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nat. Genet.* **45,** 450–455, e1 (2013).
12. Warinner, C. *et al.* Pathogens and host immunity in the ancient human oral cavity. *Nat. Genet.* **46,** 336–344 (2014).
13. Ziesemer, K. A. *et al.* Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. *Sci. Rep.* **5,** 16498 (2015).
14. Meyer, F. *et al.* The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9,** 386 (2008).
15. Herbig, A. *et al.* MALT: fast alignment and analysis of metagenomic DNA sequence data applied to the Tyrolean Iceman. Preprint at http://biorxiv.org/content/early/2016/04/27/050559 (2016).
16. Germonpré, M., Udrescu, M. & Fiers, E. The fossil mammals of Spy. *Anthropologica et Præhistorica* **123/2012,** 298–327 (2013).
17. Crégut-bonnoure, E. Nouvelles données paléogéographiques et chronologiques sur les *Caprinae* (Mammalia, Bovidae) du Pléistocène moyen et supérieur d'Europe. *Euskomedia* **57,** 205–219 (2005).
18. Patou-Mathis, M. Neanderthal subsistence behaviours in Europe. *Int. J. Osteoarchaeol.* **10,** 379–395 (2000).
19. Naito, Y. I. *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).
20. O'Regan, H. J., Lamb, A. L. & Wilkinson, D. M. The missing mushrooms: searching for fungi in ancient human dietary analysis. *J. Archaeol. Sci.* **75,** 139–143 (2016).
21. Tumwine, J. K. *et al. Enterocytozoon bieneusi* among children with diarrhea attending Mulago Hospital in Uganda. *Am. J. Trop. Med. Hyg.* **67,** 299–303 (2002).
22. Leonard, W. R. & Crawford, M. H. *The Human Biology of Pastoral Populations* (Cambridge Univ. Press, 2002).
23. Muegge, B. D. *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332,** 970–974 (2011).
24. Marsh, P. D. & Bradshaw, D. J. Dental plaque as a biofilm. *J. Ind. Microbiol.* **15,** 169–175 (1995).
25. Signat, B., Roques, C., Poulet, P. & Duffaut, D. Role of *Fusobacterium nucleatum* in periodontal health and disease. *Curr. Issues Mol. Biol.* **13,** 25–36 (2011).
26. Cho, I. & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nat. Rev. Genet.* **13,** 260–270 (2012).
27. Topić, B., Rašćić-Konjhodžić, H. & Cižek Sajko, M. Periodontal disease and dental caries from Krapina Neanderthal to contemporary man—skeletal studies. *Acta Med. Acad.* **41,** 119–130 (2012).
28. Wood, R. E. *et al.* A new date for the Neanderthals from El Sidrón Cave (Asturias, Northern Spain)*. *Archaeometry* **55,** 148–158 (2013).
29. Stringer, C. The origin and evolution of *Homo sapiens*. *Phil. Trans. R. Soc. B* **371,** 20150237 (2016).
30. Sankararaman, S., Patterson, N., Li, H., Pääbo, S. & Reich, D. The date of interbreeding between Neandertals and modern humans. *PLoS Genet.* **8,** e1002947 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Manzi, the Odontological Collection of the Royal College of Surgeons, Royal Belgian Institute of Natural Sciences, Museo Nacional de Ciencias Naturales, and Adelaide University for access to dental calculus material. We thank A. Croxford for DNA sequencing and A. Walker, J. Krause and A. Herbig for feedback. The Australian Research Council supported this work through the Discovery Project and Fellowship schemes. We acknowledge the fundamental contribution of D. Brothwell (1933-2016) to this research by initiating the archaeological study of dental calculus.

Author Contributions L.S.W., K.D. and A.C. designed the study; A.G.M., K.W.A., D.C., V.D., M.Fa., M.Fr., N.G., W.H., K.Hard., K.Harv., P.H., J.K., C.L.F., M.d.l.R., A.R., P.S., A.S., D.U. and J.W. provided samples and interpretations of associated archaeological goods; L.S.W. performed experiments; L.S.W., S.D., E.C.H., J.S., B.L., J.B., L.A., A.G.F. and A.C. performed bioinformatics analysis and interpretation of the data; D.H.H. developed bioinformatics tools; N.G., J.K., and G.T. analysed medical relevance of data; L.S.W. and A.C. wrote the paper; and all authors contributed to editing the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to L.S.W. (laura.weyrich@adelaide.edu.au) or A.C. (alan.cooper@adelaide.edu.au).

Reviewer Information *Nature* thanks P. Ungar and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Sample handling and DNA extraction. All sample storage and molecular biology procedures before PCR amplification stages were carried out at the Australian Centre for Ancient DNA facility at The University of Adelaide. All experiments were performed within UV-treated, still-air hoods located in isolated, still-air rooms that have been designed to allow highly technical ancient-DNA research to be performed with ultra-low levels of background contamination (that is, workflow and background microbial contamination are monitored, facilities are irradiated with ultraviolet light each night, and the general facility is under positive air pressure). To minimize environmental contamination, each dental calculus sample was UV-treated for 15 min on each side, soaked in 2 ml of 5% bleach for 3 min, rinsed in 90% ethanol for 1 min, and dried at room temperature for several minutes. Immediately after decontamination, DNA extraction was performed using an in-house silica-based method, as previously described[31] but with decreased buffer volumes (1.8 ml lysis buffer (1.6 ml EDTA; 200 μl SDS; 20 μl 20 mg ml^{-1} proteinase K) and 3 ml guanidine DNA-binding buffer).

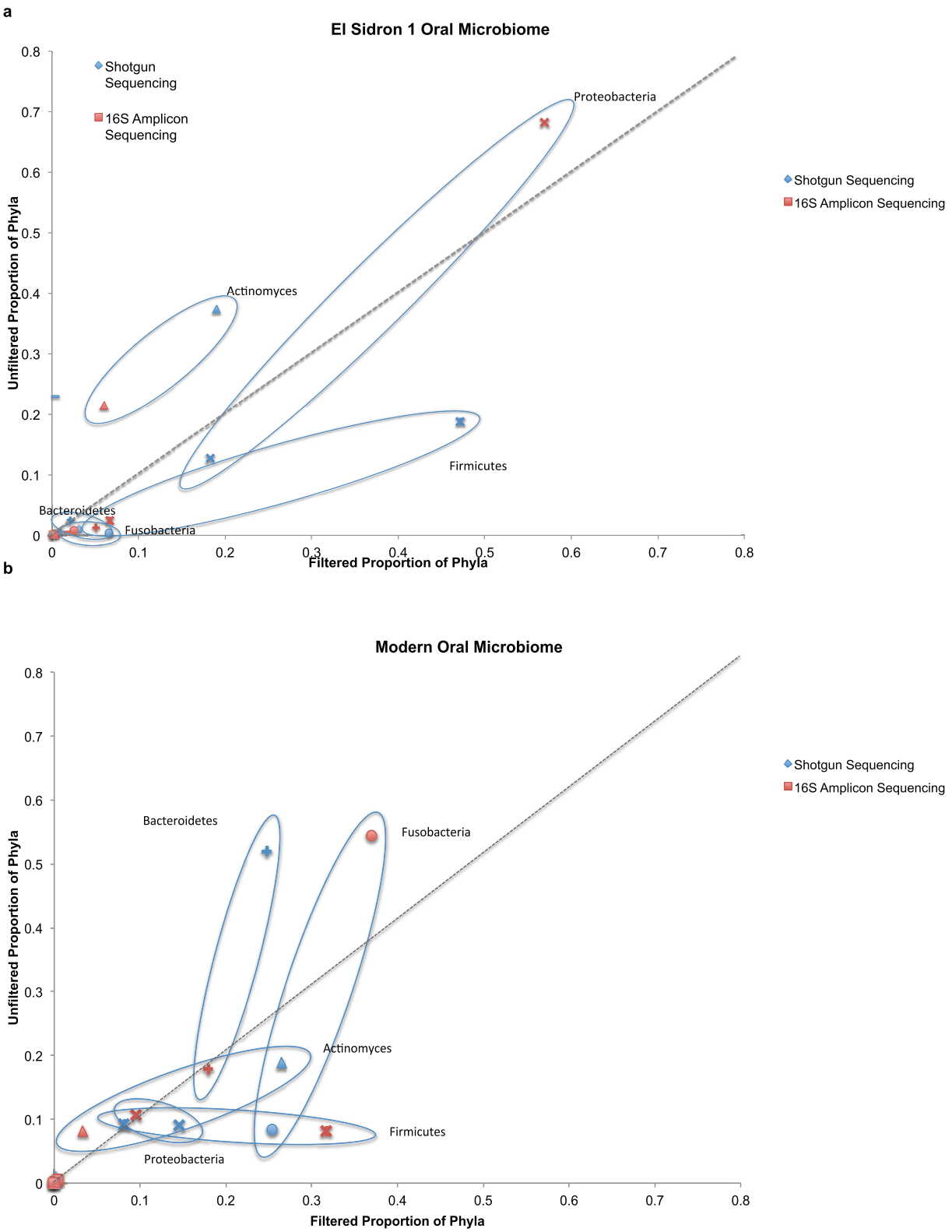
DNA library preparation and sequencing. After DNA was extracted, 16S ribosomal RNA amplicon libraries of the V4 region were constructed by PCR amplification[32]. Each sample was amplified in triplicate, and samples were pooled, Ampure cleaned, and quantified using a TapeStation and quantitative PCR (KAPA Illumina quantification kit), before sequencing with an Illumina MiSeq 300 cycle kit (approximately 40 samples per run). Frequent and repetitive extraction blank controls (EBCs) are used throughout all experimental procedures (that is, additional control samples were included during extraction, amplification, and library preparation). From the samples processed for 16S sequencing, several key specimens were selected for shotgun metagenomic sequencing. Shotgun metagenomic libraries were constructed as previously described[33], with 5-bp forward and reverse barcodes. Metagenomic shotgun libraries were Ampure cleaned, quantified using a TapeStation and quantitative PCR (KAPA Illumina quantification kit), and pooled at equimolar concentrations before sequencing.

16S rRNA amplicon library analysis. To process the 16S amplicon data, sequences were de-multiplexed using the CASAVA pipeline and joined into amplicons using fastq_joiner (ea-utils)[34]. Quality filtering and trimming was completed using Cutadapt, and sequences were then imported into QIIME 1.6.0 for analysis[35]. In QIIME, OTUs (operational taxonomic units) were clustered in UCLUST at 97%, and representative sequences were taxonomically identified using the Greengenes (gg_12_10) database[36]. After OTU selection, strict filtering was applied to all samples (see Supplementary Information Section II). Diversity was analysed in QIIME, and phylogenetic analysis was visualized in FigTree (http://tree.bio.ed.ac.uk/software/figtree). Statistical analysis was performed by anosim in QIIME, and the calculation of the Jaccard or Bray–Curtis indices, hierarchal clustering, and heatmap construction were completed in R using the vegan and gplots packages (http://cran.r-project.org).

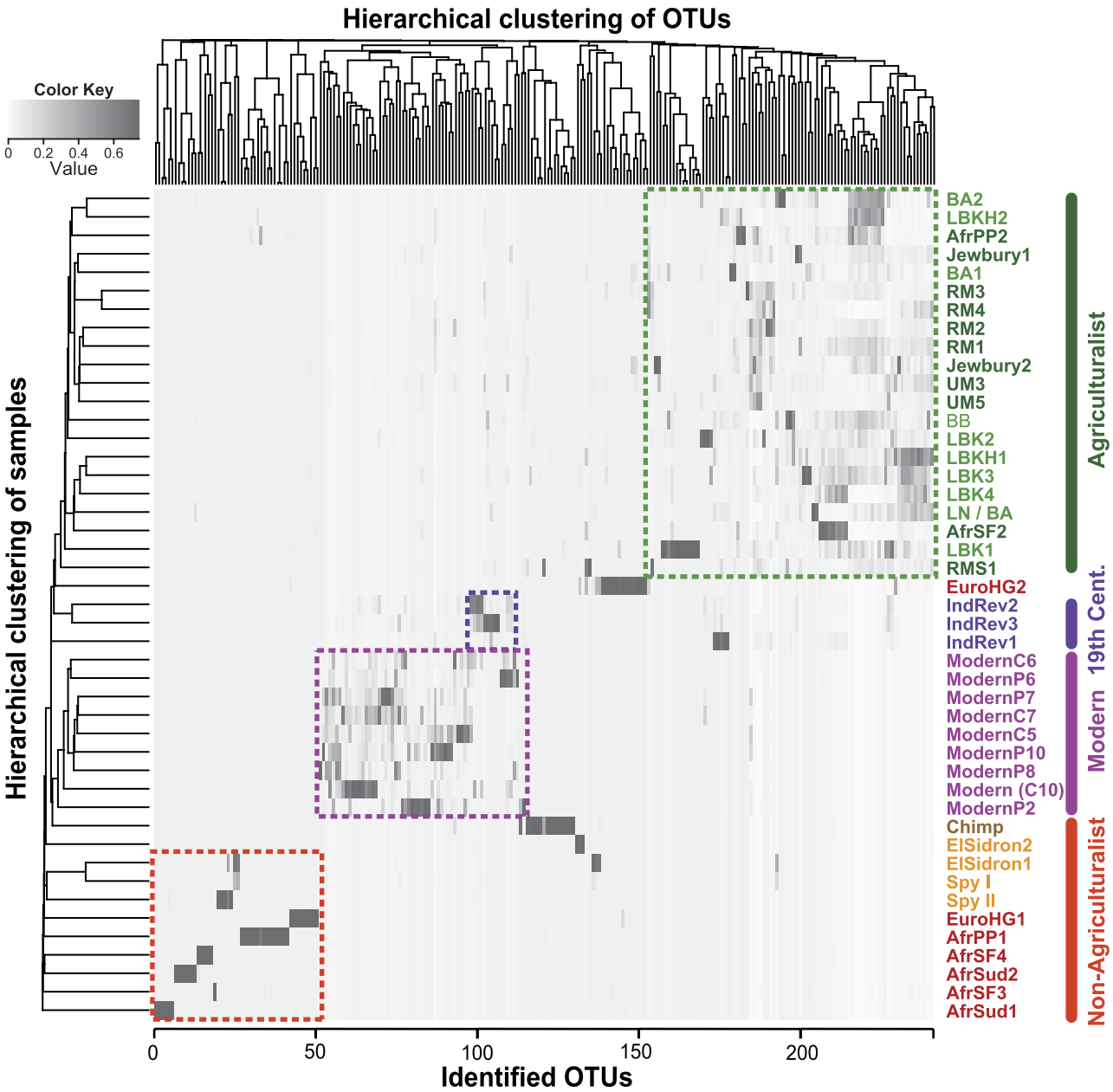
Shotgun DNA sequencing library analysis. To process shotgun metagenomic sequencing data, reads were merged with a 5-bp overlap using bbmerge, and reads matching the forward and reserve barcodes with no mismatches were retained using AdapterRemoval[37]. Taxonomic identifications were made using protein alignments in MetaPhlAn[38], MG-RAST[39], DIAMOND[40], and the new metagenome alignment tool with BLASTX-like approach (MALTX) developed in the Huson laboratory at the University of Tübingen[15]. Taxonomic assignments were then filtered using default LCA parameters in MEGAN5 (ref. 41), and data was exported at specific taxonomic classification levels (that is, phyla, genera, or species) for downstream analysis. Reference genomes were excluded if they were known to have human DNA contamination[42]. Statistical analyses were done using a Mann–Whitney *U* test (comparisons of phyla in one sample compared to other samples), a heteroscedastic *t*-test (direct comparisons between specific taxa in two samples or groups), or LefSe (identifying taxa that distinguish one group from another). Genomes were assembled by mapping to a reference genome using specific ancient DNA parameters in bwa[43], and authenticated using mapDamage 2.0 (ref. 44). Phylogenetic analyses were completed by mapping reads to reference genomes, aligning genomic sequences using progressiveMauve[45], and inferring trees in RAxML v8.1.21 (ref. 46) using the GTRGAMMA model. A Bayesian approach was used to estimate dates of divergence between strains and clades. Detailed descriptions of the procedures are available in the Supplementary Information.

Data and code availability. Raw and analysed sequence data and all analytical codes used in this study are publically available in the Online Ancient Gene Repository (https://www.oagr.org.au/) under the study 'Reconstructing Neandertal behaviour, diet, and disease using ancient DNA from dental calculus.' (http://dx.doi.org/10.4225/55/584775546a409).

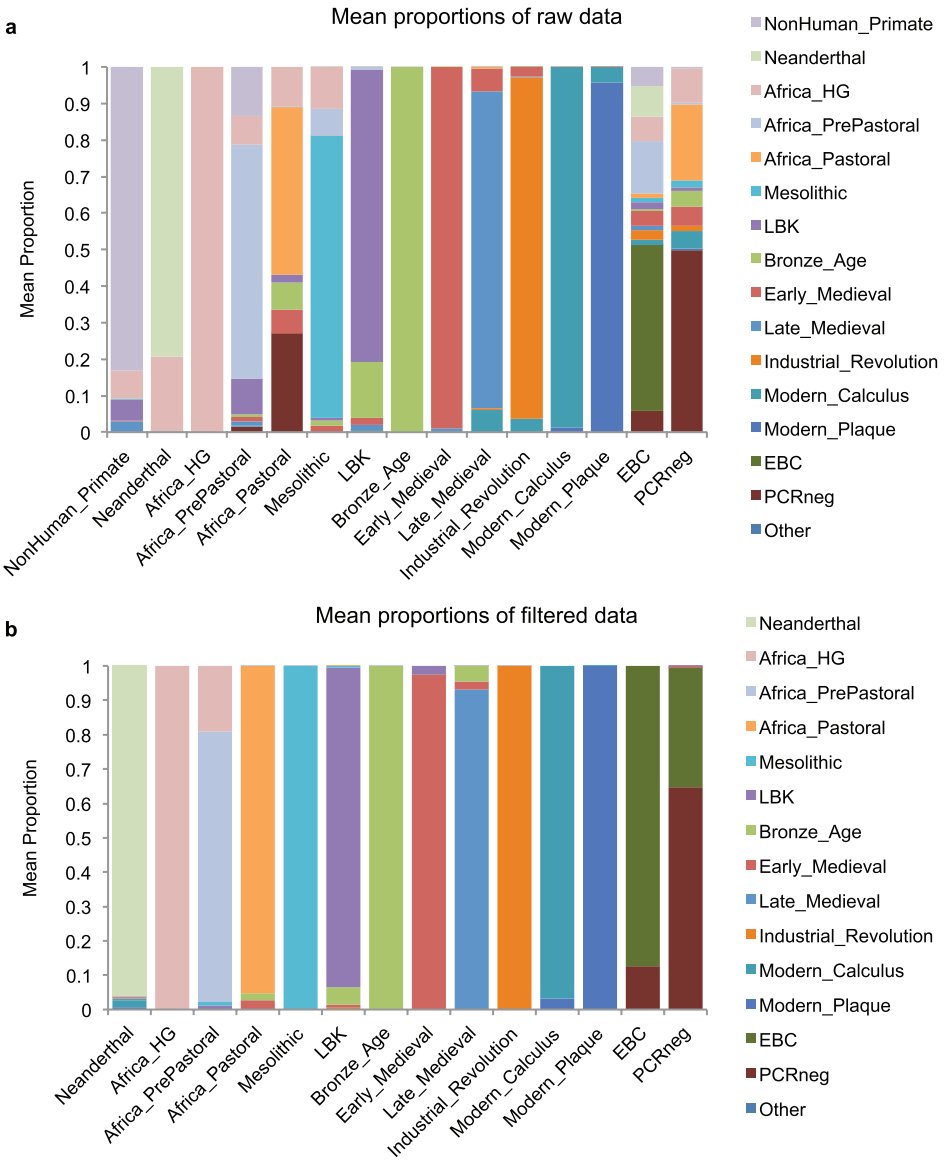
31. Brotherton, P. *et al.* Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat. Commun.* **4,** 1764 (2013).
32. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6,** 1621–1624 (2012).
33. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010,** pdb.prot5448 (2010).
34. Aronesty, E. *ea-utils: command-line tools for processing biological sequencing data.* (Expression Analysis, Durham, 2011).
35. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7,** 335–336 (2010).
36. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72,** 5069–5072 (2006).
37. Lindgreen, S. AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Res. Notes* **5,** 337 (2012).
38. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12,** R60 (2011).
39. Aziz, R. K. *et al.* The RAST server: rapid annotations using subsystems technology. *BMC Genomics* **9,** 75 (2008).
40. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12,** 59–60 (2015).
41. Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N. & Schuster, S. C. Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* **21,** 1552–1560 (2011).
42. Longo, M. S., O'Neill, M. J. & O'Neill, R. J. Abundant human DNA contamination identified in non-primate genome databases. *PLoS One* **6,** e16410 (2011).
43. Schubert, M. *et al.* Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* **13,** 178 (2012).
44. Ginolhac, A., Rasmussen, M., Gilbert, M. T., Willerslev, E. & Orlando, L. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* **27,** 2153–2155 (2011).
45. Darling, A. E., Mau, B. & Perna, N. T. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5,** e11147 (2010).
46. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22,** 2688–2690 (2006).



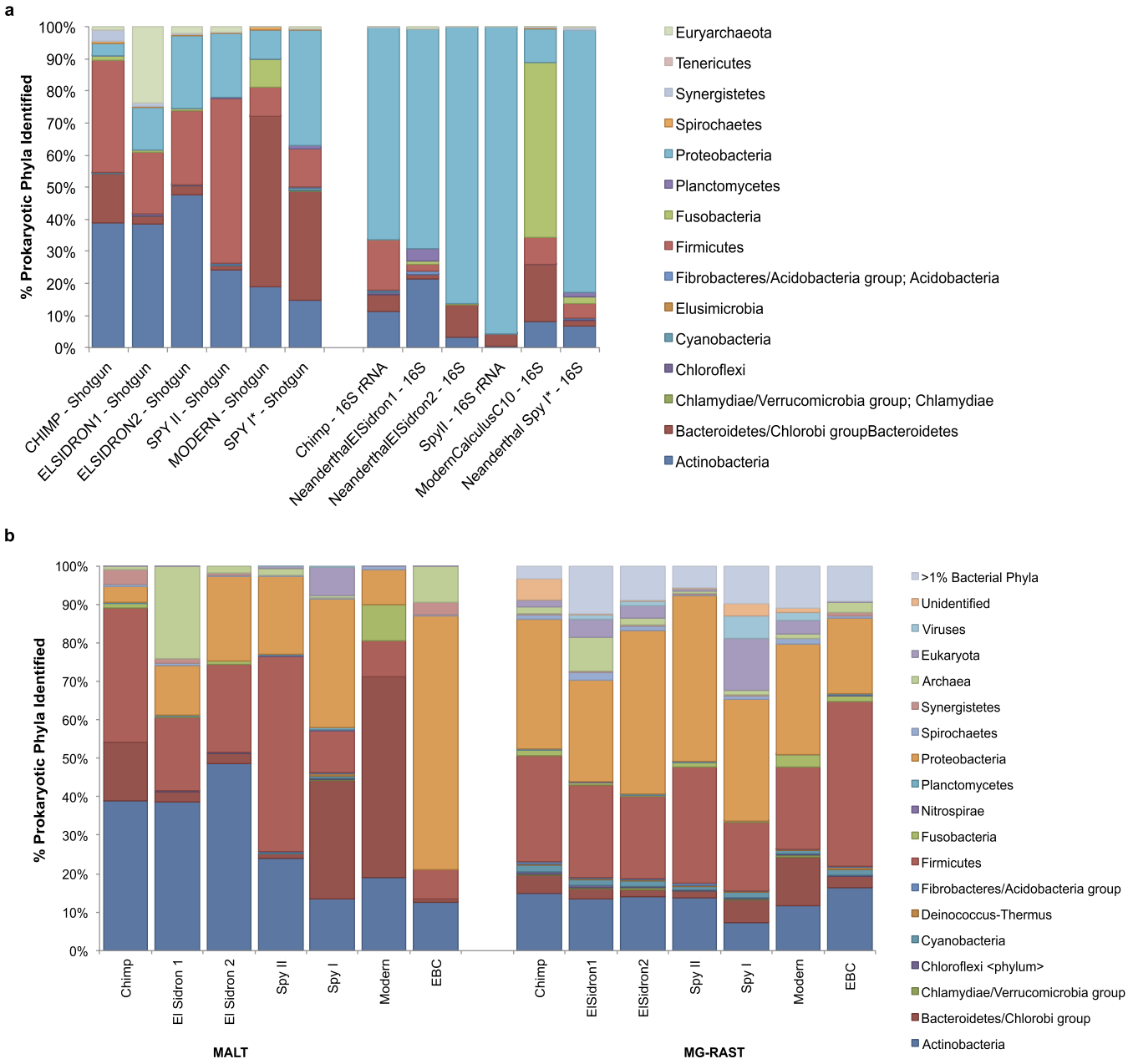
Extended Data Figure 1 | Proportions of bacterial phyla from filtered and unfiltered 16S amplicon and shotgun sequencing datasets. **a**, **b**, Proportions of bacterial phyla of El Sidrón 1(**a**) and the modern human oral microbiota were compared (**b**). Samples in blue are from shotgun sequencing datasets, whereas samples in red are from 16S amplicon datasets. The different shapes of each data point correspond to the microbial phyla, which are displayed next to each phyla grouping (for example, the cross represents Proteobacteria for both 16S and shotgun sequencing datasets).



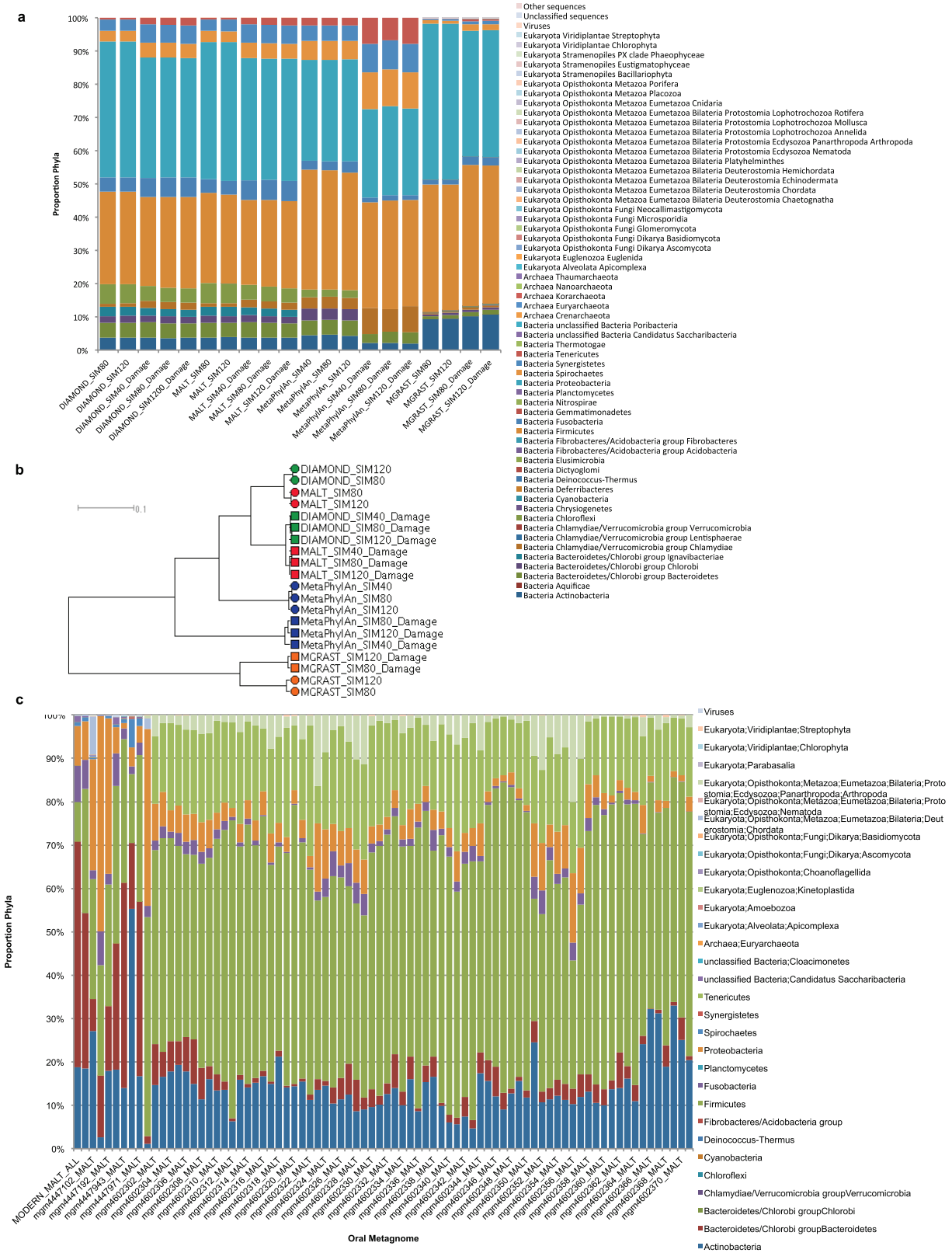
Extended Data Figure 2 | The presence/absence distance (Jaccard) calculated for each 16S OTU observed in the 99th percentile. OTUs from each sample were then clustered according to dissimilarity within each sample. Clusters of unique operational taxonomic units (OTUs) are identified (dashed lines) and labelled according to cluster relationships (red, no-agriculture; green, agriculture; purple, 19th century; fuchsia, modern time). Calculations are consistent with ancient DNA metagenomic analysis.



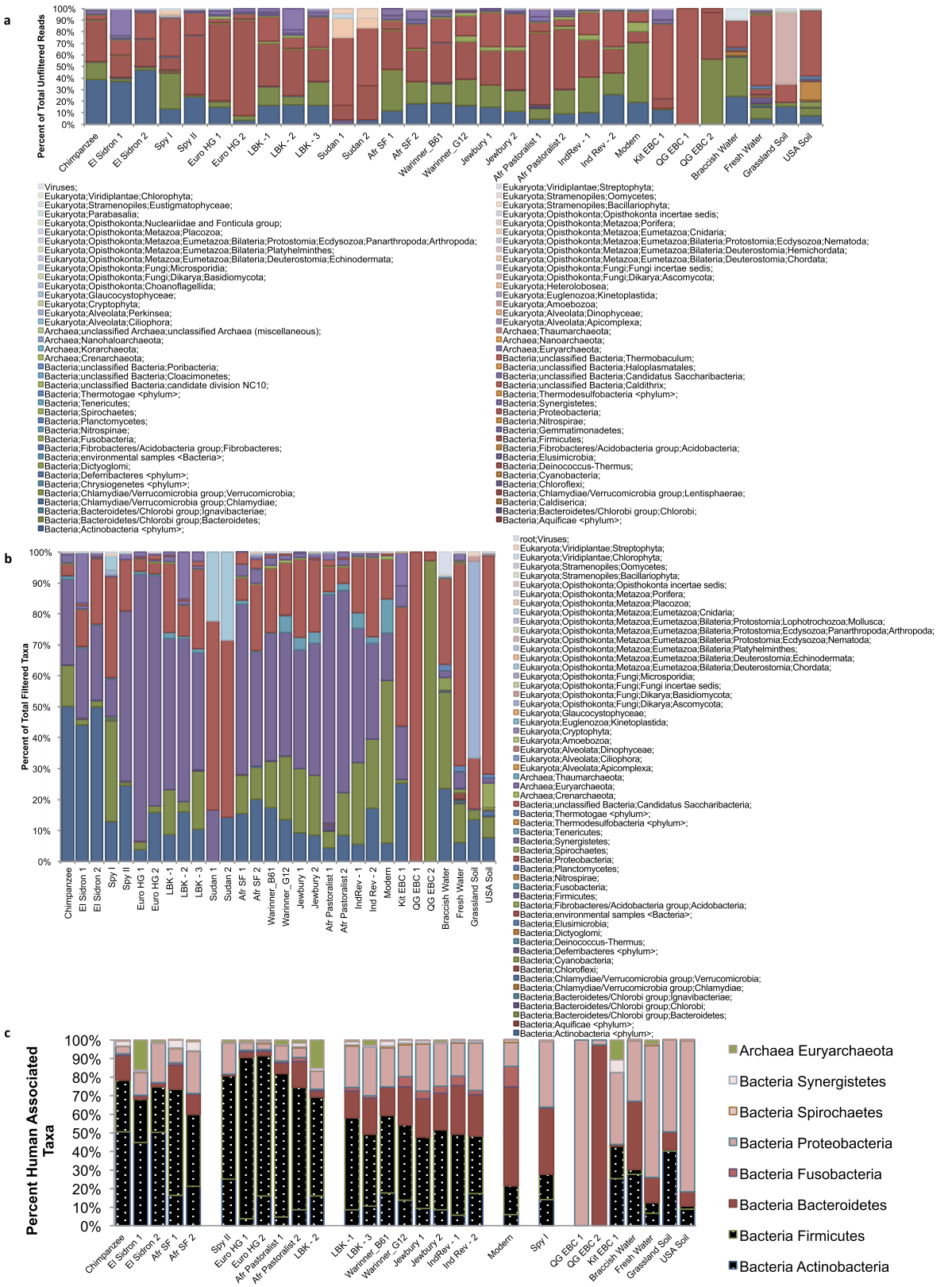
Extended Data Figure 3 | SourceTracker take-one-out analysis for all samples. a, **b**, Samples were grouped into time periods, and the proportion of each taxa originating from each sample group was inferred. Other, summed proportions across non-oral microbial groups (non-oral human microbiome, air, and soil) and unknown classification. Groups have a minimum of two samples (the non-human primate group is removed in filtered analysis as filtering reduced the sample number to one) and are displayed for the raw (unfiltered) OTU (**a**) ($n = 54$) and filtered OTU (**b**) data ($n = 42$).



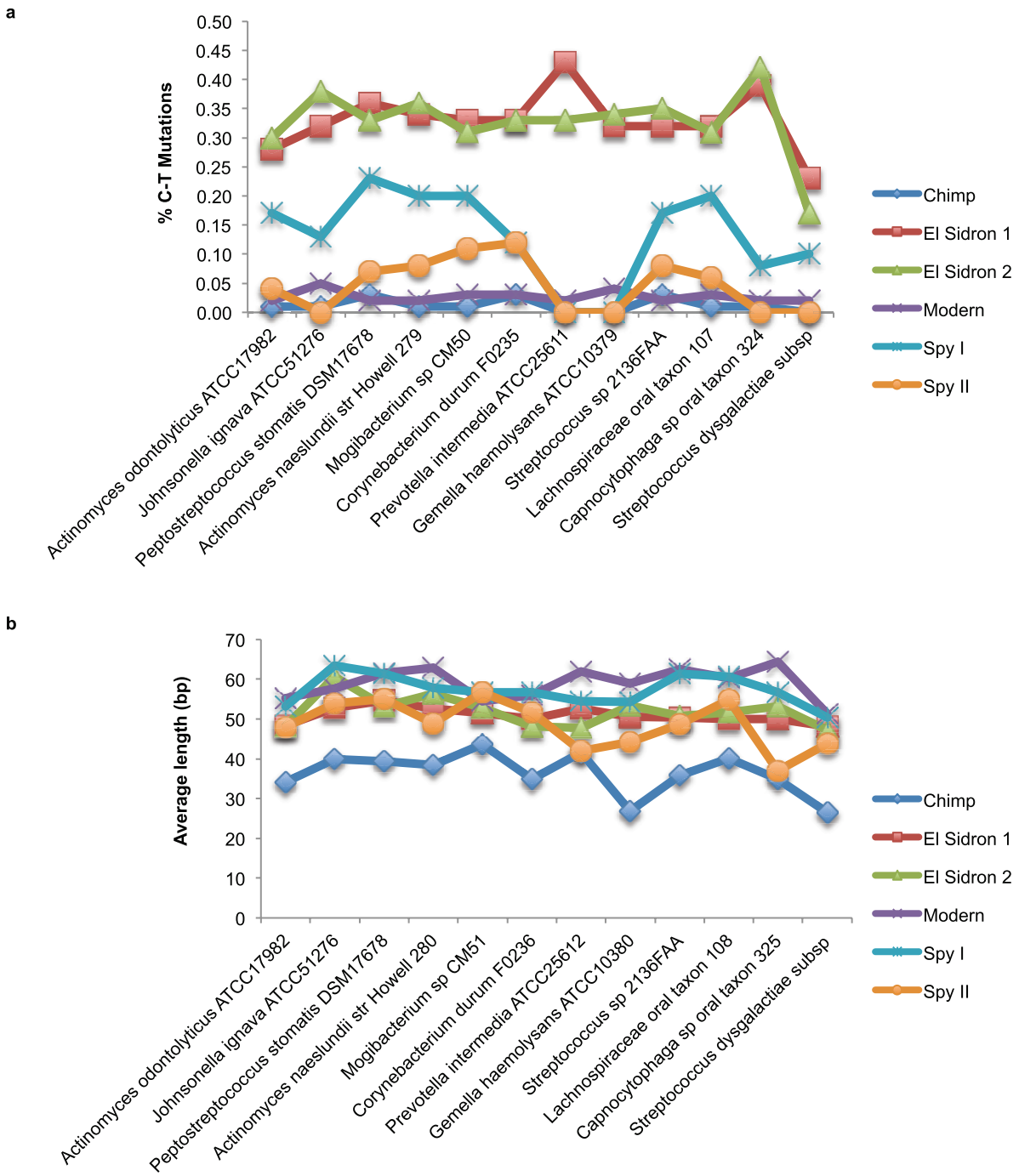
Extended Data Figure 4 | MALTX analysis compared to 16S and alternative shotgun analysis methods. a, Unfiltered prokaryotic phyla identified from 16S rRNA (QIIME) and shotgun sequencing results (MALTX) are compared. **b**, Raw shotgun sequences were analysed by MALTX and by MG-RAST, and bacterial phyla and kingdom level results are displayed.



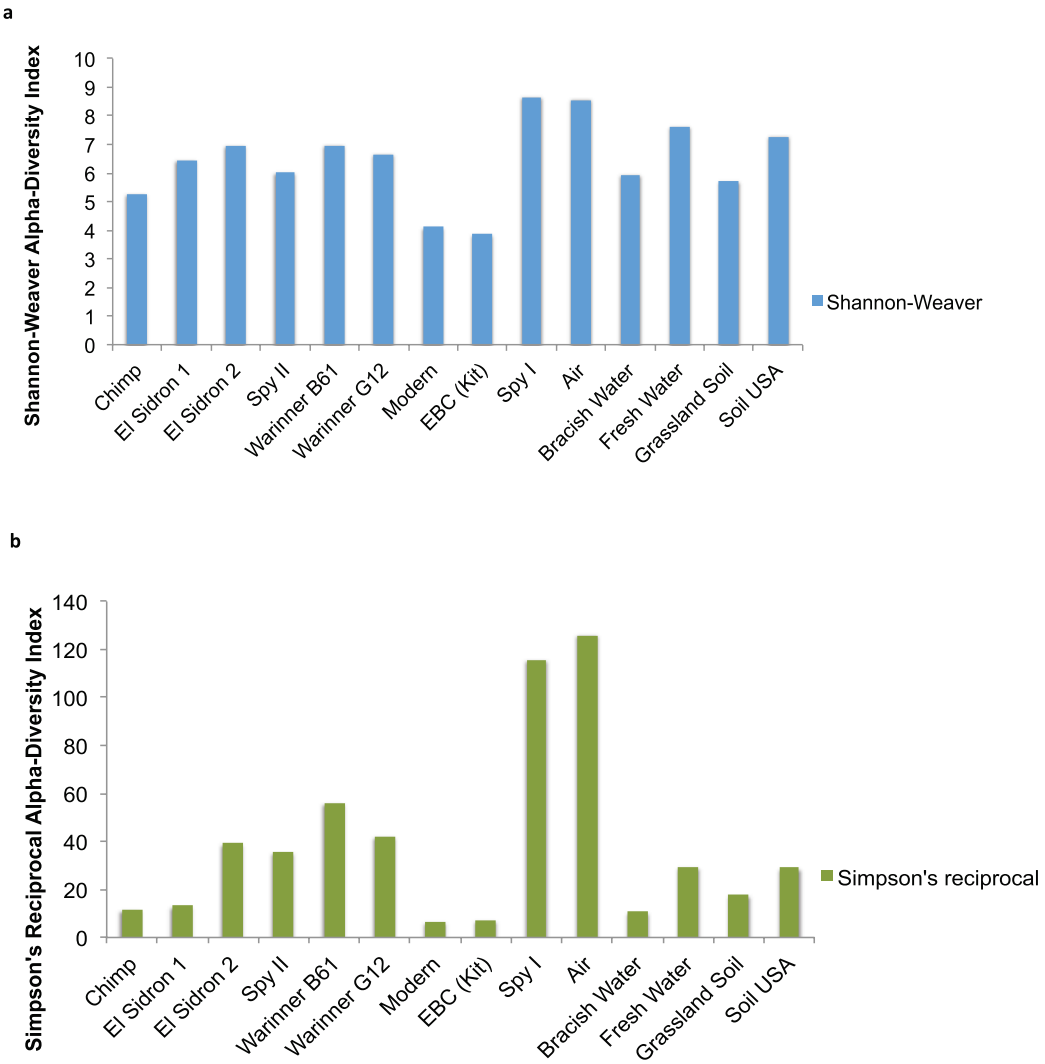
Extended Data Figure 5 | MALTX benchmarked using modern oral microbiota and simulated datasets. a, Phyla identified in simulated metagenomes (modern or ancient) are shown for four different analysis programs: MALTX, DIAMOND, MetaPhylAn, and MG-RAST. **b**, Simulated metagenomes (modern (circle) or ancient (square; damaged)) analysed using four different software (DIAMOND (green), MALT (red), MetaPhylAn (blue), MG-RAST (orange)) were UPGMA-clustered according to Bray–Curtis distances calculated from genera within samples. **c**, Phyla identified by MALTX analysis in shotgun and amplicon oral datasets obtained from this study and MG-RAST are displayed in stacked bar plots.



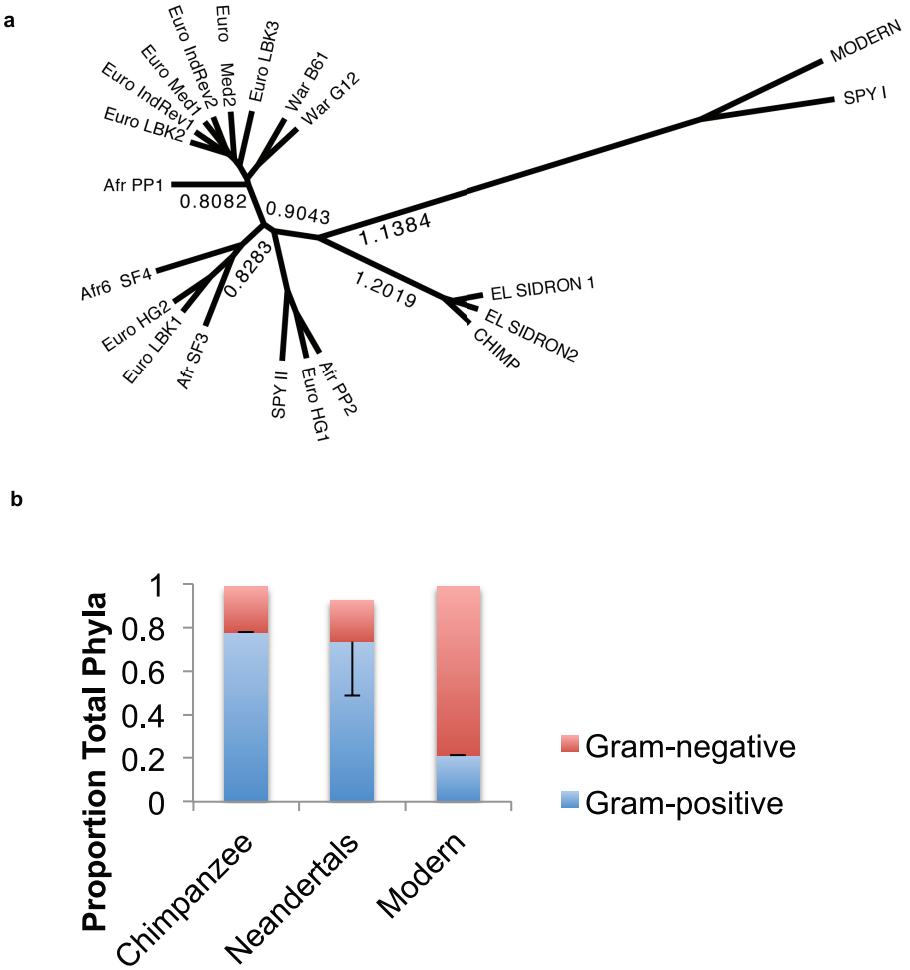
Extended Data Figure 6 | The composition of DNA sequences within ancient dental calculus in contrast to laboratory and environmental controls. a, Sequences identified by MALTX at the phyla level are displayed for dental calculus samples, extraction blank controls (EBCs), and environmental samples. Ancient dental calculus samples are ordered according to age, with the oldest specimens listed on the left. **b**, Identified reads from MALTX were filtered to remove reads corresponding to species identified in extraction blank controls from QG DNA extractions and environmental controls. **c**, Filtered data was summarized to analyse only archaea and bacterial phyla typically found in the modern oral cavity. Dental calculus samples are displayed in order of age.



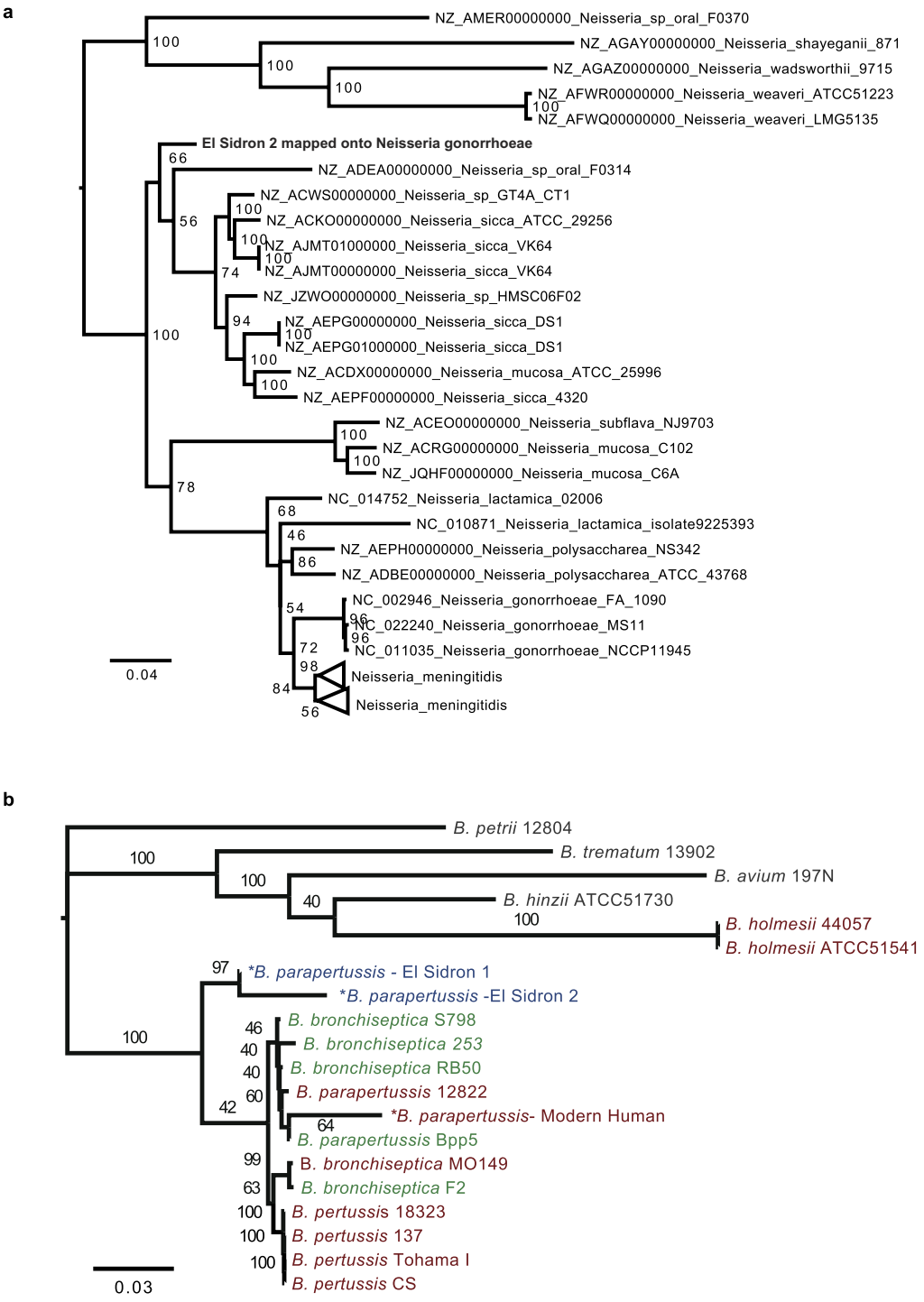
Extended Data Figure 7 | MapDamage analysis of oral bacterial species shared between Neanderthals and the modern human. a, **b**, The per cent of C–T mutations (**a**) and read length (**b**) calculated from mapped reads from each sample are shown for ten conserved species.



Extended Data Figure 8 | Alpha diversity from deeply sequenced unfiltered shotgun datasets. a, **b**, Calculations of rarefied data were carried out using Shannon–Weaver (**a**) and Simpson's reciprocal (**b**) indexes.



Extended Data Figure 9 | Neanderthal microbiota compared to other ancient and modern calculus specimens. a, UPGMA clustering of Bray–Curtis values were calculated from filtered rarefied shotgun data. **b**, The groups in **a** split on the basis of their differences in proportion of Gram-positive and Gram-negative phyla in shotgun datasets and were plotted for each group (chimpanzee and modern human, $n = 1$; Neanderthals, $n = 3$). Error bars represent s.d.



Extended Data Figure 10 | Phylogenetic analysis of unlikely bacterial pathogens observed in Neanderthal dental calculus. a, Reads from El Sidrón 2 were mapped onto shared *Neisseria* genes (that is, those gene regions shared between all of the species) and the resulting DNA fragments were aligned in MUGSY, compared to RAxML, and bootstrapped with 100 iterations. **b**, Phylogenetic analysis of whooping cough in Neanderthals was completed. Shared genomic regions within publically available *Bordetella* genomes were compared to ancient *Bordetella* reads from El Sidrón Neanderthals using RAxML with 1,000 iterations (bootstrap values).