

1 **Multi-proxy analyses of a mid-15<sup>th</sup> century ‘Middle Iron Age’ Bantu-speaker palaeo-faecal**  
2 **specimen elucidates the configuration of the ‘ancestral’ sub-Saharan African intestinal**  
3 **microbiome**

4  
5 Riaan F. Rifkin<sup>1,2,\*</sup>, Surendra Vikram<sup>1</sup>, Jean-Baptiste Ramond<sup>1,2,3</sup>, Alba Rey-Iglesia<sup>4</sup>, Tina B. Brand<sup>4</sup>,  
6 Guillaume Porraz<sup>5,6</sup>, Aurore Val<sup>6,7</sup>, Grant Hall<sup>8</sup>, Stephan Woodborne<sup>8,9</sup>, Matthieu Le Bailly<sup>10</sup>, Marnie  
7 Potgieter<sup>1</sup>, Simon J. Underdown<sup>1,2</sup>, Jessica E. Koopman<sup>1</sup>, Don A. Cowan<sup>1</sup>, Yves Van de Peer<sup>1,11,12</sup>, Eske  
8 Willerslev<sup>4,13,14</sup> & Anders J. Hansen<sup>4\*</sup>

9  
10 <sup>1</sup> Center for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology,  
11 University of Pretoria, Hatfield, South Africa

12 <sup>2</sup> Department of Anthropology and Geography, Human Origins and Palaeoenvironmental Research Group, Oxford  
13 Brookes University, Oxford, UK

14 <sup>3</sup> Department of Molecular Genetics and Microbiology, Pontificia Universidad Católica de Chile, Santiago, Chile

15 <sup>4</sup> Centre for GeoGenetics, GLOBE Institute, University of Copenhagen, Denmark

16 <sup>5</sup> CNRS, UMR 7041 ArScAn-AnTET, Université Paris-Nanterre, Paris, France

17 <sup>6</sup> Evolutionary Studies Institute, University of the Witwatersrand, Johannesburg, South Africa

18 <sup>7</sup> Department of Early Prehistory and Quaternary Ecology, University of Tübingen, Tübingen, Germany

19 <sup>8</sup> Mammal Research Institute, University of Pretoria, South Africa

20 <sup>9</sup> iThemba LABS, Johannesburg, South Africa

21 <sup>10</sup> University of Bourgogne France-Comte, CNRS UMR 6249 Chrono-environment, Besancon, France

22 <sup>11</sup> VIB Centre for Plant Systems Biology, Ghent, Belgium

23 <sup>12</sup> Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium

24 <sup>13</sup> GeoGenetics Group, Department of Zoology, University of Cambridge, Cambridge, UK

25 <sup>14</sup> Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK

26  
27 \* Corresponding authors: [ajhansen@snm.ku.dk](mailto:ajhansen@snm.ku.dk), [riaanrifkin@gmail.com](mailto:riaanrifkin@gmail.com)

28  
29 **ABSTRACT**

30 **The archaeological incidence of ancient human faecal material provides a rare opportunity to**  
31 **explore the taxonomic composition and metabolic capacity of the ancestral human intestinal**  
32 **microbiome (IM). Following the recovery of a single desiccated palaeo-faecal specimen from**  
33 **Bushman Rock Shelter in Limpopo Province, South Africa, we applied a multi-proxy analytical**  
34 **protocol to the sample. Our results indicate that the distal IM of the Neolithic ‘Middle Iron Age’ (c.**  
35 **AD 1485) Bantu-speaking individual exhibits features indicative of a largely mixed forager-agro-**  
36 **pastoralist diet. Subsequent comparison with the IMs of the Tyrolean Iceman (Ötzi) and**  
37 **contemporary Hadza hunter-gatherers, Malawian agro-pastoralists and Italians, reveals that this**  
38 **IM precedes recent adaptation to ‘Western’ diets, including the consumption of coffee, tea,**  
39 **chocolate, citrus and soy, and the use of antibiotics, analgesics and also exposure to various toxic**  
40 **environmental pollutants. Our analyses reveal some of the causes and means by which current**  
41 **human IMs are likely to have responded to recent dietary changes, prescription medications and**  
42 **environmental pollutants, providing rare insight into human IM evolution following the advent of**  
43 **the Neolithic c. 12,000 years ago.**

44  
45 **INTRODUCTION**

46 The human gastrointestinal tract (GI) harbours a dynamic population of bacteria, archaea, fungi, protozoa  
47 and viruses; the intestinal microbiota. This collection of microorganisms, comprising the human  
48 intestinal microbiome (IM) (*I*) performs critical functions in digestion, development, behaviour and  
49 immunity (2, 3). Modifications of the core IM composition (dysbiosis) have been associated with the

50 pathogenesis of inflammatory diseases and infections (3, 4), including autoimmune and allergic diseases,  
51 obesity, inflammatory bowel disease and diabetes (5). Despite its clinical importance, the factors that  
52 contribute to changes in IM taxonomic composition and functionality are not entirely understood (6, 7).  
53 This is attributed to the fact that most of what is known about the human IM is based on contemporary  
54 industrialised and ‘traditional’ human societies (8-10). In evolutionary terms, our species have subsisted  
55 by hunting and gathering for >90% of our existence (11). Evidence derived from the analyses of the IMs  
56 of traditional societies, including the Tanzanian Hadza hunter-gatherers (8), the Venezuelan Yanomami  
57 Amerindians (5), the BaAka Pygmies in the Central African Republic (12) and the Arctic Inuit (13) are  
58 thus widely viewed as representing ‘snapshots’ of ancient human IM composition. However, as exposure  
59 to Western diets, medicines and microbes cannot be excluded, one must be cautious about the use of  
60 these ethnographic cohorts as proxies for prehistoric human IMs (14).

61 The transformation of the IMs of present-day humans to their current ‘modernised’ state  
62 commenced millennia ago, with the advent of the Neolithic, which, at c. 12,000 years ago (ya), resulted  
63 in the first major human dietary transition (15). But precisely how our IMs changed following the advent  
64 of the Neolithic, and the Industrial Revolution after c. AD 1760, remains ambiguous (16-18). In this  
65 regard, the analyses of ancient human IMs provide a unique view into the co-evolution of microbes and  
66 human hosts, host microbial ecology and changing human IM-related health states through time (2).  
67 Indeed, over the past 15 million years, multiple lineages of intestinal bacterial taxa arose via co-  
68 speciation with African hominins and non-human primates, *i.e.*, chimpanzees, bonobos and gorillas (19).  
69 The departure of behaviourally ‘fully-modern’ *Homo sapiens* from Africa c. 75,000 years ago (kya)  
70 resulted in the global dispersal of our species (20). Significantly, various microbes accompanied these  
71 human dispersals ‘out of Africa’ (21, 22). Since the ancestral human IM is estimated to comprise a  
72 taxonomically and metabolically more diverse array of microbes than those found in contemporary  
73 societies (6, 10), the IMs of pre-Clovis North Americas (23), pre-Columbian Puerto Rican Amerindians  
74 (24) and pre-Columbian Andeans (25) represent more accurate indications of ancient human IM  
75 composition. These studies have provided significant insight into the structure, function and evolution of  
76 the human IM, highlighting the influence of dietary changes on the intestinal microbial ecology of  
77 contemporary humans (2, 6, 7). These have also provided essential baseline information for  
78 understanding the evolutionary processes implicated in the taxonomic configuration and metabolic  
79 capacity of both healthy and dysbiotic human IMs.

80 Despite the fact that African populations are not underrepresented in studies concerning  
81 ‘traditional’ human IMs (8, 12, 26), there is, currently, no information concerning the taxonomic  
82 diversity and metabolic capacity of ancestral (*i.e.*, archaeologically-derived) African IMs. To gain insight  
83 into the ancient African human IM, the prehistoric incidence of intestinal parasites, pathogenic microbes  
84 and antibiotic resistance genes, we performed shotgun metagenomic sequencing of a prehistoric (pre-  
85 colonial) faecal specimen recovered from a Middle Iron Age (c. AD 1485) context at Bushman Rock  
86 Shelter (BRS) in Limpopo Province, South Africa. Comparison with ancient (Ötzi), traditional (Hadza  
87 and Malawian) and contemporary ‘Western’ (Italian) IM datasets indicate that the IM of the BRS  
88 individual represents a unique taxonomic and metabolic configuration observed in neither contemporary  
89 African, nor European, populations (see Methods).

90

## 91 **RESULTS AND DISCUSSION**

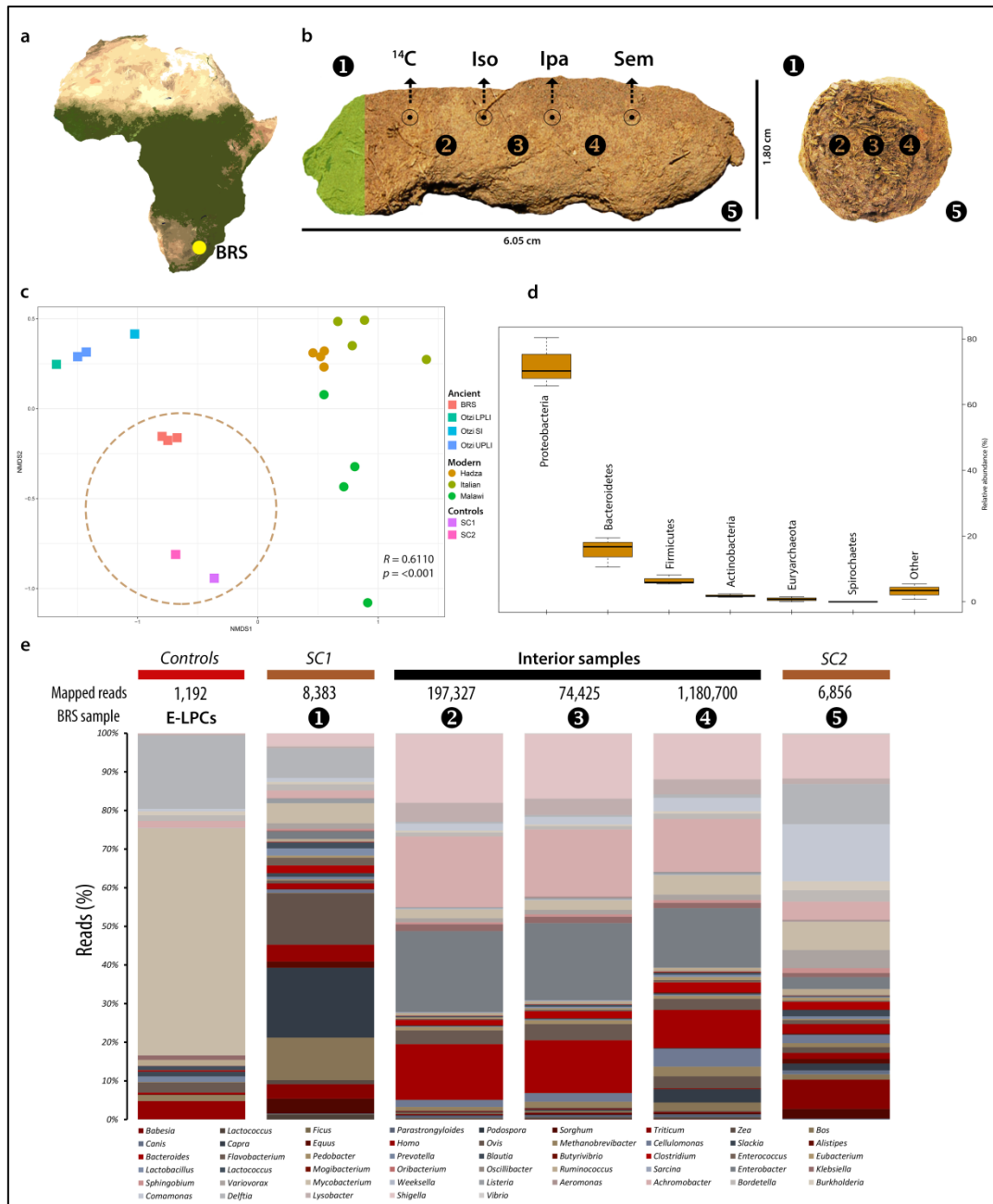
### 92 **Specimen provenience and ancient subsistence reconstruction**

93 The palaeo-faecal specimen was recovered *in situ* from an exposed stratigraphic section at BRS (27)  
94 (Fig. 1a). This large dolomitic rock-shelter is situated on the edge of the Great Escarpment in the  
95 Drakensberg chain. The occupation level from which the specimen derives comprises the uppermost  
96 archaeological unit of the rock-shelter designated ‘Angel’ in ‘Block A’. Layer 1 (‘Angel’) relates to the  
97 arrival of Bantu-speaking Iron Age agro-pastoralists in the region after c. 1,800 ya (Fig. S1). This

98 occupation therefore reflects the advent of the Neolithic in South Africa, which entailed the introduction  
99 of domesticated taxa such as sorghum (*Sorghum bicolor*), cattle (*Bos taurus*) and various other Iron Age-  
100 related species and cultural practices (*e.g.*, ceramic and iron-smelting technologies) into the region (28).  
101 All the preceding archaeological layers at BRS are representative of occupations by Holocene (*e.g.*, the  
102 Oakhurst techno-complex at ~10 kya), Terminal Pleistocene (the Robberg techno-complex at ~20 kya)  
103 and Pleistocene (*i.e.*, the Pietersburg techno-complex ~80 kya) hunter-gatherers (27, 29).

104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145

146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181



182 **Fig. 1. Provenience of, sub-sampling protocol applied to and microbial taxa detected in the BRS palaeofaecal specimen.** **A)** The location of Bushman Rock Shelter (BRS) in Limpopo Province, South Africa, **B)** Lateral  
183 (left) and cross-sectional (right) views of the specimen indicating the sub-sampling protocol applied to facilitate  
184 DNA extraction, including ‘sedimentary control’ sample 1 (‘SC1’), faecal samples 2, 3 and 4 and sedimentary  
185 control sample 5 (SC2),  $^{14}\text{C}$  AMS dating ( $^{14}\text{C}$ ), isotope analyses (Iso), intestinal parasitic analyses (Ipa), scanning  
186 electron microscopy (Sem) and the preservation of a voucher sample (indicated in green shading), **C)** Non-metric  
187 multi-dimensional scaling (NMDS) plot comparing the taxonomic community structure (by weighted Bray-Curtis  
188 dissimilarity analysis) of the BRS specimen (*i.e.*, BRS2, BRS3 and BRS4) and the sediment controls (SC1 and  
189 SC2) with the ancient (Ötzi) (indicated as SI ‘small intestine’, LPLI ‘lower part of the lower intestine’ and UPLI  
190 ‘upper part of the lower intestine’), traditional (Hadza and Malawian) and modern (Italian) IM datasets (taxa were  
191 filtered to the occurrence of  $>3$  in at least 20% of the samples resulting in the inclusion of 371 taxa) ( $R = 0.6110$   
192 indicates ANOSIM analysis which revealed significant differences ( $p < 0.001$ ) between the ancient and modern IM  
193 samples), **D)** Box-and-whisker plot indicating the relative abundance of intestinal bacterial phyla detected in the  
194 BRS specimen (*i.e.*, BRS2, BRS3 and BRS4) (‘other’ comprises phyla with  $<0.6\%$  relative abundance), **E)** Bar-  
195 chart providing an overview of all environmental, commensal and pathogenic genera identified in the BRS  
196

197 specimen (BRS2, BRS3 and BRS4) and information concerning the DNA extraction and library preparation  
198 negative controls (E-LPCs) and modern and ancient sedimentary controls (SC1 and SC2) (data derived from Table  
199 1, Table 2 and Table S1) (see Methods).

200

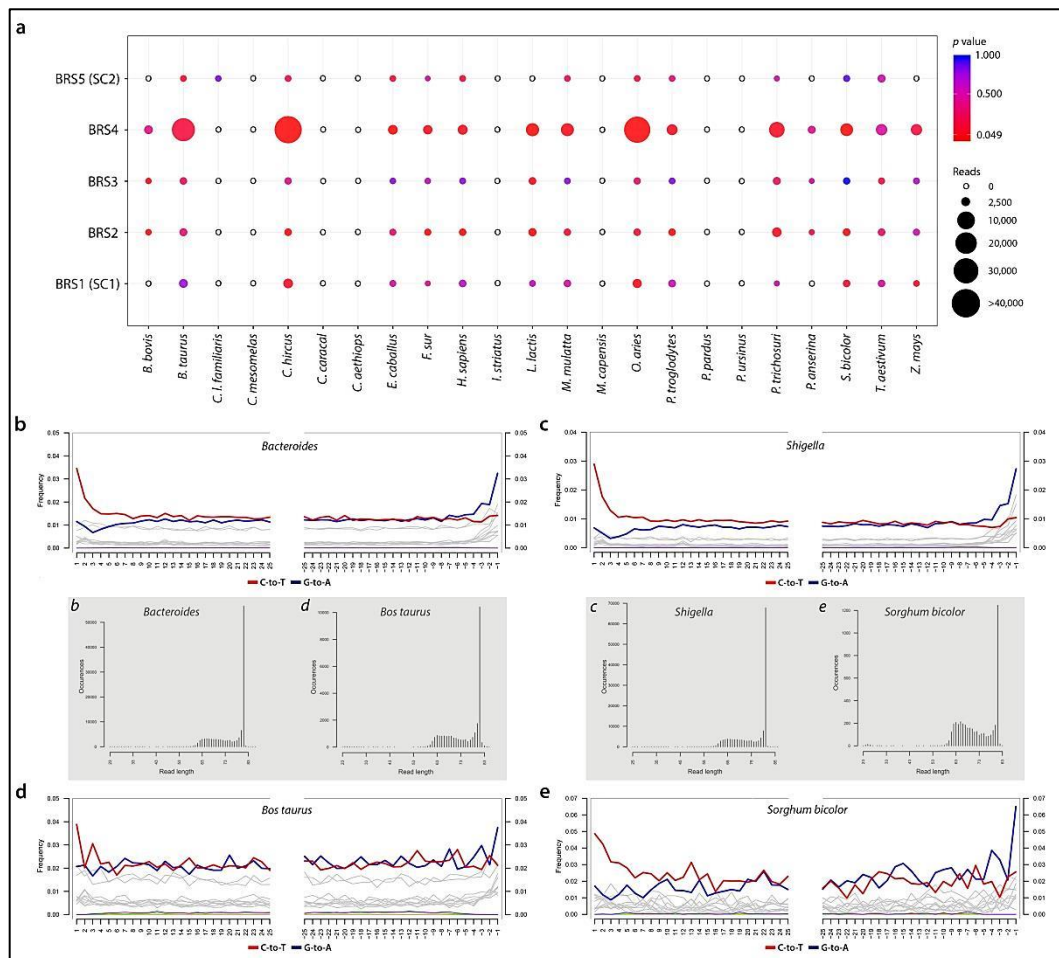
201 Following recovery of the specimen using latex gloves and stainless-steel forceps (whilst  
202 wearing a biologically-impervious body-suit and surgical face-mask) it was sealed in a sterile plastic  
203 ziplock bag and stored at ~4°C. Sediment control sample BRS1 ('SC1') was collected from the surface  
204 of the rock-shelter (~25 cm above the specimen) and BRS5 ('SC2') from the levels preceding the Iron  
205 Age (~25 cm below the specimen) in the Oakhurst occupation dated to *c.* 10 kya (27, 29). These were  
206 used as 'controls' to assess ecological and faecal community composition for biological plausibility and  
207 also the likelihood of sedimentary aDNA (sedaDNA) leaching. Sub-sampling was performed in ancient  
208 DNA (aDNA) laboratories at the Centre for GeoGenetics, University of Copenhagen (Denmark),  
209 applying established protocols (30) (Fig. S2). Prior to sub-sampling, the outer surface or cortex (~5mm)  
210 of the specimen was removed with a scalpel and excluded from further analyses, primarily as it was in  
211 contact with surrounding sediment and could therefore have been contaminated by soil-derived microbes  
212 (Fig. S2). To address within-sample variability, three faecal sub-samples (*i.e.*, BRS2, BRS3 and BRS4)  
213 were taken from different sites within the specimen. From the remaining one-third of the specimen, four  
214 sub-samples were taken for radiocarbon (<sup>14</sup>C) dating, isotopic analysis, and microscopic intestinal  
215 parasitic and scanning electron microscopy (SEM) analyses. One-sixth of the specimen was preserved (at  
216 -20°C) as a voucher sample (Fig. 1b).

217 To ascertain whether the palaeo-faecal specimen derives from a human individual, all other  
218 potential source species were eliminated. Given the limited number (1,967) of aDNA sequence reads  
219 mapped to *H. sapiens*, likely due to the removal of the exterior cortex prior to sub-sampling (in which  
220 most human-derived (nuclear and mitochondrial) DNA would be expected to reside), metagenome  
221 assembly could not be performed. Morphologically, the specimen resembles several candidate species  
222 (31), although no DNA sequence reads for indigenous felids (*e.g.*, leopard (*Panthera pardus*), caracal  
223 (*Caracal caracal*) *etc.*), mustelids (honey badger (*Mellivora capensis*), polecat (*Ictonyx striatus*) *etc.*),  
224 jackal (*Canis mesomelas*) or domestic dogs (*C. lupus familiaris*), and none for the indigenous primates,  
225 *i.e.*, vervet monkeys (*Cercopithecus aethiops*) or baboon (*Papio ursinus*), were detected. Reads related to  
226 non-human primates, *i.e.*, *Pan troglodytes* (the common chimpanzee) and *Macaca mulatta* (rhesus  
227 macaque) are likely the result of false-positive identifications, as these taxa do not currently occur in the  
228 region, nor would they have in the past (32). The incidence of statistically-significant (*i.e.*, verified  
229 ancient) C-T *p*-values for the 1,967 reads mapped to *H. sapiens* supports the conclusion that the faecal  
230 specimen derives from a human individual (Table S1).

231 To confirm the association of the faecal specimen with the archaeological context from which it  
232 was recovered, two direct radiocarbon (<sup>14</sup>C) Accelerator Mass Spectrometry (AMS) dates were generated  
233 from two sub-samples taken from within the specimen (Fig. 1b) (see Methods). The dates of 470 ± 44  
234 years before present (BP) (IT-C-1020) and 460 ± 35 years BP (IT-C-1077) indicate that the sample was  
235 deposited *c.* AD 1485 (Table S2). This date falls within the South African Middle Iron Age (spanning  
236 AD 1300-1840), follows the demise of the nearby Kingdom of Zimbabwe at *c.* AD 1450 (33) and closely  
237 precedes first contact with European seafarers in AD 1488 (34).

238 Prior to the identification of environmental and subsistence-related taxa, all exotic taxa,  
239 including kiwi (*Apteryx*), carp (*Cyprinus*), salmon (*Oncorhynchus*), pig (*Sus*), chicken (*Gallus*) and rice  
240 (*Oryza*) were identified and excluded from further analyses. The evaluation of taxa present in the DNA  
241 extraction (*n* = 1) and library preparation (*n* = 1) negative controls (E-LPCs) indicated that instances of  
242 environmental contamination were restricted largely to taxa widely cited as either 'contaminants' or as  
243 derived from false-positive identifications (35, 36). The authenticity of microbial and macrobial  
244 sequence-derived taxa was determined by statistical aDNA sequence damage estimation (37),

245 comparison to E-LPCs, DNA read-length characteristics and ecological conformity. Using high-quality  
 246 filtered reads for DNA damage estimation analyses with PMDtools (37), this process facilitated the  
 247 validation of forty-seven taxa represented by 1,470,662 reads as ancient (Fig. 2, Table 1, Table 2) (Table  
 248 S1) (see Methods). Subject to the availability of sufficient numbers of high-quality ‘mappable’ aDNA  
 249 sequence reads, we employed mapDamage (38, 39) to validate the authenticity of taxa by determining  
 250 the incidence of C-T and G-A substitution rates at the 5’-ends and 3’-ends of strands (see Methods).  
 251 DNA damage analyses of sequence reads derived from the genera *Bacteroides* and *Shigella* (Fig. 2b, c)  
 252 and also *B. taurus* and *S. bicolor* (Fig. 2d, e) indicate that the nucleotide composition at the ends of the  
 253 analysed reads exhibits the typical pattern expected for ancient DNA (38, 39).



283 **Fig. 2. DNA damage estimation analyses and authentication of environmental and subsistence-related taxa**  
 284 **detected in the BRS palaeo-faecal specimen. A)** Dot-plot indicating the occurrence of statistically-significant C-T  
 285 *p*-values calculated for environmental- and subsistence-related taxa detected in BRS (1 (SC1), BRS2, BRS3, BRS4  
 286 and BRS5 (SC2)) (circle sizes and colours represent mapped read-counts and *p*-value significance) and ancient  
 287 DNA fragmentation patterns shown within the first 25 bp from read ends for the genera **B) *Bacteroides*** and **C)**  
 288 *Shigella* and the species **D) *Bos taurus*** and **E) *Sorghum bicolor*** (fragment size distributions for each taxon is  
 289 indicated in the grey inset and labelled *b, c, d* and *e*) (Table S1) (Fig. S3).

290  
 291 In dietary terms, the Bantu-speaker agro-pastoralist diet comprised not only domesticated animal  
 292 and plant taxa, but also various hunted and gathered indigenous species, including antelope, fish, plants  
 293 and fruits (40). The presence of subsistence-related reads derived from sorghum (*S. bicolor*), cluster figs  
 294 (*Ficus sur*), goat (*Capra hircus*), sheep (*Ovis aries*) and beef (*B. taurus*) are indicative of taxa that were  
 295 consumed shortly before (*i.e.*, 24 to 36 hours) stool deposition by the BRS individual. Based on bulk

296 untreated  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  values obtained from isotopic analyses, the individual had a predominantly C4-  
297 based meal with a minor C3-based contribution (see Methods). This concurs with the aDNA evidence  
298 indicating the presence of sorghum, wild figs and beef. Sorghum is a C4 plant with published  $\delta^{13}\text{C}$  values  
299 of -14.0 to -11.0‰, and cattle are grazers (*i.e.*, C4 consumers). The  $\delta^{13}\text{C}$  values (-16.79‰) are higher  
300 than -18‰, which is considered the threshold for a predominantly terrestrial diet. These do not however  
301 preclude the occasional consumption of freshwater resources, including fish, given the close proximity  
302 (~1.5 km) of the shelter to the perennial Ohrigstad River (27) (Table S3) (Fig. S4). The incidence of  
303 cattle (*B. taurus*), cattle-specific microbes, *i.e.*, *Lactococcus lactis* (a component in fermented milks) and  
304 *Babesia bovis* (the causative agent of babesiosis) are also representative of a Bantu-speaker pastoralist  
305 subsistence economy. The non-authenticated incidence of *Podospora anserina* (a dung-colonising  
306 fungus) is interpreted as symptomatic of post-depositional saprophytic processes.

307

### 308 **Identifying commensal intestinal microbiota**

309 It is estimated that the human IM harbours ~150 to ~400 IM species (41), most of which belong to the  
310 phyla *Actinobacteria*, *Bacteroidetes*, *Firmicutes* and *Proteobacteria* (6). The variability of microbial  
311 taxonomic abundance, however, influences the identification of the common core IM (42). As many  
312 microbes are capable of transient integration into the IM, where they influence the composition and  
313 metabolic activity of resident IM communities (43), essentially environmentally-derived genera, *i.e.*,  
314 *Bacillus*, *Dietzia*, *Microbacterium*, *Paracoccus*, *Pseudomonas*, *Staphylococcus* and *Streptomyces* were  
315 omitted from further analyses.

316 On the basis of taxa detectable at sequencing depth, metagenomic comparison of the shotgun  
317 reads with the National Center for Biotechnology Information (NCBI) BLASTn non-redundant  
318 nucleotide (*nt*) database using MEGAN Community Edition (CE) v6.10.10 and the Burrows-Wheeler  
319 Aligner (BWA) facilitated the identification of 691,303 reads (1.48% of all sequence reads) representing  
320 thirty-six ancient commensal IM genera. Subsequent statistical DNA damage estimation resulted in the  
321 elimination of 12 non-authenticated bacterial genera (*i.e.*, 29,012 reads) from the dataset, including  
322 *Bifidobacterium*, *Coprococcus*, *Dorea*, *Faecalibacterium*, *Mollicutes*, *Neisseria*, *Parabacteroides*,  
323 *Phascolarctobacterium*, *Romboutsia*, *Roseburia*, *Ruminiclostridium*, *Tissierellia* and *Treponema* (see  
324 Methods) (Table S4). Based on the remaining authenticated 693,082 sequence reads exhibiting an  
325 average read-length of 66.83 base-pairs (bp), twenty-four ancient IM taxa (Table 1) were identified. It is  
326 of interest to note that, whereas the BRS1 ‘surface control’ sample (SC1) yielded authenticated reads  
327 derived from ancient microbial IM taxa (*i.e.*, *Enterobacter*, *Enterococcus* and *Slackia*), the much older  
328 BRS5 control (SC2) (*i.e.*, the Oakhurst occupation dated to *c.* 10 kya), did not (Table 1, 2) (Table S1).  
329 The IM of the BRS individual (*i.e.*, including only the interior sub-samples BRS2, BRS3 and BRS4) was  
330 determined to be dominated by the two phyla *Proteobacteria* (41.73%) and *Bacteroidetes* (31.25%),  
331 followed by *Firmicutes* (13.44%), *Actinobacteria* (8.89%) and *Euryarchaeota* (4.68%) (Table 1) (Fig.  
332 1d). At genus level, the bulk (71.82%) of reads was ascribed to *Enterobacter* (34.45%), *Bacteroides*  
333 (22.36%), *Cellulomonas* (8.68%) and *Flavobacterium* (6.33%). In addition to *Clostridium* (4.93%) and  
334 *Methanobrevibacter* (4.68%), all other genera exhibit <5% relative abundance. We note that the use of  
335 ‘relative abundance’ as a measure of taxonomic representation has been a standard means by which  
336 differences in taxonomic composition or ‘abundance’ in IM datasets is analysed, verified and compared,  
337 and that various notable IM studies have adhered to the use of ‘relative abundance’ as standard analytical  
338 protocol (6, 8, 9, 12, 24, 25, 44). In addition, while cognisant of the compositional complexity of  
339 microbiome samples (44) and of the possible influence of the fragmented nature of ancient microbial  
340 DNA on taxonomic classification (45), we note that ancient DNA damage have been revealed to exert a  
341 minimal influence on species detection and on the ‘relative abundance’ of IM taxa in simulated ancient  
342 and modern datasets (46).

343 **Table 1. DNA sequence reads for twenty-four authenticated commensal IM taxa detected in the BRS palaeo-**  
344 **faecal specimen.** Statistically-significant (*i.e.*, verified ancient) C-T *p*-values are indicated in bold. BWA mapping  
345 was performed using high-quality filtered reads for DNA damage estimation analyses using PMDtools ('C-T *p*-  
346 values') (see Methods). Additional read-length information for individual taxa is provided in Table S5.

347

### 348 **Identification of ancient pathogenic microbial taxa**

349 There is an estimated 1,407 recognised species of human pathogens (47), many of which influence not  
350 only health and immune responses (48), but also cognitive development (49) and social behaviour (50).  
351 On the basis of taxa detectable at sequencing depth, metagenomic comparison of the shotgun reads with  
352 the NCBI BLASTn non-redundant nucleotide (*nt*) database using MEGAN and BWA facilitated the  
353 identification of 625,001 'mappable' reads (1.34% of all sequence reads) representing twelve ancient  
354 pathogenic taxa (Table 2). Only authenticated ancient taxa were retained (exhibiting an average read-  
355 length of 67.55 bp), the authenticity of which was determined by assessing the incidence of statistically-  
356 significant ( $p = <0.05$ ) C-T substitutions at the 5' ends of sequence reads (Table 2).

357 The occurrence of authenticated ancient reads homologous to *Listeria* and restricted to BRS1  
358 (SC1), suggests that this taxon, although ancient, is most likely environmental (*i.e.*, sedimentary) and that  
359 it does not derive from the faecal specimen. The incidence of authenticated DNA reads for  
360 *Mycobacterium* in BRS2, BRS3 and BRS4, and not in SC1 or SC2, is indicative of the general  
361 environmental (*i.e.*, sedimentary) presence of this genus. However, although not a known member of the  
362 intestinal microbiota, and given that not all species are pathogenic, the presence of authenticated  
363 'ancient' reads derived from *Mycobacterium* within the faecal specimen cannot be precluded as  
364 symptomatic of an infection. Some species, particularly *M. avium*, is known to invade intestinal  
365 epithelial cells and has been implicated in ulcerative colitis (51). Similarly, the incidence of authenticated  
366 reads for the genera *Comamonas*, *Lysobacter* and *Shigella* in BRS2, BRS3 and BRS4, *Aeromonas* in  
367 BRS2 and BRS4 and *Vibrio* in BRS4, confirms that these taxa derive from the faecal specimen and are,  
368 consequently, representative of an ancient gastro-intestinal infection. *Achromobacter* occurs in all sub-  
369 samples except for the ancient (*c.* 10 kya) control (BRS5/SC2) which did not yield any authenticated  
370 microbial pathogenic taxa. The notable abundance of commensal (41.73 %) (Table 1) and pathogenic  
371 (87.67%) (Table 2) members of the *Proteobacteria* in the BRS IM are of interest as it has been  
372 established that an increase in *Proteobacteria* is indicative of IM dysbiosis and metabolic disease (52).  
373 Compared to primary human IM phyla, *i.e.*, *Actinobacteria*, *Bacteroidetes* and *Firmicutes*, the relative  
374 abundance of *Proteobacteria* in the IM is, however, highly variable. While an increase in the abundance  
375 of *Proteobacteria*, especially members of the *Enterobacteriaceae* (*i.e.*, *Klebsiella*, *Salmonella* and  
376 *Shigella*) (53) is a feature of bacterial dysbiosis, the human IM also contains members of commensal  
377 *Proteobacteria*, *i.e.*, *Enterobacter*, *Klebsiella*, *Sphingobium* and *Variovorax*. Under 'healthy' conditions,  
378 the relative abundance of *Proteobacteria* in the human IM can increase to ~45% without observable  
379 clinical implications (51).

380 Microscopic analysis aimed to determine the presence of intestinal parasites, namely helminths  
381 and protozoa, did not yield conclusive results (see Methods). Although this concurs with the aDNA  
382 results, the analyses of a single sub-sample might not have been sufficient to detect intestinal parasitic  
383 remains. Conversely, not all members of a population would necessarily be infected by intestinal  
384 parasites, possibly because of either natural resistance or limited exposure to contaminant sources.  
385 Similarly, while SEM analyses did not result in the detection of parasitic remains, it did facilitate the  
386 recognition of degraded plant fragments, and also of desiccated bacterial cells and saprophytic  
387 organisms, the latter of which likely represent both ancient and modern organisms, respectively (Fig.  
388 S5).

389

390



391 **Table 2. DNA sequence reads for twelve authenticated pathogenic taxa detected in the BRS palaeo-faecal**  
392 **specimen.** Statistically-significant (*i.e.*, verified ancient) C-T *p*-values are indicated in bold text. BWA mapping  
393 was performed using high-quality filtered reads for DNA damage estimation analyses using PMDtools ('C-T *p*-  
394 values') (see Methods). Additional read-length information for individual taxa is provided in Table S5.

395

### 396 **Ancient and modern IM taxonomic comparisons**

397 In terms of taxonomic composition, the ancient samples (BRS and Ötzi) exhibit spatial separation from  
398 the 'traditional' (Hadza, Malawian) and modern (Italian) comparative cohorts. Hierarchical clustering  
399 using complete-linkage based on Spearman's correlations, produced a clear separation of ancient and  
400 modern populations. ANOSIM analysis revealed significant differences between the ancient and modern  
401 IM samples ( $R = 0.9098$ ;  $p = <0.001$ ) for 371 taxa (Fig. S6) (see Methods). As stated, and bearing in mind  
402 the compositional complexity of IM samples (44) and the conceivable influence of fragmented DNA on  
403 taxonomic classification (45), ancient DNA damage result in minor differences in species detection and  
404 in comparisons concerning the 'relative abundance' of microbial taxa identified in ancient and modern  
405 IM datasets (46). As with weighted Bray-Curtis analysis based on the relative abundance of all identified  
406 IM taxa (Fig. 1c), we note that un-weighted Bray-Curtis analyses based on the 'presence-absence' of IM  
407 taxa exhibits correspondingly clear differences between the sedimentary controls (SC1 and SC2) and the  
408 ancient (*i.e.*, BRS and Ötzi), 'traditional' (Hadza, Malawian) and modern (Italian) comparative IM  
409 cohorts (ANOSIM  $R = 0.8361$ ;  $p = <0.001$ ) (Fig. S7). With regards potential contamination derived from  
410 the surrounding archaeological sedimentary matrix in the BRS palaeo-faecal specimen, comparison of  
411 the incidence of the 24 authenticated ancient IM taxa (Table 1) indicate that the surrounding sedimentary  
412 matrix (BRS1 'SC1' and BRS5 'SC2') and the DNA extraction and library preparation negative controls  
413 (E-LPCs) are not significant sources of microbial taxa identified in the palaeo-faecal specimen (BRS2,  
414 BRS3 and BRS4) (Fig. S8).

415 Metagenomic comparison of all analysed shotgun reads revealed that IM of the BRS individual  
416 (*i.e.*, BRS2, BRS3 and BRS4) is characterised by a *Firmicutes/Bacteroidetes* (F/B) ratio significantly  
417 skewed towards *Bacteroidetes* (at 31.25%), as opposed to *Firmicutes* (at 13.44%) (Fig. 1d) (Table 1).  
418 The F/B ratio (54) is widely considered as significant in human IM composition, with dysbiosis  
419 associated with inflammation, obesity, and metabolic diseases (55). Although this significance is  
420 controversial (56), we note that the BRS F/B ratio (0.4) does not resemble those reported for modern  
421 'traditional' Bantu-speaking Africans in Burkina Faso (2.8) (57) nor that calculated here for the East  
422 African Hadza (2.6) (8). This can likely be attributed to the fact that 'traditional' diets rich in starches  
423 (*e.g.*, potatoes, yams and sweet potatoes) have been shown to increase the F/B ratio, including increases  
424 in relative abundance of *Firmicutes* and enzymatic pathways and metabolites involved in lipid  
425 metabolism (58).

426 The IMs of modern humans have furthermore been stated to comprise one of three 'enterotypes',  
427 based on prevailing genera, *i.e.*, *Bacteroides*, *Prevotella* or *Ruminococcus* (59). Some taxa relate to long-  
428 term diets, such as *Bacteroides*, which is associated with the consumption of large amounts of protein  
429 and animal fat, and *Prevotella*, which is indicative of a high plant-derived carbohydrate intake (60).  
430 Similarly, *Ruminococcus* prevails in individuals who consume significant amounts of polyunsaturated  
431 fats, *e.g.*, marine fish, vegetable oils and nuts and seeds. The enterotypic composition of the BRS IM  
432 diverges from that reported for 'traditional' Africans (8, 12, 15, 26, 57, 61). In relation to *Ruminococcus*  
433 (1.57%) and *Prevotella* (0.63%), the BRS IM is characterised by a predominance of *Bacteroides*  
434 (22.36%) (*i.e.*, 'Enterotype 1') which concurs with a diet rich in protein and animal fat and which lends  
435 support our interpretation of the BRS *Bacteroidetes*-dominated F/B ratio. While this corresponds to data  
436 reported for the West African BaAka (12), it differs from the IM taxonomic composition reported for  
437 modern African cohorts, including the Tanzanian Hadza (8) and children in Burkina Faso (57), which  
438 exhibits higher abundance of *Prevotella* (Fig. 1e). The sizable incidence of *Flavobacterium* (6.33%) in

439 the BRS IM likely relates to the fact that members of this genus are resistant to dietary phenolic  
440 compounds derived from largely ‘medicinal’ plant taxa, including phenolic acids, flavonoids, tannins,  
441 curcuminoids, coumarins, lignans, quinones *etc.* (62). This genus also occurs in the IMs of non-human  
442 primates, including baboons (*P. ursinus*) and gorillas (*Gorilla gorilla*) (63). In relation to its substantial  
443 presence in the BRS IM, members of this genus might also have played a role in the elimination of  
444 aflatoxins present in milk, cheese, grains and figs. *Methanobrevibacter* (4.68%) is the most abundant  
445 archaeon in the human IM (64). Besides consuming fermentation products produced by saccharolytic  
446 bacteria, archaeal methanogenesis also improves the efficiency of polysaccharide fermentation.

447 The BRS IM furthermore exhibits enrichment towards *Cellulomonas* (8.68%) which degrades  
448 cellulose (65), *Clostridium* (4.93%) which is essential for IM resistance to infection and dysbiosis (66)  
449 and *Pedobacter* (1.75%) and *Prevotella* (0.63%) which, resembling non-pathogenic *Treponema*, are  
450 cellulose and xylan hydrolyzers (8). *Alistipes* (0.18%) is associated with a protein-rich diet and involved  
451 in amino acid fermentation (61) and *Butyrivibrio* (0.51%) ferments sugars, cellodextrins and cellulose  
452 (67). Since the antiquity of *Treponema* (*Spirochaetes*) in the BRS IM could not be verified, we cannot  
453 substantiate the premise that *Treponema* is inherently characteristic of all ‘traditional’ IMs (8, 9). Less  
454 abundant taxa, *i.e.*, *Ruminococcus* (1.57%), *Eubacterium* (1.46%) and *Enterococcus* (1.20%) are  
455 implicated in the digestion of starches and vitamin synthesis. *Sarcina* (0.15%) synthesizes microbial  
456 cellulose and occurs in high numbers amongst yam-farming African Pygmy hunter-gatherers and  
457 traditional populations in Papua New Guinea (68).

458

#### 459 **Ancient and modern IM metabolic comparisons**

460 Despite having highly divergent IM taxonomic compositions, functional gene profiles are relatively  
461 similar amongst different contemporary human populations. Accordingly, microbial community (*i.e.*,  
462 taxonomic) composition does not afford a thorough understanding of microbial IM community function  
463 (*i.e.*, metabolic capacity) (69). To ascertain statistically-significant differences between the metabolic  
464 capacities of ancient (pre-industrial) and modern ‘Westernised’ human IMs, we explored and compared  
465 the functional IM capacities of two ancient (BRS and Ötzi), two traditional (Hadza and Malawian) and  
466 one modern ‘Western’ (Italian) human cohorts (see Methods). This was performed only for the 24  
467 ancient authenticated IM taxa (Table 1). ANOSIM analysis revealed significant difference ( $R = 0.5395$ ;  $p$   
468  $= 0.001$ ) in the metabolic capacity of 24 authenticated ancient taxa for the ancient and modern cohorts.  
469 Spearman's correlation was performed on the taxa linked to the KEGG categories (*i.e.*, 1,487 KO gene  
470 categories linked to specific IM taxa). Although not contiguous to the faecal specimen, the considerable  
471 differences in the incidence of particular KEGG Orthology (KO) genes in BRS1 (SC1) and BRS5 (SC2)  
472 (the younger surface-derived and the ancient Oakhurst (*c.* 10 kya) sediment samples), and BRS2, BRS3  
473 and BRS4, eliminates the sedimentary matrix as a source for the greater proportion of KO genes  
474 identified in the faecal specimen (Fig. 3). The differential distribution of the 24 authenticated ancient IM  
475 taxonomic categories (Table 1), particularly in terms of the taxa detected in SC1 and SC2 *vs.* those  
476 detected in the BRS specimen (BRS2, BRS3 and BRS4) ( $R = 0.8361$ ;  $p = <0.001$ ), lends support to this  
477 conclusion (Fig. S7). In SC1 (BRS1), only three taxa (*i.e.*, *Enterobacter*, *Enterococcus* and *Slackia*)  
478 could be authenticated as ancient. No authenticated ancient taxa were recovered from SC2 (BRS5).

479

480

481

482

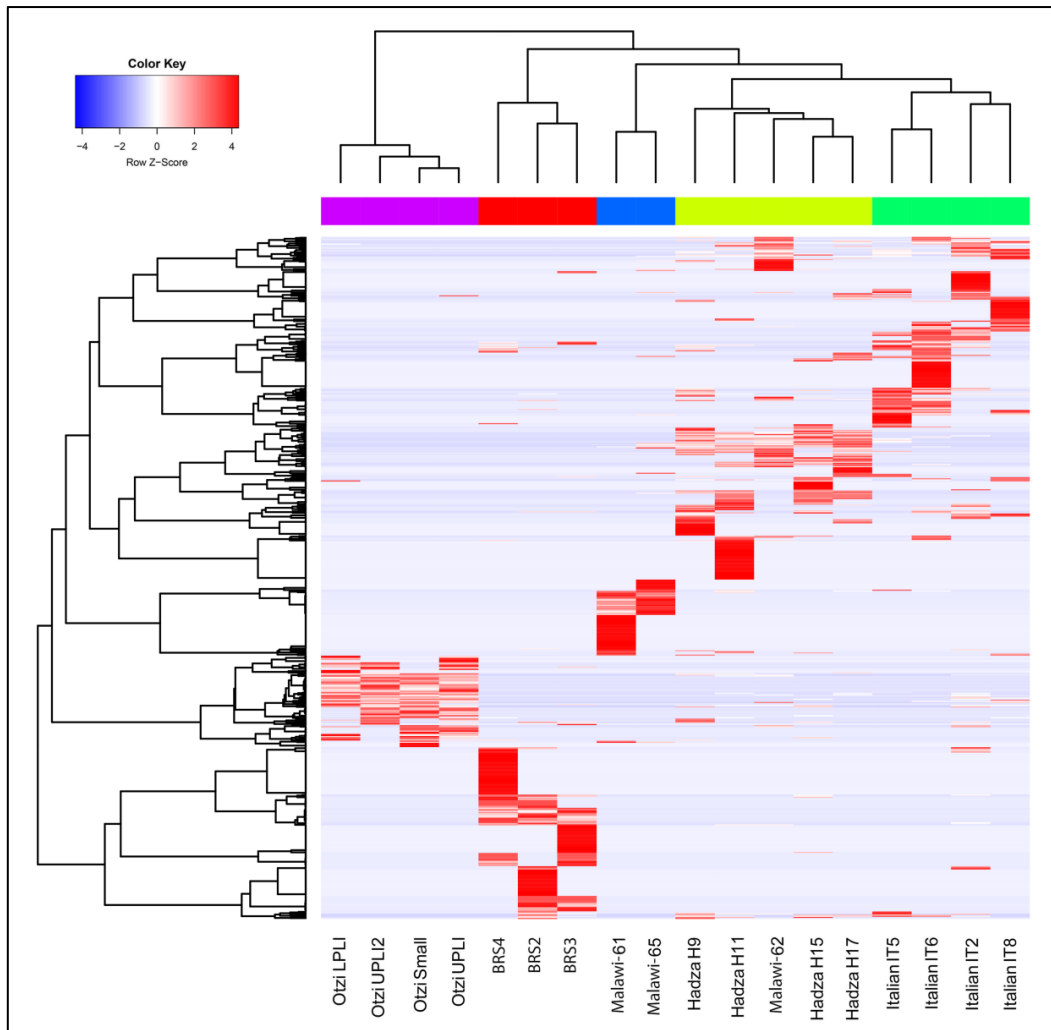
483

484

485

486

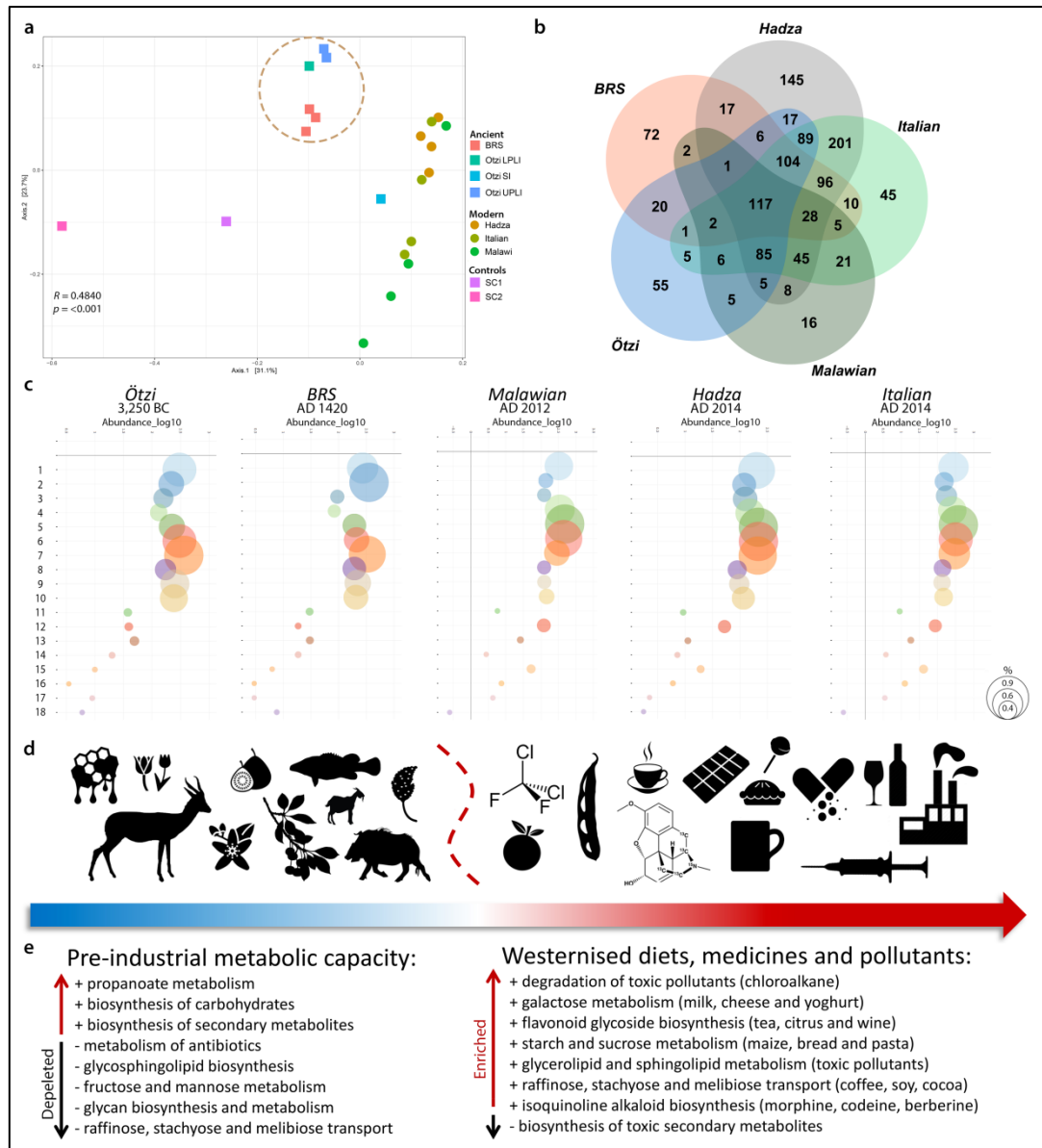
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537



**Fig. 3. Functional (metabolic) comparison of the ancient (BRS and Ötzi), ethnographic (Hadza and Malawian) and contemporary (Italian) faecal-derived human IMs based on KO-gene analyses for the twenty-four ancient authenticated IM taxa listed in Table 1 (Table S10).** In SC1 (BRS1), only three taxa (*i.e.*, *Enterobacter*, *Enterococcus* and *Slackia*) could be authenticated as ancient, and no authenticated ancient taxa were recovered from SC2 (BRS5). The heat-map based on Spearman's correlation coefficients comparing differences in functionality for the BRS IM (*i.e.*, BRS2, BRS3 and BRS4) with the ancient, traditional and contemporary IM datasets.

ANOSIM analysis revealed significant differences ( $R = 0.4840$ ;  $p = <0.001$ ) in the metabolic capacity of the ancient and modern comparative cohorts (Fig. 4a). Based on the analyses of the KO gene categories occurring only in the 24 authenticated ancient taxa (*i.e.*, 1,487) (Table 1), 72 taxa-specific KO genes are identified as unique to the BRS IM (Fig. 4b). Metagenomic comparison of the shotgun reads with the BLASTx NCBI non redundant protein (*nr*) database using DIAMOND v0.8.36.98 and MEGAN CE v6.10.10 revealed that 117 taxa-specific KO genes (7.86%) are shared between all (*i.e.*, ancient and modern) IM cohorts. While this is indicative of the relative temporal stability of a core commensal human IM community, it also reflects the variable (adaptable) and transient nature of human commensal IM community composition and metabolic capacity (3, 39, 70). This responsive adaptability is also echoed by the variable co-abundance of metabolic pathways identified in the BRS IM and the ancient and modern cohorts (Fig. 4c) (Table S7, S8, S9).

538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589



**Fig. 4. Graphic summary of dietary- and environmentally-induced differences in the metabolic capacities of the ancient and modern IM datasets analysed in this study.** **A)** Principal coordinates analysis (PCoA) comparison of the metabolic (functional) capacity of the BRS specimen (*i.e.*, BRS2, BRS3 and BRS4) and the sediment controls (SC1 and SC2) with the ancient (Ötzi) (SI ‘small intestine’, LPLI ‘lower part of the lower intestine’ and UPLI ‘upper part of the lower intestine’), traditional (Hadza and Malawian) and modern (Italian) IM datasets (KEGG categories were filtered for occurrence of >3 in at least 20% of the samples), **B)** Venn diagram indicating the relative abundance of IM taxa-linked KO genes identified in the ancient (*i.e.*, BRS2, BRS3, BRS4 and Ötzi), traditional (Hadza and Malawian) and modern (Italian) comparative cohorts, calculated as based on the 24 authenticated ancient IM taxa indicated in Table 1, **C)** Bubble-charts indicating the co-abundance ( $\log_{10}$ ) of eighteen (labelled ‘1’ to ‘18’) metabolic IM capacities for the ancient, traditional and modern IM cohorts (bubble sizes are representative of the relative abundance of KEGG categories (see scale on right) and comprise 1) glycolysis/gluconeogenesis, 2) citrate cycle, 3) fructose/mannose metabolism, 4) galactose metabolism, 5) starch/sucrose metabolism, 6) amino sugar and nucleotide sugar metabolism, 7) pyruvate metabolism, 8) glyoxylate/dicarboxylate metabolism, 9) propanoate metabolism, 10) butanoate metabolism, 11) synthesis and degradation of ketone bodies, 12) sphingolipid metabolism, 13) biosynthesis of unsaturated fatty acids, 14) n-glycan biosynthesis, 15) glycosphingolipid biosynthesis (-globo), 16) glycosphingolipid biosynthesis (-ganglio), 17) chloroalkane/ chloroalkene degradation and 18) naphthalene degradation) (Table S7), **D)** Dissimilarities in ancient

590 and modern IM metabolic capacities are related to recent (historical) changes in human dietary composition and  
591 exposure to toxic environmental pollutants (as indicated by the icons and the blue and red arrow) and **E**  
592 Differences in IM metabolic capacities are contrasted in terms of the up- and down-regulation of IM metabolic  
593 capacities as an ‘ancient’ vs. ‘modern’ comparative summary (see Methods).

594

595 In our comparison of the functional IM capacities of the two ancient (BRS and Ötzi), two  
596 traditional (Hadza and Malawian) and one modern ‘Western’ (Italian) human IM cohorts, we also  
597 focussed our analyses only on the 24 authenticated ancient IM taxa as indicated in Table 1 (Table S8, S9,  
598 S10). In relation to the modern (Italian, Hadza and Malawian) and ancient (Ötzi) datasets, the BRS IM  
599 (*i.e.*, BRS2, BRS3 and BRS4) appears to exhibit enrichment of KO genes implicated in the biosynthesis  
600 of secondary metabolites, including K00163 (Kruskal-Wallis value ( $H$ ) = 14.151 and  $p$ -value ( $p$ ) = 0.002),  
601 K00164 ( $H$  = 14.096,  $p$  = 0.002), K00163 ( $H$  = 15.812,  $p$  = 0.014), K00600 ( $H$  = 11.243,  $p$  = 0.004),  
602 K00568 ( $H$  = 11.706,  $p$  = 0.003) and K00457 ( $H$  = 14.762,  $p$  = 0.002) (Table S10). K00568 and K00457  
603 are also implicated in the biosynthesis of terpenoid-quinones. The capacity to biosynthesise toxic  
604 secondary metabolites (*e.g.*, polyketides, isoprenoids, aromatics (phenylpropanoids) or alkaloids) is  
605 essential when dietary sources comprise largely natural unprocessed foods. The BRS IM also exhibits  
606 enrichment of genes implicated in glyoxylate and dicarboxylate metabolism (K03781:  $H$  = 15.275,  $p$  =  
607 0.018 and K00600) and the citric acid cycle (CAC) (K00164:  $H$  = 15.650,  $q$  = 0.015 and K02274:  $H$  =  
608 15.928,  $p$  = 0.014). Glyoxylate and dicarboxylate glyoxylate are involved in the biosynthesis of  
609 carbohydrates, and CAC facilitates the release of energy from dietary carbohydrates, proteins and fats.  
610 Genes involved in glycolysis and gluconeogenesis (K00163:  $H$  = 15.812,  $p$  = 0.015) (pyruvate  
611 dehydrogenase E1 component) are also enriched.

612 The BRS IM exhibits depletion of KO genes involved in raffinose, stachyose and melibiose  
613 transport (*e.g.*, K10119:  $H$  = 15.934,  $p$  = 0.015, K10118:  $H$  = 15.640,  $p$  = 0.015 and K10117:  $H$  = 16.383,  $p$   
614 = 0.011) (Table S8). Soybeans are primary dietary sources of raffinose and stachyose, and melibiose  
615 occurs in coffee, cacao and processed soy (71). Deglycosylation by intestinal epithelial cell beta-  
616 glucosidases is a critical step in the metabolism of dietary flavonoid glycosides derived specifically from  
617 tea, citrus and wine (K05349:  $H$  = 15.068,  $p$  = 0.019). The BRS IM also contrasts with the modern cohort  
618 in terms of the depletion of genes involved in glycan (sugar-chain) biosynthesis and metabolism (alpha-  
619 mannosidase) (K01206:  $H$  = 14.443,  $p$  = 0.025) (beta-galactosidase) (K01190:  $H$  = 15.777,  $p$  = 0.015).  
620 While modern oligosaccharides are derived largely from processed ‘table sugar’ comprising mainly  
621 sucrose and fructose, foremost natural sources of sugar, comprising fruits and honey, would not have  
622 been consistently available for consumption. Correspondingly, KO genes involved in starch and sucrose  
623 metabolism (K00705:  $H$  = 15.318,  $q$  = 0.018 and K00975:  $H$  = 15.438,  $p$  = 0.017) (starch phosphorylase)  
624 (K00688:  $H$  = 13.496,  $p$  = 0.035) and fructose and mannose metabolism (K01193:  $H$  = 14.468,  $p$  = 0.025)  
625 (6-phosphofructokinase 1) (K00847:  $H$  = 16.009,  $p$  = 0.014) (fructokinase) are also depleted.

626 Whereas the Hadza IM is enriched in genes involved in fructose and mannose metabolism, the  
627 Italian IM is enriched in genes involved in the metabolism of simple sugars *e.g.*, glucose, galactose and  
628 sucrose (26). KO genes involved in the degradation of toxic pollutants (*i.e.*, chloroalkane, chloroalkene  
629 and naphthalene) and butanoate metabolism (K04072:  $H$  = 13.468,  $p$  = 0.036 and K00128:  $H$  = 15.970,  $p$  =  
630 0.013) are also depleted in the BRS IM. This is noteworthy, as there are no known natural sources of  
631 chlorinated paraffins (CPs) (72). CPs, including chloro-alkanes ( $C_{10-13}$ ), are widely used in the production  
632 of refrigerants, solvents, plasticisers and fire-retarding agents (73). Naphthalene ( $C_{10}H_8$ ), a polycyclic  
633 aromatic hydrocarbon, is derived from petroleum distillation and used in the manufacture of plastics,  
634 resins, fuels and insecticides. In addition, the depletion of KO genes involved in galactose, glycerolipid  
635 and sphingolipid metabolism and glycosphingolipid biosynthesis (K07407:  $H$  = 14.352,  $p$  = 0.026) (alpha-  
636 galactosidase) (K01190) (beta-galactosidase) is significant as it provides insight into the influence of  
637 exposure to modern environmental pollutants on IM composition and metabolic capacity.

638 Although it is challenging to infer ancient diet from ancient IM data, there is a growing  
639 understanding of the role of dietary choices on IM composition (2). The enrichment, in the BRS IM, of  
640 genes serving specific metabolic processes, including the biosynthesis of secondary metabolites,  
641 xenobiotic biodegradation, the metabolism of terpenoids and polyketides, propionate and butanoate and  
642 lysine degradation and the synthesis of ketone bodies is noteworthy as it suggest that the BRS metabolic  
643 profile is indicative of a diet rich in unprocessed natural resources, conceivably comprising medicinal  
644 plant substances (given the capacity for biosynthesising secondary metabolites and biodegrading  
645 xenobiotics), and encompassing irregular dietary intake. In addition, and given the enrichment of KO  
646 genes implicated in the synthesis and degradation of ketone bodies (*i.e.*, K00626), the metabolic profile  
647 of the BRS individual approximates that induced by a ketogenic diet (74), characterised by high-fat,  
648 adequate-protein and low-carbohydrate dietary consumption and accompanied by prolonged exercise and  
649 periods of low dietary intake or unintentional ‘fasting’ (Fig. 4d) (Table S9). Whether this metabolic  
650 profile resembles that generally referred to as a ‘palaeo-diet’ is unclear, as this would have entailed the  
651 exclusion of dairy, grains and legumes, nutritional categories which did indeed form part of the BRS diet.  
652 The BRS IM, and also that of Ötzi, are furthermore characterized by depletion of KO genes involved in  
653 the metabolism of antibiotics (*e.g.*, K11358:  $H = 15.320$ ,  $p = 0.017$ ), including aminocoumarin antibiotics  
654 and the metabolism of isoquinoline alkaloids, including the opiates morphine and codeine, as well as the  
655 antibiotic berberine.

656 In contrast to the ancient IMs, the modern (Hadza, Malawian and Italian) IM is characterized by  
657 enrichment of KO genes involved in raffinose, stachyose and melibiose transport (*e.g.*, K10119, K10118  
658 and K10117) indicative of a diet comprising soy, coffee, cacao and dietary flavonoid glycosides derived  
659 from tea, citrus and wine (K05349) (Table S10). This group also exhibits enrichment in genes concerning  
660 galactose metabolism, *i.e.*, K07407 ( $H = 14.352$ ,  $p = 0.0259$ ) (alpha-galactosidase), K00849 ( $H = 15.553$ ,  $p$   
661  $= 0.016$ ) (galactokinase) and K00965 ( $H = 15.694$ ,  $p = 0.015$ ) (UDPglucose--hexose-1-phosphate  
662 uridylyltransferase). Galactose is metabolized from milk sugar (lactose, a disaccharide glucose and  
663 galactose), the primary dietary source of which is milk and yogurt. KO genes involved in starch and  
664 sucrose (K00705) and amino sugar and nucleotide sugar metabolism (K00965 and K00849:  $H = 15.553$ ,  
665  $p = 0.016$ ) are also enriched, and so are genes involved in glycine, serine, threonine and methane and  
666 antibiotic metabolism. The enrichment of genes involved in glycerolipid and sphingolipid metabolism  
667 and glycosphingolipid biosynthesis (*e.g.*, K01190 and K07407) (alpha-galactosidase) likely reflects the  
668 impact of modern environmental pollutants on IM composition and metabolic capacity. It would be of  
669 interest to determine whether this does in fact represent a ‘population-wide’ functional response to  
670 exposure to toxic compounds ubiquitous in modern industrialized urban environments

671 In summary, significant differences between the ancient and modern IM metabolic capacity  
672 comprise the ability of the modern IMs to metabolise opiates (morphine, codeine) and antibiotics  
673 (berberine), raffinose, stachyose and melibiose indicative of a diet comprising soy, coffee, cacao, dietary  
674 flavonoid glycosides derived from tea, citrus and wine and glycerolipid and sphingolipid metabolism.  
675 Since these compounds were not present at the time of deposition of the BRS specimen, our results  
676 document the evolutionary influence of dietary changes, medicinal treatments and environmental  
677 pollutants on the IM taxonomic composition and metabolic capacity of contemporary human populations  
678 (Fig. 4e) (Table S8, S9).

### 679 680 **Antibiotic-resistance genes**

681 Our results furthermore confirm reports of antibiotic-resistance genes (ARGs) previously recovered from  
682 ethnographic cohorts and archaeological faecal samples (5, 25, 75). Following analysis of the ancient and  
683 modern resistomes using Resistome Analyser (<https://github.com/cdeanj/resistomeanalyzer>) (see  
684 Methods), we identified a total of 15 functional ARGs, four of which occurs in the BRS IM *i.e.*, BRS4.  
685 These include the prokaryotic protein synthesis elongation factor Tu (EF-Tu) (*tufA* and *tufB*), fluoro-

686 quinolone-resistant DNA topoisomerase (*parE*) and daptomycin-resistant *rpoB* (Table S11) (Fig. S9).  
687 Several bacterial taxa (*e.g.*, *Escherichia coli*, *Staphylococcus aureus* and *Streptomyces collinus*) have  
688 duplicate genes for the gene encoding EF-Tu (*tufA* and *tufB*) which confers resistance to the antibiotic  
689 kirromycin. These genes are also present in the Ötzi (sample UPLI), Hadza (H9 and H11) and Italian  
690 (IT6) datasets. The daptomycin-resistant *rpoB* gene encodes the  $\beta$  subunit bacterial RNA polymerase and  
691 is the site of mutations that confer resistance to the daptomycin antibacterial agents. Daptomycin-  
692 resistant *rpoB* is present in the Ötzi (SI, LPLI and UPLI) and Hadza (H11) datasets, but absent from the  
693 Malawian and Italian cohorts. Certain mutations in the RNA polymerase  $\beta$  subunit have been found to  
694 reduce the susceptibility of methicillin-resistant *S. aureus* (MRSA) for the antibiotics daptomycin and  
695 vancomycin (76). Several ARGs are limited in occurrence to the modern comparative cohorts. Some,  
696 including *gyrA* which confers fluoroquinolone resistance to *Neisseria gonorrhoeae* and *Ureaplasma*  
697 *urealyticum*, triclosan resistance to *Salmonella enterica*, and *pbp4B*, a penicillin-binding protein and the  
698 target of  $\beta$ -lactam antibiotics and *marA*, are shared only with the Hadza cohort. *Pbp2*, a point mutation in  
699 *N. meningitidis* which confers resistance to  $\beta$ -lactam, and a penicillin-binding protein found in  
700 *Streptococcus pneumoniae*, also occurs only in the Hadza dataset. *Cat*, which confers resistance to broad-  
701 spectrum phenicol antibiotics by antibiotic inactivation, occurs only in the Italian cohort. *Cat* has also  
702 been detected amongst isolated Amerindians (5).

703

## 704 CONCLUSIONS

705 In this study, we performed a comprehensive analysis of an ancient palaeo-faecal specimen derived from  
706 a 15<sup>th</sup> century Iron Age (Neolithic) South African Bantu-speaking hunter-agro-pastoralist. Although  
707 representative of the IM composition, metabolic capacity and ARG configuration of the distal (*i.e.*, the  
708 colon including the cecum, rectum and anal canal) IM of a single human individual, following particular  
709 dietary consumption and excreted at a single point in time, the characterisation of an authenticated  
710 ancient African Bantu-speaker IM is an important step towards understanding the ancestral (*i.e.*, pre-  
711 colonial African) state of the human IM. Our analyses designate a diet atypical of what is generally  
712 expected from a Neolithic (Iron Age) IM (17), instead comprising taxa indicative of a mixed forager-  
713 agro-pastoralist diet, supporting the role of dietary habits in shaping human IM composition.

714 It must be emphasised that the Neolithic of South Africa is different from the ‘classic’ Near-  
715 Eastern Neolithic, as foraging and hunting did play a prominent role in the subsistence configuration of  
716 southern African Iron Age communities (40). It must also be emphasised that the samples derived from  
717 the interior of the BRS specimen (*i.e.*, BRS2, BRS3 and BRS4) ‘clusters’ with the ancient comparative  
718 samples (*i.e.*, those derived from Ötzi) in terms of both taxonomic composition (ANOSIM analysis  
719 revealed significant differences between the ancient and modern IM samples ( $R = 0.8361$ ;  $p = <0.001$ ) for  
720 731 taxa (Fig. S7) and metabolic capacity (ANOSIM analysis revealed significant differences ( $R =$   
721  $0.4840$ ;  $p = <0.001$ ) (Fig. 4a), and not with the sedimentary controls (SC1 and SC2) or the modern  
722 comparative (*i.e.*, Hadza, Malawian and Italian) cohorts. In contrasting contemporary (Hadza, Malawian  
723 and Italian) and ancient (BRS and Ötzi) human IM taxonomic composition and metabolic capacity, it is  
724 evident that the changes brought about by modern human dietary composition, exposure to toxic  
725 pollutants and the excessive use of antibiotics, almost certainly resulted in positive selection for bacterial  
726 taxa involved in specific metabolic IM activities (69, 77). While this does not correlate directly with  
727 geography, it does exhibit a temporal trend towards the selection of KO genes in direct response to a  
728 number of specific changes in human dietary behaviour and environmental interaction and modification.

729 The IM of the BRS individual represents a unique taxonomic and metabolic configuration not  
730 observed in either contemporary African or European populations. Several studies have found that IM  
731 composition differs between Western urbanized and indigenous rural populations, and that these  
732 dissimilarities frequently correlate with dietary characteristics. In this instance, the diet of the BRS

733 individual, based on hunting, foraging and also agricultural and pastoral resources, differs from the  
734 typical Western diet comprising preservatives and food-enhancers, as well as coffee, chocolate, soy, wine  
735 and citrus. In terms of modern human hygiene-practices, it has been suggested that regular contact with  
736 ‘old friends’ (including both pathogenic and commensal environmental bacteria) is significantly  
737 diminished in Western countries (78) and that, given our extensive evolutionary history with microbes  
738 (19), this diminishes the capacity of the modern human IM to mediate allergic reactions and autoimmune  
739 and inflammatory diseases. It is evident that the ubiquitous use of antibiotics has altered the properties of  
740 formerly commensal bacteria and of the human IM (79). We therefore hypothesise that, by compelling  
741 commensal IM residents to respond to the introduction of antibiotics prescribed for pathogenic taxa,  
742 artificially-introduced dysbiosis has significantly modulated the pathogenic potential of commensal taxa,  
743 resulting in long-term deleterious impacts on optimal human IM functioning.

744 The IM of the BRS individual also provides evidence for recent human IM adaptation to  
745 environmental pollutants. The emergence of xenobiotic degradation pathways involved in naphthalene,  
746 chloroalkane and chloroalkene, benzoate and xylene degradation is likely a population-wide functional  
747 response of the IM to exposure to toxic and foreign compounds that are ubiquitous in industrialized  
748 urban environments. The respective enrichment and depletion of several KO genes implicated in the  
749 metabolism of morphine and codeine, as well as additives and supplements including pyruvate, l-arginine  
750 and beta-alanine, are also indicative of the adaptive capacity of the human IM. Given these modern  
751 influences, the contemporary human IM appears to be predisposed towards shifting to a state of  
752 dysbiosis. Such altered states of equilibrium frequently result in the pathogenesis of inflammatory  
753 diseases and infections, including autoimmune and allergic diseases, obesity and diabetes. The IM of the  
754 BRS individual also corroborates the premise that ARGs are a feature of the human IM, regardless of  
755 exposure to currently-available commercial antibiotics.

756 In conclusion, the large number of taxonomically (92.63%) and metabolically (88.51%)  
757 unassigned reads in the BRS palaeo-faecal specimen analysed here, granting that this might, to some  
758 extent, be a result of aDNA damage and the inability to ‘map’ all reads to existing comparative  
759 sequences (44, 45), is suggestive of substantial unknown IM taxonomic diversity and metabolic  
760 functionality (Table S12). In the future, the identification of these taxa and metabolic capacities might  
761 have significant implications for identifying health risks specific to the sub-Saharan African Bantu-  
762 speaker population which has increased in prevalence with the adoption of Western diets, medical  
763 treatments and exposure to modern pollutants. Given that sub-Saharan Africans living outside Africa  
764 exhibits a high prevalence of complex diseases, the comparison of ancient African IM data to those of  
765 modern Africans might facilitate not only retrospective disease diagnosis, but also the identification of  
766 IM-related risk factors that contribute to the onset of certain diseases.

767

## 768 **METHODS**

### 769 **Accelerator mass spectrometry (AMS) dating**

770 Two sub-samples derived from the interior of the specimen were subjected to accelerator mass  
771 spectrometry (AMS) dating. The samples were pre-treated using the standard acid-base-acid approach  
772 (80) performed at 70°C. Carbon was oxidised using off-line combustion in the presence of excess CuO  
773 and Ag, and the resulting CO<sub>2</sub> was reduced to graphite through Fe reduction at 600°C (81). The graphite  
774 was measured at the iThemba LABS AMS facility using Oxalic Acid II and Coal as the reference and  
775 background, respectively. We report <sup>14</sup>C ages in conventional radio-carbon years BP (*i.e.*, before present  
776 refers to AD 1950).

777

### 778 **Dietary isotope analyses**

779 To investigate the dietary composition of the BRS individual, one sub-sample derived from the interior  
780 of the specimen were subjected to isotopic analyses. The sample was homogenised using a mortar and



781 pestle and then divided in to three sub-samples. The first was left untreated and the second was subjected  
782 to a lipid extraction process using a 2:1 chloroform/ethanol solution to remove any lipids present (82).  
783 The sample was covered with 25 ml of the solution and the mixture agitated in an ultra-sonic bath for 15  
784 minutes and then left overnight. The solvent was then decanted and the sample dried at 70°C prior to  
785 weighing for analysis. The third sample was covered with 25 ml 1% hydrochloric acid (HCl) to remove  
786 inorganic carbonates (78), agitated for 15 min and left overnight. The acid was then decanted and the  
787 sample repeatedly washed (6 times) with distilled water before drying at 70°C. Aliquots of the samples  
788 weighing between 0.80 mg and 0.90 mg were weighed using a Mettler Toledo MX5 micro-balance. The  
789 weighed material was placed in tin capsules that had been pre-cleaned in toluene. All the samples were  
790 run in triplicate. Samples for isotopic analyses were combusted at 1020 °C using an elemental analyser  
791 (Flash EA 1112 Series) coupled to a Delta V Plus stable light isotope ratio mass spectrometer via a  
792 ConFlo IV system (Thermo Fischer, Bremen, Germany), housed at the Stable Isotope Laboratory,  
793 University of Pretoria. Two laboratory running standards (Merck Gel:  $\delta^{13}\text{C} = -20.26\text{‰}$ ,  $\delta^{15}\text{N} = 7.89\text{‰}$ ,  
794  $\text{C}\% = 41.28$ ,  $\text{N}\% = 15.29$  and DL-Valine:  $\delta^{13}\text{C} = -10.57\text{‰}$ ,  $\delta^{15}\text{N} = -6.15\text{‰}$ ,  $\text{C}\% = 55.50$ ,  $\text{N}\% = 11.86$ ) and a  
795 blank sample were run after every 11 unknown samples. Data corrections were performed using the  
796 values obtained for the Merck Gel during each run and the values for the DL-Valine standard provide the  
797  $\pm$  error/precision for each run. The precision for the BRS analyses was  $> 0.04\text{‰}$  and  $0.05\text{‰}$  for nitrogen  
798 and carbon respectively. These running standards are calibrated against international standards, *i.e.*,  
799 National Institute of Standards and Technology (NIST): NIST 1557b (bovine liver), NIST 2976 (muscle  
800 tissue) and NIST 1547 (peach leaves). All results are referenced to Vienna Pee-Dee Belemnite for carbon  
801 isotope values, and to air for nitrogen isotope values. Results are expressed in delta notation using a per  
802 mille scale using the standard equation ' $\delta\text{X}(\text{‰}) = [(R_{\text{sample}}/R_{\text{standard}}) - 1]$ ' where  $\text{X} = ^{15}\text{N}$  or  $^{13}\text{C}$  and  $\text{R}$   
803 represents  $^{15}\text{N}/^{14}\text{N}$  or  $^{13}\text{C}/^{12}\text{C}$  respectively.

804

### 805 **Intestinal parasite detection**

806 In addition to genomic taxonomic profiling, we also performed microscopic analysis to determine the  
807 incidence of intestinal parasitic helminths and protozoa. The extraction protocols applied in palaeo-  
808 parasitology used to extract parasitic markers (*i.e.*, eggs or oocysts) typically entails rehydration,  
809 homogenisation and micro-sieving (83). The sub-sample (~5 g) was placed in a rehydration solution  
810 comprising 50 ml 0.5% TSP solution and 50 ml 5% glycerinated water for seven days, after which it was  
811 ground and passed through an ultrasonic bath for 1 minute. The sample was then filtered in a sieving  
812 column comprising mesh sizes of 315  $\mu\text{m}$ , 160  $\mu\text{m}$ , 50  $\mu\text{m}$  and 25  $\mu\text{m}$  in aperture diameter. Because of  
813 the typical size of most intestinal parasite eggs range from 30  $\mu\text{m}$  to 160  $\mu\text{m}$  long and 15  $\mu\text{m}$  to 90  $\mu\text{m}$   
814 wide, only the two last sieves (*i.e.*, 50  $\mu\text{m}$  and 25  $\mu\text{m}$ ) were subjected to microscopic analyses.

815

### 816 **Scanning electron microscopy (SEM)**

817 We immobilised 0.5 g palaeo-faecal material on double-sided carbon tape (SPI supplies). Excess loose  
818 particles were blown off with compressed argon gas and coated with gold using an Emitech K450X  
819 sputter coater (Quorum Technologies, UK). SEM images were acquired on a Zeiss Ultra Plus Field  
820 Emission Scanning Electron microscopes (Carl Zeiss, Oberkochen, Germany), at an accelerating voltage  
821 of 1kV.

822

### 823 **Ancient DNA extraction and library preparation**

824 All pre-PCR amplification steps were carried out in aDNA laboratories at the Centre for GeoGenetics  
825 applying established aDNA protocols (84). Extractions for shotgun metagenome sequencing were carried  
826 out using a phenol-chloroform- and kit-based extraction protocol optimized for ancient sedimentary and  
827 faecal samples. This entailed dissolving a total of ~16 g of palaeo-faecal matter in 40 ml digestion buffer.  
828 Seven libraries, comprising BRS1 (SC1), BRS2, BRS3, BRS4, BRS5 (SC2) and two negative controls

829 (*i.e.*, an extraction and library preparation control referred to as ‘E-LPCs’) were constructed using the  
830 NEBNext DNA Library Prep Master Mix for 454 (E6070) and sequenced on an Illumina HiSeq 2500  
831 platform at the Danish National High-Throughput DNA Sequencing Centre. The libraries were  
832 sequenced twice to improve DNA recovery, producing 14.563 Gbp of data (17,979,669 reads) for BRS1,  
833 29.201 Gbp (36,050,926 reads) for BRS2, 0.709 Gbp (8,754,692 reads) for BRS3, 95.323 Gbp  
834 (117,683,541 reads) for BRS4 and 9.54 Gbp (110,485,402 reads) for BRS5. The merged E-LPCs yielded  
835 2.946 Gbp of data (3,637,328 reads) (Table S13).

836

### 837 **Sequence processing and microbial taxonomic profiling**

838 The potential for retrieving ancient IM data from palaeo-faecal remains is confounded by technical and  
839 biological variables (14, 44, 45). In technical terms, the detection of >90 microbial genera in DNA  
840 extraction and library preparation controls suggest that reagent and laboratory contamination can  
841 influence sequence-based IM analyses (84, 85, 86). The choice of DNA extraction protocols can also  
842 impact metagenomic compositional profiles (87). In biological terms, IM research generally focuses on  
843 the microbial community of the large intestine as expressed in stools, despite the fact that the 6.5 meter  
844 human digestive tract consists of three organs, *i.e.*, the stomach, small intestine and large intestine (6).  
845 Since microbial communities change along the length of the GI, differences exist between oral, intestinal  
846 and faecal taxonomic profiles in both modern (88) and ancient (25) instances. Moreover, and besides the  
847 influence on IM taxonomic composition of diet (57), age (61), seasonal variation (89) and host immuno-  
848 modulation (90), stool consistency also influence IM taxonomic composition (91). Differences in  
849 taxonomic composition between the cores and cortices of specimens have also been documented, with  
850 larger proportions of soil-derived taxa present in the cortices (24). In post-depositional terms, the  
851 retrieval of ancient IM data is confounded by on-going microbial activity and also environmental  
852 contamination (92, 93, 94). Instances of reverse contamination, *i.e.*, from a faecal specimen into the  
853 surrounding sediment, are also probable as the exchange of microbes between a stool and the  
854 surrounding sediment would certainly occur.

855 With these concerns in mind, taxonomic profiling was preceded by several data pre-processing  
856 steps. First, raw sequence reads were processed to remove all Illumina PhiX spikes, human reads and all  
857 exact duplicate reads present in the extraction ( $n = 1$ ) and library preparation ( $n = 1$ ) negative controls  
858 (E-LPCs) using BBDuk (95). Second, barcodes, adapters, reads shorter than 25 base-pairs (bp) and  
859 ‘quality score’ <25 were removed from the dataset (96, 97) using AdapterRemoval V2 (98). Taxonomic  
860 binning was then performed via comparison of the shotgun reads with the BLASTn v2.2.31+ NCBI *nt*  
861 database (99). Taxa were identified using MEGAN CE v6.10.10 (97) by using the weighted lowest  
862 common ancestor (‘wlca’) option and the default percent-to-cover value setting (‘80’) with parameter  
863 values set as follows: min. bit score: 50, expect value (*e*-value):  $1.0e-10$ , top percent: 10, min. support:  
864 10 and min complexity: 0.45. Species identifications were based on significant hits (bit score  $\geq 50$ ) and on  
865 MEGAN parameters established at ‘identities’: 97%, ‘positives’: 100% and no (0%) ‘gaps’. Comparisons  
866 of BRS IM sequence reads with those derived from other (comparative) IMs was performed by the  
867 subsampling in of the reads to the lowest number of reads present in any sample (*i.e.*, BRS1 (SC1) with  
868 56,682 filtered sequence reads).

869

### 870 **Ancient DNA authentication**

871 Molecular damage following death is a standard feature of all aDNA molecules. The accumulation of  
872 deaminated cytosine (uracil) within the overhanging ends of aDNA templates results in increasing  
873 cytosine (C) to thymine (T) misincorporation rates toward read starts, with matching guanine (G) to  
874 adenine (A) misincorporations increasing toward read ends in double-stranded library preparations (100,  
875 101). MapDamage (38, 39) is widely used to determine the incidence of cytosine (C) to thymine (T) and  
876 guanine (G) to adenine (A) substitution rates at the 5’-ends and 3’-ends of strands. MapDamage is not

877 however optimised for ancient samples lacking high genome coverage that would permit the  
878 identification of all possible misincorporations (101), and DNA damage patterns cannot be calculated for  
879 taxa with insufficient read counts, *i.e.*, <150 (101, 102). In addition, DNA fragmentation rates vary  
880 according to environmental conditions and the types of organisms involved (103, 104), often resulting in  
881 ‘alternative’ damage patterns (105-107). We therefore also validated the antiquity of putatively ancient  
882 taxa by a statistical method that compares post-mortem damage patterns indicative of the cytosine to  
883 thymine (C-T) substitutions at the 5’ ends of sequence reads (37, 108). High-quality filtered reads were  
884 aligned to comparative genomes (Table S13) using BWA (-n 0.02; -l 1024) (98) and duplicate sequence  
885 reads were removed using the Picard tools script MarkDuplicates  
886 (<https://broadinstitute.github.io/picard/>). Resulting alignments were used to perform statistical DNA  
887 damage estimation analyses which entailed the calculation of goodness-of-fit *p*-values ( $p = <0.05$ )  
888 indicative of significant cytosine to thymine (C-T) substitutions at the 5’ ends of sequence reads using  
889 PMDtools (<https://omictools.com/pmdtools-tool>) (37).

890 The authentication of aDNA sequence reads was furthermore based on the comparison of reads  
891 derived from negative (*i.e.*, extraction and library preparation) controls (84, 93, 94, 100, 101). In addition  
892 to the detection of >90 microbial taxa derived from reagents and laboratory contamination (85, 86), the  
893 probability that negative controls can become cross-contaminated during sample processing (103)  
894 complicates the authentication process. To establish the antiquity of microbial taxa occurring in the E-  
895 LPCs (*i.e.*, *Arthrobacter*, *Blautia*, *Klebsiella*, *Lactobacillus*, *Prevotella* and *Ruminococcus*), we  
896 compared the read yields in the BRS specimen with those in the E-LPCs. Since aDNA sequences are  
897 shorter than those derived from modern organisms, most frequently via contamination (104-108), DNA  
898 read length was also used as criteria in the authentication process. This criterion is particularly relevant in  
899 instances that exclude the use of bleach for the removal of contaminants (84), as is the case here. Lastly,  
900 evaluating ecological conformity, *i.e.*, excluding DNA reads that are derived from either non-indigenous  
901 taxa, foreign contaminants or false-positive identifications, such as *Apteryx* (kiwi), *Cyprinus* (carp),  
902 *Oncorhynchus* (salmon) and *Oryza* (rice) was used to assess taxonomic community composition for  
903 biological plausibility (93, 94). The authenticity (*i.e.*, prehistoric provenience) of microbial and  
904 macrobial sequence-derived taxa was therefore evaluated according to 1) validating the existence of C-T  
905 and G-A substitution rates, 2) statistical aDNA sequence damage estimation, 3) comparison to negative  
906 DNA E-LPCs, 4) DNA read length characteristics and 5) ecological conformity.

907

### 908 **Functional metabolic profiling**

909 To discern the metabolic pathways associated with the dietary and environmental factors characteristic of  
910 the Bantu-speaking forager-agro-pastoralist in question, we identified reads assigned to functional genes  
911 in the shotgun metagenome dataset. To ascertain the presence of microbial groups either positively or  
912 negatively correlated with specific metabolic profiles, we determined functional categories based on  
913 DIAMOND v0.36.98 BLASTx comparisons to the NCBI non-redundant protein database. The GI  
914 accessions were used to identify the Kyoto Encyclopaedia of Genes and Genomes (KEGG) orthologies  
915 (<https://www.genome.jp/kegg/>) in MEGAN CE v6.10.10 (97) by using the weighted lowest common  
916 ancestor (‘wlca’) option and the default percent-to-cover value setting (‘80’) with the parameters as  
917 follows: ‘*e*-value cut-off’: 1.0e-4, ‘top percent’:10, ‘min support’:10, ‘min complexity’:0.45 and  
918 ‘identity’: 97%. Comparisons of KEGG orthologies was performed following the subsampling to the  
919 lowest number of reads for any sample (*i.e.*, BRS1 (SC1) with 56,682 sequence reads).

920

### 921 **Comparative IM datasets**

922 To gain insight into the taxonomic and functional (metabolic) differences between the BRS IM ( $n = 4$ )  
923 and other ancient and modern (including ‘traditional’) IMs, we compared BRS with data derived from  
924 the shotgun metagenome sequencing of Malawian agro-pastoralists (51) (MG-RAST

925 <http://metagenomics.anl.gov/> accession number ‘qiime:621’), Tanzanian Hadza hunter-gatherers (26)  
926 (NCBI SRA SRP056480 Bioproject ID PRJNA278393), contemporary Italians (26) and the Copper Age  
927 (dated to *c.* 3,250 BC) Alpine (Tyrolean) Iceman (Ötzi) (109) (European Nucleotide Archive accession  
928 number ERP012908). For comparison, four metagenome samples (comprising two males and two  
929 females) were randomly selected from each of the comparative cohorts ( $n = 12$ ).

930

### 931 **Detecting antibiotic-resistance genes**

932 The resistome is defined as the complete set of antibiotic resistant genes (ARGs) presented in a microbial  
933 community, which is important for understanding the proliferation of pathogen antibiotic resistance (26).  
934 Sequence reads were aligned to the MEGARes database (110) using the Burrows-Wheeler Alignment  
935 tool (BWA) (111). The BRS IM resistome was analysed using Resistome Analyser  
936 (<https://github.com/cdeanj/resistomeanalyzer>) by applying the default threshold of ‘80’ to determine the  
937 gene significance and in order to decrease false positive gene identifications. Relative abundance for  
938 each of the resistance genes was calculated and sub-sampled to the lowest number of sequences (*i.e.*, the  
939 Ötzi ‘small intestine’ (Ötzi\_SI\_F) comprising 2,377 sequences) in any sample (*i.e.*, 28,524 sequences for  
940 a total of twelve samples). Because of low numbers of gene assignments, seven samples were excluded  
941 from the analysis.

942

### 943 **Statistical analyses**

944 Statistical analysis of the ancient and modern IM samples was performed by filtering the taxa and KEGG  
945 orthologies for >3 occurrences in at least 20% of the samples. The data was Hellinger-transformed and  
946 the Bray-Curtis dissimilarity matrix was used for the *vegdist* and *anosim* functions in the VEGAN  
947 (<https://www.rdocumentation.org/packages/vegan/versions/2.4-2>) package of R. The ordination plots  
948 were generated in the Phyloseq package (112) of R. Functional (metabolic) group significance tests were  
949 performed in Qiime v1.9.1 package (113) for the ancient and modern comparative cohorts and only gene  
950 categories with significant *p*-value ( $p < 0.05$ ) were included in the analyses. The *p*-values were corrected  
951 using false discovery rate (FDR) (114) to corrected *p* values (*q*) and Kruskal-Wallis values (*H*) to  
952 determine the significance of differences between samples. Hierarchical clustering using complete-  
953 linkage based on Spearman’s correlations was performed and visualised in R using ‘gplots’  
954 (<https://www.rdocumentation.org/packages/gplots/versions/3.0.1>). Heat-maps were generated using  
955 Spearman’s correlation and complete linkage method for microbial taxa and antibiotic resistance genes.  
956 Taxa were filtered for occurrence of >3 in at least 20% of the samples, and ARG data was filtered for the  
957 occurrence of >5 in at least 20% samples. Heat-map for the KEGG orthologies linked to 24 taxa (1,487  
958 categories) was produced by using Spearman’s correlation and Ward-linkage method.

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973 **SUPPLEMENTARY MATERIALS**

974

975 **Supplementary Figures**

976 Fig. S1. Additional information concerning the archaeological provenance of the BRS faecal specimen.

977 Fig. S2. Processing of the faecal specimen at the Centre for GeoGenetics, Copenhagen, Denmark.

978 Fig. S3. Dot-plot indicating the occurrence of statistically-significant C-T *p*-values.

979 Fig. S4. Biplot of  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  stable isotope values obtained for the BRS specimen.

980 Fig. S5. SEM analysis detected bacterial cells, plant fragments and saprophytic organisms.

981 Fig. S6. Heat-map indicating differences in taxonomic community structure for IM datasets.

982 Fig. S7. Comparing 'relative abundance' and 'presence-absence' as taxonomic representation.

983 Fig. S8. Comparison of the incidence of the twenty-four authenticated ancient IM taxa.

984 Fig. S9. Heat-map indicating the presence of fifteen functional ARGs identified.

985

986 **Supplementary Tables**

987 Table S1. Sequence reads for environmental- and subsistence-related taxa detected.

988 Table S2. Information concerning ( $^{14}\text{C}$ ) Accelerator Mass Spectrometry (AMS) dating.

989 Table S3a. Processing protocol and results for isotope analyses.

990 Table S3b. Results for isotope analyses (Merck standard).

991 Table S3c. Results for isotope analyses (DL-Valine standard).

992 Table S4. Abundance of bacterial taxonomic categories in the IM datasets.

993 Table S5. Sequence read-length distribution for taxa identified in this study.

994 Table S6. Significant KEGG pathways in the comparative IM datasets analysed.

995 Table S7. Enrichment and depletion of KO metabolic gene categories in the comparative IM sample cohorts based on *p*-value ( $p = <0.05$ ) designation.

997 Table S8. Enrichment and depletion of KO metabolic gene categories in the comparative IM sample cohorts based on false discovery rate (FDR) corrected *p*-values ( $q = <0.05$ ).

999 Table S9. Enrichment and depletion of KO metabolic gene categories in the ancient and modern comparative IM sample cohort as calculated for the twenty-four authenticated ancient IM taxa.

1000 Table S10. Comparison of relative abundance of antibiotic resistance genes in the comparative IM cohorts.

1001 Table S11. Raw and filtered high-quality sequence read counts as related to the comparative IM datasets.

1002 Table S12. Information concerning the comparative NCBI genomes used during this study.

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021 **REFERENCES**

- 1022 1. J. Lederberg, A. T. McCray, ‘Ome sweet’ omics: A genealogical treasury of words. *Scientist* 15, 8  
1023 (2010).
- 1024 2. C. Warinner, C. Speller, M. J. Collins, C. M. Lewis, Ancient human microbiomes. *J Hum Evol.* 79,  
1025 125-136 (2015).
- 1026 3. E. Thursby, N. Juge, Introduction to the human gut microbiota. *Biochem J.* 474, 1823-1836 (2017).
- 1027 4. L. W. van den Elsen, H. C. Poyntz, L. S. Weyrich, W. Young, E. E. Forbes-Blom, Embracing the gut  
1028 microbiota: The new frontier for inflammatory and infectious diseases. *Clin Transl Immunology* 6,  
1029 doi: 10.1038/cti.2016.91 (2017).
- 1030 5. J. C. Clemente, E. C. Pehrsson, M. J. Blaser, K. Sandhu, Z. Gao, B. Wang, M. Magris, G. Hidalgo,  
1031 M Contreras, Ó. Noya-Alarcón, O. Lander, J. McDonald, M. Cox, J. Walter, P. L. Oh, J. F. Ruiz, S.  
1032 Rodriguez, N. Shen, S. J. Song, J. Metcalf, R. Knight, G. Dantas, M. G. Dominguez-Bello, The  
1033 microbiome of uncontacted Amerindians. *Sci Adv.* 1, doi: 10.1126/sciadv.1500183 (2015).
- 1034 6. E. R. Davenport, J. G. Sanders, S. J. Song, K. R. Amato, A. G. Clark, R. Knight, The human  
1035 microbiome in evolution. *BMC Biol.* 15, doi: 10.1186/s12915-017-0454-7 (2017).
- 1036 7. M. J. Blaser, The past and future biology of the human microbiome in an age of extinctions. *Cell*  
1037 172, 1173-1177 (2018).
- 1038 8. S. L. Schnorr, M. Candela, S. Rampelli, M. Centanni, C. Consolandi, G. Basaglia, S. Turrone, E.  
1039 Biagi, C. Peano, M. Severgnini, J. Fiori, R. Gotti, G. De Bellis, D. Luiselli, P. Brigidi, A. Mabulla,  
1040 F. Marlowe, A. G. Henry, A. N. Crittenden, Gut microbiome of the Hadza hunter-gatherers. *Nat*  
1041 *Commun.* 5, doi: 10.1038/ncomms4654 (2014).
- 1042 9. A. J. Obregon-Tito, R. Y. Tito, J. Metcalf, K. Sankaranarayanan, J. C. Clemente, L. K. Ursell, Z.  
1043 Zech Xu, W. Van Treuren, R. Knight, P. M. Gaffney, P. Spicer, P. Lawson, L. Marin-Reyes, O.  
1044 Trujillo-Villarreal, M. Foster, E. Gujja-Poma, L. Troncoso-Corzo, C. Warinner, A. T. Ozga, C. M.  
1045 Lewis, Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat Commun.*  
1046 6505, doi: 10.1038/ncomms7505 (2015).
- 1047 10. C. Warinner, C. M. Lewis, Microbiome and health in past and present human populations. *Am*  
1048 *Anthrop.* 117, 740-741 (2015).
- 1049 11. R. Lee, R. H. Daly, Cambridge encyclopaedia of hunters and gatherers. Cambridge, Cambridge  
1050 University Press ISBN 9780521609197 (1999).
- 1051 12. A. Gomez, K. J. Petzelkova, M. B. Burns, C. J. Yeoman, A. R. Amato, K. Vlckova, D. Modry, A.  
1052 Todd, C. A. Jost Robinson, M. J. Remis, M. G. Torralba, E. Morton, J. D. Umaña, F. Carbonero, H.  
1053 R. Gaskins, K. E. Nelson, B. A. Wilson, R. M. Stumpf, B. A. White, S. R. Leigh, Gut microbiome of  
1054 coexisting BaAka Pygmies and Bantu reflects gradients of traditional subsistence patterns. *Cell Rep.*  
1055 14, 2142-2153 (2016).
- 1056 13. C. Girard, N. Tromas, M. Amyot, B. J. Shapiro, Gut microbiome of the Canadian Arctic Inuit.  
1057 *mSphere* 2, doi: 10.1128/mSphere.00297-16 (2017).
- 1058 14. A. Y. Voigt, P. I. Costea, J. R. Kultima, S. S. Li, G. Zeller, S. Sunagawa, P. Bork, Temporal and  
1059 technical variability of human gut metagenomes. *Genome Biol.* 16, doi: 10.1186/s13059-015-0639-8  
1060 (2015).
- 1061 15. J. Walter, R. Ley, The human gut microbiome: Ecology and recent evolutionary changes. *Annu Rev*  
1062 *Microbiol.* 65, 411-429 (2011).
- 1063 16. M. J. Blaser, S. Falkow, What are the consequences of the disappearing human microbiota? *Nat Rev*  
1064 *Microbiol.* 7, 887-894 (2009).
- 1065 17. C. J. Adler, K. Dobney, L. S. Weyrich, J. Kaidonis, A. W. Walker, W. Haak, C. J. Bradshaw, G.  
1066 Townsend, A. Soltysiak, K. W. Alt, J. Parkhill, A. Cooper, Sequencing ancient calcified dental  
1067 plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial  
1068 revolutions. *Nat Genet.* 45, 450-455 (2013).

- 1069 18. S. Quercia, M. Candela, C. Giuliani, S. Turrone, D. Luiselli, S. Rampelli, P. Brigidi, C. Franceschi,  
1070 M. G. Bacalini, P. Garagnani, C. Pirazzini, From lifetime to evolution: Timescales of human gut  
1071 microbiota adaptation. *Front Microbiol.* 5, doi: 10.3389/fmicb.2014.00587 (2014).
- 1072 19. A. H. Moeller, A. Caro-Quintero, D. Mjungu, A. V. Georgiev, E. V. Lonsdorf, M. N. Muller, A. E.  
1073 Pusey, M. Peeters, B. H. Hahn, H. Ochman, Cospeciation of gut microbiota with hominids. *Science*  
1074 353, 380-382 (2016).
- 1075 20. H. Reyes-Centeno, K. Harvati, G. Jäger, Tracking modern human population history from linguistic  
1076 and cranial phenotype. *Sci Rep.* 6, doi: 10.1038/srep36645 (2016).
- 1077 21. C. Houldcroft, J. B. Ramond, R. F. Rifkin, S. J. Underdown, Migrating microbes: What pathogens  
1078 can tell us about population movements and human evolution. *Ann Hum Biol.* 44, 397-407 (2017).
- 1079 22. B. Linz, F. Balloux, Y. Moodley, A. Manica, H. Liu, P. Roumagnac, D. Falush, C. Stamer, F.  
1080 Prugnolle, S. W. van der Merwe, Y. Yamaoka, D. Y. Graham, E. Perez-Trallero, T. Wadstrom, S.  
1081 Suerbaum, M. Achtman, An African origin for the intimate association between humans and  
1082 *Helicobacter pylori*. *Nature* 445, 915-918 (2007).
- 1083 23. M. T. Gilbert, D. L. Jenkins, A. Götherstrom, N. Naveran, J. J. Sanchez, M. Hofreiter, P. F.  
1084 Thomsen, J. Binladen, T. F. Higham, R. M. Yohe, R. Parr, L. S. Cummings, E. Willerslev, DNA  
1085 from pre-Clovis human coprolites in Oregon, North America. *Science* 320, 786-789 (2008).
- 1086 24. R. J. Cano, J. Rivera-Perez, G. A. Toranzos, T. M. Santiago-Rodriguez, Y. M. Narganes-Storde, L.  
1087 Chanlatte-Baik, E. García-Roldán, L. Bunkley-Williams, S. E. Massey, Paleomicrobiology:  
1088 Revealing fecal microbiomes of ancient indigenous cultures. *PLoS One* 9, doi:  
1089 10.1371/journal.pone.0106833 (2014).
- 1090 25. T. M. Santiago-Rodriguez, G. Fornaciari, S. Luciani, S. E. Dowd, G. A. Toranzos, I. Marota, R. J.  
1091 Cano, Gut microbiome of an 11<sup>th</sup> century A.D. Pre-Columbian Andean mummy. *PLoS One* 10, doi:  
1092 10.1371/journal.pone.0138135 (2015).
- 1093 26. S. Rampelli, S. L. Schnorr, C. Consolandi, S. Turrone, M. Severgnini, C. Peano, P. Brigidi, A. N.  
1094 Crittenden, A. G. Henry, M. Candela, Metagenome sequencing of the Hadza hunter-gatherer gut  
1095 microbiota. *Curr Biol.* 25, 1682-1693 (2015).
- 1096 27. G. Porraz, A. Val, L. Dayet, P. de la Pena, K. Douze, C. E. Miller, M. Murungi, C. Tribolo, V.  
1097 Schmid, C. Sievers, Bushman Rock Shelter (Limpopo, South Africa): A perspective from the edge of  
1098 the Highveld. *S Afr Archaeol Bull.* 70, 166-179 (2015).
- 1099 28. T. Russell, F. Silva, J. Steele, Modelling the spread of farming in the Bantu-speaking regions of  
1100 Africa: an archaeology-based phylogeography. *PLoS One* 9, doi: 10.1371/journal.pone.0087854  
1101 (2014).
- 1102 29. G. Porraz, A. Val, C. Tribolo, N. Mercier, P. de la Peña, M. M. Haaland, M. Igreja, C. E. Miller, V.  
1103 Schmid, The MIS5 Pietersburg at '28' Bushman Rock Shelter, Limpopo Province, South Africa.  
1104 *PLoS One* 13, doi: <https://doi.org/10.1371/journal.pone.0202853> (2018).
- 1105 30. J. R. Wood, J. M. Wilmshurst, A protocol for subsampling Late Quaternary coprolites for multi-  
1106 proxy analysis. *Quat Sci Rev.* 138, 1-5 (2014).
- 1107 31. C. Walker, Signs of the wild: A field guide to the spoor and signs of the mammals of southern  
1108 Africa. Cape Town: Struik Nature (2015).
- 1109 32. R. D. Estes, The behavior guide to African mammals: Including hoofed mammals, carnivores,  
1110 primates. Berkeley, University of California Press pp. 1-611 (2012).
- 1111 33. T. N. Huffman, Mapungubwe and Great Zimbabwe: The origin and spread of social complexity in  
1112 southern Africa. *J Anthropol Archaeol.* 28, 37-54 (2007).
- 1113 34. M. Samuelson, Rendering the Cape-as-port: Sea-mountain, Cape of Storms/Good Hope, Adamastor  
1114 and local-world literary formations. *J S Afr Stud.* 42, 523-537 (2016).

- 1115 35. M. G. Campana, N. Robles García, F. J. Rühli, N. Tuross, False positives complicate ancient  
1116 pathogen identifications using high-throughput shotgun sequencing. *BMC Res Notes* 7, doi:  
1117 10.1186/1756-0500-7-111 (2014).
- 1118 36. E. Willerslev, L. Davison, M. Moora, M. Zobel, E. Coissac, M. E. Edwards, E. D. Lorenzen, M.  
1119 Vestergård, G. Gussarova, J. Haile, J. Craine, L. Gielly, S. Boessenkool, L. S. Epp, P. B. Pearman,  
1120 R. Cheddadi, D. Murray, K. A. Bråthen, N. Yoccoz, H. Binney, C. Cruaud, P. Wincker, T. Goslar, I.  
1121 Greve Alsos, E. Bellemain, A. Krag Brysting, R. Elven, J. H. Sønstebo, J. Murton, A. Sher, M.  
1122 Rasmussen, R. Rønn, T. Mourier, A. Cooper, J. Austin, P. Möller, D. Froese, G. Zazula, F.  
1123 Pompanon, D. Rioux, V. Niderkorn, A. Tikhonov, G. Savvinov, R. G. Roberts, R. D. E. MacPhee,  
1124 M. T. P. Gilbert, K. H. Kjær, L. Orlando, C. Brochmann, P. Taberlet, Fifty thousand years of Arctic  
1125 vegetation and megafaunal diet. *Nature* 506, 47-51 (2014).
- 1126 37. C. J. Weiß, M. Dannemann, K. Prüfer, H. A. Burbano, Contesting the presence of wheat in the  
1127 British Isles 8,000 years ago by assessing ancient DNA authenticity from low-coverage data. *Elife* 4,  
1128 doi: 10.7554/eLife.10005 (2015).
- 1129 38. A. Ginolhac, M. Rasmussen, M. T. P. Gilbert, E. Willerslev, L. Orlando, mapDamage: Testing for  
1130 damage patterns in ancient DNA sequences. *Bioinformatics* 27, 2153-2155 (2011).
- 1131 39. H. Jónsson, A. Ginolhac, M. Schubert, P. L. Johnson, L. Orlando, mapDamage 2.0: Fast approximate  
1132 Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682-1684 (2013).
- 1133 40. S. Badenhorst, Intensive hunting during the Iron Age of Southern Africa. *J Hum Palaeoecol.* 20, 41-  
1134 45 (2015).
- 1135 41. J. Lloyd-Price, G. Abu-Ali, C. Huttenhower, The healthy human microbiome. *Genome Med.* 8, doi:  
1136 <https://doi.org/10.1186/s13073-016-0307-y> (2016).
- 1137 42. J. Qin, R. Li, J. Raes, M. Arumugam, K. Solvsten Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F.  
1138 Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng,  
1139 Y. Xie, J. Tap, P. Lepage, M. Bertalan, J. M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H. B.  
1140 Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, S. Li, M. Jian, Y.  
1141 Zhou, Y. Li, X. Zhang, S. Li, N. Qin, H. Yang, J. Wang, S. Brunak, J. Doré, F. Guarner, K.  
1142 Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, MetaHIT Consortium, P. Bork, S. D. Ehrlich,  
1143 J. Wang, A human gut microbial gene catalog established by metagenomic sequencing. *Nature* 464,  
1144 59-65 (2010).
- 1145 43. M. Derrien, J. E. van Hylckama Vlieg, Fate, activity, and impact of ingested bacteria within the  
1146 human gut microbiota. *Trends Microbiol.* 23, 354-366 (2015).
- 1147 44. G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, J. J. Egozcue, Microbiome datasets are  
1148 compositional: And this is not optional. *Front Microbiol.* 8, doi: [https://doi.org/10.3389/fmicb.](https://doi.org/10.3389/fmicb.2017.02224)  
1149 2017.02224 (2017).
- 1150 45. L. Kistler, R. Ware, O. Smith, M. Collins, R. G. Allaby, A new model for ancient DNA decay based  
1151 on paleogenomic meta-analysis. *Nucleic Acids Res.* 45, 6310-6320 (2017).
- 1152 46. I. M. Velsko, L. A. F. Frantz, A. Herbig, G. Larson, C. Warinner, Selection of appropriate  
1153 metagenome taxonomic classifiers for ancient microbiome research. *mSystems* 3, doi:  
1154 10.1128/mSystems.00080-18 (2018).
- 1155 47. M. E. J. Woolhouse, Where do emerging pathogens come from? *Microbe* 1, 511-515 (2006).
- 1156 48. Y. Nédélec, J. Sanz, G. Baharian, Z. A. Szpiech, A. Pacis, A. Dumaine, J. C. Grenier, A. Freiman, A.  
1157 J. Sams, S. Hebert, A. Pagé Sabourin, F. Luca, R. Blekhman, R. D. Hernandez, R. Pique-Regi, J.  
1158 Tung, V. Yotova, L. B. Barreiro, Genetic ancestry and natural selection drive population differences  
1159 in immune responses to pathogens. *Cell* 167, 657-669 (2017).
- 1160 49. S. E. Kessler, T. R. Bonnell, R. W. Byrne, C. A. Chapman, Selection to outsmart the germs: The  
1161 evolution of disease recognition and social cognition. *J Hum Evol.* 108, 92-109 (2017).



- 1162 50. R. Thornhill, C. L. Fincher, The parasite-stress theory of values and sociality. London: Springer pp.  
1163 1-440 (2014).
- 1164 51. E. S. Pierce, Could *Mycobacterium avium* subspecies *paratuberculosis* cause Crohn's disease,  
1165 ulcerative colitis...and colorectal cancer? BMC Infect. Agents Cancer 13, doi:  
1166 <https://doi.org/10.1186/s13027-017-0172-3> (2018).
- 1167 52. N. R. Shin, T. W. Whon, J. W. Bae, *Proteobacteria*: Microbial signature of dysbiosis in gut  
1168 microbiota. Trends Biotechnol. 33, 496-503 (2015).
- 1169 53. E. R. Hughes, M. G. Winter, B. A. Buerkop, L. Spiga, T. Furtado de Carvalho, W. Zhu, C. C. Gillis,  
1170 L. Büttner, M. P. Smoot, C. L. Behrendt, S. Cherry, R. L. Santos, L. V. Hooper, S. E. Winter,  
1171 Microbial respiration and formate oxidation as metabolic signatures of inflammation-associated  
1172 dysbiosis. Cell Host Microbe. 21, 208-219 (2017).
- 1173 54. R. E. Ley, P. J. Turnbaugh, S. Klein, J. I. Gordon, Microbial ecology: Human gut microbes  
1174 associated with obesity. Nature 444, 1022-1023 (2006).
- 1175 55. J. C. Clemente, L. K. Ursell, L. W. Parfrey, R. Knight, The impact of the gut microbiota on human  
1176 health: An integrative view. Cell 148, 1258-1270 (2012).
- 1177 56. J. R. Marchesi, D. H. Adams, F. Fava, G. D. Hermes, G. M. Hirschfield, G. Hold, M. N. Quraishi, L.  
1178 Kinross, H. Smidt, K. M. Tuohy, L. V. Thomas, E. G. Zoetendal, A. Hart, The gut microbiota and  
1179 host health: A new clinical frontier. Gut 65, 330-339 (2016).
- 1180 57. C. De Filippo, D. Cavalieri, M. Di Paola, M. Ramazzotti, J. B. Poullet, S. Massart, S. Collini, G.  
1181 Pieraccini, P. Lionetti, Impact of diet in shaping gut microbiota revealed by a comparative study in  
1182 children from Europe and rural Africa. Proc Natl Acad Sci U S A. 107, 14691-14696 (2010).
- 1183 58. T. V. Maier, M. Lucio, L. H. Lee, N. C. VerBerkmoes, C. J. Brislawn, J. Bernhardt, R. Lamendella,  
1184 J. E. McDermott, N. Bergeron, S. S. Heinzmann, J. T. Morton, A. González, G. Ackermann, R.  
1185 Knight, K. Riedel, R. M. Krauss, P. Schmitt-Kopplin, J. K. Jansson, Impact of dietary resistant starch  
1186 on the human gut microbiome, metaproteome, and metabolome. MBio 8, doi: 10.1128/mBio.01343-  
1187 17 (2017).
- 1188 59. M. Arumugam, J. Raes, E. Pelletier, D. Le Paslier, T. Yamada, D. R. Mende, G. R. Fernandes, J.  
1189 Tap, T. Bruls, J.-M. Batto, M. Bertalan, N. Borruel, F. Casellas, L. Fernandez, L. Gautier, T.  
1190 Hansen, M. Hattori, T. Hayashi, M. Kleerebezem, K. Kurokawa, M. Leclerc, F. Levenez, C.  
1191 Manichanh, H. B. Nielsen, T. Nielsen, N. Pons, J. Poulain, J. Qin, T. Sicheritz-Ponten, S. Tims, D.  
1192 Torrents, E. Ugarte, E. G. Zoetendal, J. Wang, F. Guarner, O. Pedersen, W. M. de Vos, S. Brunak, J.  
1193 Doré, MetaHIT Consortium, J. Weissenbach, S. D. Ehrlich, P. Bork, Enterotypes of the human gut  
1194 microbiome. Nature 473, 174-180 (2011).
- 1195 60. G. D. Wu, J. Chen, C. Hoffmann, K. Bittinger, Y. Y. Chen, S. A. Keilbaugh, M. Bewtra, D. Knights,  
1196 W. A. Walters, R. Knight, R. Sinha, E. Gilroy, K. Gupta, R. Baldassano, L. Nessel, H. Li, F. D.  
1197 Bushman, J. D. Lewis, Linking long-term dietary patterns with gut microbial enterotypes Science  
1198 334, 105-108 (2011).
- 1199 61. T. Yatsunenko, F. E. Rey, M. J. Manary, I. Trehan, M. G. Dominguez-Bello, M. Contreras, M.  
1200 Magris, G. Hidalgo, R. N. Baldassano, A. P. Anokhin, A. C. Heath, B. Warner, J. Reeder, J.  
1201 Kuczynski, J. G. Caporaso, C. A. Lozupone, C. Lauber, J. C. Clemente, D. Knights, R. Knight, J. I.  
1202 Gordon, Human gut microbiome viewed across age and geography. Nature 486, 222-227 (2012).
- 1203 62. M. J. Rodríguez-Vaquero, M. R. Alberto, M. C. Manca de Nadra, Antibacterial effect of phenolic  
1204 compounds from different wines. Food Control 18, 93-101 (2007).
- 1205 63. E. A. McKenney, M. Ashwell, J. E. Lambert, V. Fellner, Fecal microbial diversity and putative  
1206 function in captive western lowland gorillas (*Gorilla gorilla gorilla*), common chimpanzees (*Pan*  
1207 *troglodytes*), Hamadryas baboons (*Papio hamadryas*) and binturongs (*Arctictis binturong*). Integr  
1208 Zool. 9, 557-569 (2014).

- 1209 64. B. S. Samuel, E. E. Hansen, J. K. Manchester, P. M. Coutinho, B. Henrissat, R. Fulton, P. Latreille,  
1210 K. Kim, R. K. Wilson, J. I. Gordon, Genomic and metabolic adaptations of *Methanobrevibacter*  
1211 *smithii* to the human gut. Proc Natl Acad Sci U S A. 104, 10643-10648 (2007).
- 1212 65. T. Chen, W. Long, C. Zhang, S. Liu, L. Zhao, B. R. Hamaker, Fiber-utilizing capacity varies in  
1213 *Prevotella*- versus *Bacteroides*-dominated gut microbiota. Sci. Rep. 7, doi: 10.1038/s41598-017-  
1214 02995-4 (2017).
- 1215 66. L. R. Lopetuso, F. Scaldaferri, V. Petito, A. Gasbarrini, Commensal *Clostridia*: Leading players in  
1216 the maintenance of gut homeostasis. Gut Pathog. 5, doi: 10.1186/1757-4749-5-23 (2013).
- 1217 67. C. Chassard, E. Delmas, C. Robert, A. Bernalier-Donadille, The cellulose-degrading microbial  
1218 community of the human gut varies according to the presence or absence of methanogens. FEMS  
1219 Microbiol Ecol. 74, 205-213 (2010).
- 1220 68. E. R. Morton, J. Lynch, A. Froment, S. Lafosse, E. Heyer, M. Przeworski, R. Blekhman, L. Ségurel,  
1221 Variation in rural African gut microbiota is strongly correlated with colonization by *Entamoeba* and  
1222 subsistence. PLoS Genet. 11, doi: 10.1371/journal.pgen.1005658 (2015).
- 1223 69. C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson, R. Knight, Diversity, stability and  
1224 resilience of the human gut microbiota. Nature 489, 220-230 (2012).
- 1225 70. P. Vangay, A. J. Johnson, T. L. Ward, G. A. Al-Ghalith, R. R. Shields-Cutler, B. M. Hillmann, S. K.  
1226 Lucas, L. K. Beura, E. A. Thompson, L. M. Till, R. Batres, B. Paw, S. L. Pergament, P. Saenyakul,  
1227 M. Xiong, A. D. Kim, G. Kim, D. Masopust, D. Knight, US immigration westernizes the human gut  
1228 microbiome. Cell 175, doi: 10.1016/j.cell.2018.10.029 (2018).
- 1229 71. K. Tomita, T. Nagura, Y. Okuhara, H. Nakajima-Adachi, N. Shigematsu, T. Aritsuka, S.  
1230 Kaminogawa, S. Hachimura, Dietary melibiose regulates *th* cell response and enhances the induction  
1231 of oral tolerance. Biosci Biotechnol Biochem. 71, 2774-2280 (2007).
- 1232 72. H. Fiedler, Short-chain chlorinated paraffins: Production, use and international regulations. In: Boer  
1233 J. (ed.) Chlorinated paraffins. The handbook of environmental chemistry. Berlin, Springer pp. 1-41  
1234 (2010).
- 1235 73. J. Glüge, Z. Wang, C. Bogdal, M. Scheringer, K. Hungerbühler, Global production, use, and  
1236 emission volumes of short-chain chlorinated paraffins: A minimum scenario. Sci Total Environ. 573,  
1237 1132-46 (2016).
- 1238 74. A. Paoli, A. Bianco, K. A. Grimaldi, A. Lodi, G. Bosco, Long term successful weight loss with a  
1239 combination biphasic ketogenic Mediterranean diet and Mediterranean diet maintenance protocol.  
1240 Nutrients 5, 5205-5217 (2013).
- 1241 75. V. M. D'Costa, C. E. King, L. Kalan, M. Morar, W. W. Sung, C. Schwarz, D. Froese, G. Zazula, F.  
1242 Calmels, R. Debruyne, G. B. Golding, H. N. Poinar, G. D. Wright, Antibiotic resistance is ancient.  
1243 Nature 477, 457-461 (2011).
- 1244 76. R. Kaur, V. Gautam, P. Ray, G. Singh, L. Singhal, R. Tiwari, Daptomycin susceptibility of  
1245 methicillin resistant *Staphylococcus aureus* (MRSA) Indian J Med Res. 136, 676-677 (2012).
- 1246 77. K. Ye, F. Gao, D. Wang, O. Bar-Yosef, A. Keinan, Dietary adaptation of FADS genes in Europe  
1247 varied across time and geography. Nat Ecol Evol. 1, doi: 10.1038/s41559-017-0167 (2017).
- 1248 78. G. A. W. Rook, L. R. Brunet, Microbes, immunoregulation, and the gut. Gut 54, 317-320 (2005).
- 1249 79. B. Stecher, L. Maier, W. D. Hardt, 'Blooming' in the gut: How dysbiosis might contribute to  
1250 pathogen evolution. Nat Rev Microbiol. 11, 277-284 (2013).
- 1251 80. M. Stuiver, P. D. Quay, Atmospheric <sup>14</sup>C changes resulting from fossil fuel CO<sup>2</sup> release and cosmic  
1252 ray flux variability. Earth and Planetary Science 53, 349-362 (1981).
- 1253 81. M. Nemeč, L. Wacker, H. Gäggeler, Optimization of the graphitization process at AGE-1.  
1254 Radiocarbon 52, 1380-1393 (2010).

- 1255 82. S. Woodborne, D. Huchzermeyer, D. Govender, D. J. Pienaar, G. Hall, J. G. Myburgh, A. R.  
1256 Deacon, J. Venter, N. Lubcker, Ecosystem change and the Olifants River crocodile mass mortality  
1257 events. *Ecosphere* 3, doi: <http://dx.doi.org/10.1890/ES12-00170.1> (2012).
- 1258 83. B. Dufour, M. Le Bailly, Testing new parasite egg extraction methods in paleoparasitology and an  
1259 attempt at quantification. *Int J Paleopathol.* 3, 199-203 (2013).
- 1260 84. C. Warinner, A. Herbig, A. Mann, J. A. Fellows Yates, C. L. Weiß, H. A. Burbano, L. Orlando, J.  
1261 Krause, A robust framework for microbial archaeology. *Annu Rev Genomics Hum Genet.* 18, 321-  
1262 356 (2017).
- 1263 85. S. J. Salter, M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, J.  
1264 Parkhill, N. J. Loman, A. W. Walker, Reagent and laboratory contamination can critically impact  
1265 sequence-based microbiome analyses. *BMC Biol.* 12, doi: 10.1186/s12915-014-0087-z (2014).
- 1266 86. A. P. Lauder, A. M. Roche, S. Sherrill-Mix, A. Bailey, A. L. Laughlin, K. Bittinger, R. Leite, M. A.  
1267 Elovitz, S. Parry, F. D. Bushman, Comparison of placenta samples with contamination controls does  
1268 not provide evidence for a distinct placenta microbiota. *Microbiome* 4, doi:  
1269 <https://doi.org/10.1186/s40168-016-0172-3> (2016).
- 1270 87. P. I. Costea, G. Zeller, S. Sunagawa, E. Pelletier, A. Alberti, F. Levenez, M. Tramontano, M.  
1271 Driessen, R. Hercog, F. Jung, J. Roat Kultima, M. R. Hayward, L. P. Coelho, E. Allen-Vercoe, L.  
1272 Bertrand, M. Blaut, J. R. M. Brown, T. Carton, S. Cools-Portier, M. Daigneault, M. Derrien, A.  
1273 Druesne, W. M. de Vos, B. B. Finlay, H. J. Flint, F. Guarner, M. Hattori, H. Heilig, R. A. Luna, J.  
1274 van Hylckama Vlieg, J. Junick, I. Klymiuk, P. Langella, E. Le Chatelier, V. Mai, C. Manichanh, J. C.  
1275 Martin, C. Mery, H. Morita, P. W. O'Toole, C. Orvain, K. Raosaheb Patil, J. Penders, S. Persson, N.  
1276 Pons, M. Popova, A. Salonen, D. Saulnier, K. P. Scott, B. Singh, K. Slezak, P. Veiga, J. Versalovic,  
1277 L. Zhao, E. G. Zoetendal, S. D. Ehrlich, J. Dore, P. Bork, Towards standards for human fecal sample  
1278 processing in metagenomic studies. *Nat Biotechnol.* 35, 1069-1076 (2017).
- 1279 88. J. C. Stearns, M. D. J. Lynch, D. B. Senadheera, H. C. Tenenbaum, M. B. Goldberg, D. G.  
1280 Cvitkovitch, K. Croitoru, G. Moreno-Hagelsieb, J. D. Neufeld, Bacterial biogeography of the human  
1281 digestive tract. *Sci. Rep.* 1, doi: 10.1038/srep00170 (2011).
- 1282 89. S. A. Smits, C. G. Gonzalez, J. Leach, A. Manjuran, E. D. Sonnenburg, J. S. Lichtman, J.  
1283 Chagalucha, M. G. Dominguez-Bello, G. Reid, R. Knight, J. E. Elias, J. L. Sonnenburg, Seasonal  
1284 cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 353, 802-806  
1285 (2017).
- 1286 90. G. P. Donaldson, S. M. Lee, S. K. Mazmanian, Gut biogeography of the bacterial microbiota. *Nat*  
1287 *Rev Microbiol.* 14, 20-32 (2016).
- 1288 91. D. Vandeputte, G. Falony, S. Vieira-Silva, R. Y. Tito, M. Joossens, J. Raes, Stool consistency is  
1289 strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth  
1290 rates. *BMJ Gut* 65, 57-62 (2016).
- 1291 92. S. J. Ott, M. Musfeldt, K. N. Timmis, J. Hampe, D. F. Wenderoth, S. Schreiber, *In vitro* alterations  
1292 of intestinal bacterial microbiota in fecal samples during storage. *Diagn Microbiol Infect Dis.* 50,  
1293 237-245 (2004).
- 1294 93. R. Y. Tito, S. Macmil, G. Wiley, F. Najjar, L. Cleeland, C. Qu, P. Wang, F. Romagne, S. Leonard, A.  
1295 Jiménez Ruiz, K. Reinhard, B. A. Roe, C. M. Lewis Jr., Phylotyping and functional analysis of two  
1296 ancient human microbiomes. *PLoS One* 3, doi: 10.1371/journal.pone.0003703 (2008).
- 1297 94. R. Y. Tito, D. Knights, J. Metcalf, A. J. Obregon-Tito, L. Cleeland, F. Najjar, B. Roe, K. Reinhard,  
1298 K. Sobolik, S. Belknap, M. Foster, P. Spicer, R. Knight, C. M. Lewis Jr., Insights from  
1299 characterizing extinct human gut microbiomes. *PLoS One* 12, doi:  
1300 <https://doi.org/10.1371/journal.pone.0051146> (2012).
- 1301 95. B. Bushnell, BMap. <http://sourceforge.net/projects/bbmap/> (2014).

- 1302 96. S. Sawyer, J. Krause, K. Guschanski, V. Savolainen, S. Pääbo, Temporal patterns of nucleotide  
1303 misincorporations and DNA fragmentation in ancient DNA. *PLoS One* 7, doi:  
1304 10.1371/journal.pone.0034131 (2012).
- 1305 97. J. Dabney, M. Knapp, I. Glocke, M. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. García,  
1306 S. Pääbo, J. Arsuag, M. Meyer, Complete mitochondrial genome sequence of a Middle Pleistocene  
1307 cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A.* 110, 15758-  
1308 15763 (2013).
- 1309 98. M. Schubert, S. Lindgreen, L. Orlando, AdapterRemoval v2: Rapid adapter trimming, identification,  
1310 and read merging. *BMC Res Notes.* 9, doi: 10.1186/s13104-016-1900-2 (2016).
- 1311 99. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J*  
1312 *Mol Biol.* 215, 403-410 (1990).
- 1313 100. D. H. Huson, S. Beier, I. Flade, A. Górská, M. El-Hadidi, S. Mitra, H. Ruscheweyh, R. Tappu,  
1314 MEGAN Community Edition: Interactive exploration and analysis of large-scale microbiome  
1315 sequencing data. *PLoS Comput Biol.* 12, doi: 10.1371/journal.pcbi.1004957 (2016).
- 1316 101. A. W. Briggs, U. Stenzel, P. L. Johnson, R. E. Green, J. Kelso, K. Prüfer, M. Meyer, J. Krause,  
1317 M. T. Ronan, M. Lachmann, S. Pääbo, Patterns of damage in genomic DNA sequences from a  
1318 Neandertal. *Proc Natl Acad Sci U S A.* 104, 14616-14621 (2007).
- 1319 102. F. V. Seersholm, M. W. Pedersen, M. J. Søre, H. Shokry, S. S. Mak, A. Ruter, M. Raghavan, W.  
1320 Fitzhugh, K. H. Kjær, E. Willerslev, M. Meldgaard, C. M. O. Kapel, A. J. Hansen DNA evidence of  
1321 bowhead whale exploitation by Greenlandic Paleo-Inuit 4,000 years ago. *Nat Commun.* 7, doi:  
1322 10.1038/ncomms13389 (2016).
- 1323 103. Å. J. Vågane, A. Herbig, M. G. Campana, N. M. Robles García, C. Warinner, S. Sabin, M. A.  
1324 Spyrou, A. Andrades Valtueña, D. Huson, N. Tuross, K. I. Bos, J. Krause, *Salmonella enterica*  
1325 genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat Ecol Evol.* 2, 520-528  
1326 (2017).
- 1327 104. J. C. Tackney, B. A. Potter, J. Raff, M. Powers, W. S. Watkins, D. Warner, J. D. Reuther, J. D.  
1328 Irish, D. H. O'Rourke, Two contemporaneous mitogenomes from terminal Pleistocene burials in  
1329 eastern Beringia. *Proc Natl Acad Sci U S A.* 112, doi: 10.1073/pnas.1511903112 (2015).
- 1330 105. J. Dabney, M. Meyer, S. Pääbo, Ancient DNA damage. *Mol Ecol Res.* 17, doi: 10.1111/1755-  
1331 0998.12595 (2017).
- 1332 106. C. L. Weiß, V. J. Schuenemann, J. Devos, G. Shirsekar, E. Reiter, B. A. Gould, J. R.  
1333 Stinchcombe, J. Krause, H. A. Burbano, Temporal patterns of damage and decay kinetics of DNA  
1334 retrieved from plant herbarium specimens. *R Soc Open Sci.* 3, doi: 10.1098/rsos.160239 (2016).
- 1335 107. H. Malmström, E. M. Svensson, M. T. P. Gilbert, E. Willerslev, A. Götherström, G. Holmlund,  
1336 More on contamination: The use of asymmetric molecular behavior to identify authentic ancient  
1337 human DNA. *Mol Biol Evol.* 24, 998-1004 (2007).
- 1338 108. K. Prüfer, U. Stenzel, M. Hofreiter, S. Pääbo, J. Kelso, R. E. Green, Computational challenges in  
1339 the analysis of ancient DNA. *Genome Biol.* 11, doi: 10.1186/gb-2010-11-5-r47 (2010).
- 1340 109. G. A. Lugli, C. Milani, L. Mancabelli, F. Turroni, C. Ferrario, S. Duranti, D. van Sinderen, M  
1341 Ventura, Ancient bacteria of the Ötzi's microbiome: A genomic tale from the Copper Age.  
1342 *Microbiome* 5, doi: 10.1186/s40168-016-0221-y (2017).
- 1343 110. S. M. Lakin, C. Dean, N. R. Noyes, A. Dettenwanger, A. S. Ross, E. Doster, P. Rovira, Z. Abdo,  
1344 K. L. Jones, J. Ruiz, K. E. Belk, P. S. Morley, C. Boucher, MEGAREs: An antimicrobial resistance  
1345 database for high throughput sequencing. *Nucleic Acids Res.* 45, doi: 10.1093/nar/gkw1009 (2017).
- 1346 111. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform.  
1347 *Bioinformatics* 25, 1754-1760 (2009).
- 1348 112. P. J. McMurdie, S. Holmes, phyloseq: an R package for reproducible interactive analysis and  
1349 graphics of microbiome census data. *PLoS One* 8, doi: 10.1371/journal.pone.0061217 (2013).

- 1350 113. J. G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N.  
1351 Fierer, A. G. Pena, J. K Goodrich, J. I. Gordon, G. A. Huttenhower, S. T. Kelley, D. Knights, J. E. Koenig,  
1352 R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J.  
1353 Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight, QIIME allows  
1354 analysis of high-throughput community sequencing data. *Nat Methods* 7, doi: 10.1038/nmeth.f.303  
1355 (2010).
- 1356 114. L. Jiang, A. Amir, J. T. Morton, R. Heller, E. Arias-Castro, R. Knight, Discrete false-discovery  
1357 rate improves identification of differentially abundant microbes. *mSystems* 2, doi:  
1358 10.1128/mSystems.00092-17 (2017).

1359  
1360 **Acknowledgements:** We thank Clemens Weiß (University of Tübingen), Fredrik Seersholm (University  
1361 of Copenhagen) and Pedro Lebre (University of Pretoria) for informative discussions and analytical  
1362 support.

1363  
1364 **Funding:** RFR acknowledges funding provided by a National Geographic Society Scientific Exploration  
1365 Grant (Nr. NGS-371R-18). GP, AV, RFR and JBR acknowledge funding provided by the French  
1366 Ministère des Affaires Étrangères and the French Institute of South Africa (IFAS) and RFR and JBR  
1367 acknowledge research mobility assistance provided by the NRF (South Africa) and DRIS (University of  
1368 Pretoria). We thank the South African heritage Resources Agency (SAHRA) for the provision of export  
1369 permits (Permit ID: 2364).

1370  
1371 **Author contributions:** RFR, AJH, JBR and SV conceived the study and composed the manuscript. SV  
1372 performed the bioinformatic and statistical analyses and RFR and SV generated the figures. JBR, ARI,  
1373 TBB and RFR performed aDNA extractions, library preparations and sequencing. GP and AV provided  
1374 access to the analysed specimen. GH performed the isotope analyses, and SW performed the <sup>14</sup>C dating.  
1375 MLB carried out intestinal parasitic analyses. MP and RFR performed SEM analyses. EW, DAC and  
1376 YVDP provided funding and access to analytical facilities. AJH, JBR and RFR conceived the analytical  
1377 protocol and AJH supervised the experimental protocol. All authors contributed to the completion of the  
1378 final manuscript.

1379  
1380 **Competing interests:** The authors declare that they have no competing interests. The funding sponsors  
1381 had no role in the design of the study, the collection, analyses and interpretation of data, in the writing of  
1382 the manuscript or in the decision to distribute the results.

1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398

**Table 1. DNA sequence reads for twenty-four authenticated commensal IM taxa detected in the BRS palaeo-faecal specimen.**

Phylum	Genus	E-LPC reads	BWA aligned (mapped) sequence reads									Mean read length (bp)	C-T p-values				
			BRS1 SC1	BRS2	BRS3	BRS4	BRS5 SC2	Total	Total %	BRS IM Total	BRS IM %		BRS1 SC1	BRS2	BRS3	BRS4	BRS5 SC2
<i>Euryarchaeota</i>	<i>Methanobrevibacter</i>	19	7	1952	1109	29147	69	32303	4.67	32377	4.68	70.60	-	0.193028900	0.073998830	<b>0.002632932</b>	0.567650800
<i>Actinobacteria</i>	<i>Cellulomonas</i>	0	65	3445	1626	54666	143	59945	8.67	60097	8.68	66.60	0.921500700	<b>0.002823314</b>	0.153993200	<b>0.006118848</b>	0.509736300
	<i>Slackia</i>	0	5	89	29	1343	8	1474	0.21	1482	0.21	61.80	<b>0.007399490</b>	0.935197800	0.311623400	0.508960500	-
<i>Bacteroidetes</i>	<i>Alistipes</i>	2	4	139	46	1069	15	1275	0.18	1290	0.18	65.40	0.675706100	0.420822200	0.053055550	<b>0.031112450</b>	0.606658100
	<i>Bacteroides</i>	5	114	28316	10115	115446	162	154158	22.30	154342	22.36	72.20	0.721165000	<b>0.001696204</b>	<b>0.000000390</b>	<b>0.000622655</b>	0.142918300
	<i>Flavobacterium</i>	32	58	6986	3125	33452	72	43725	6.33	43803	6.33	64.40	0.216522900	<b>0.000235190</b>	<b>0.000039500</b>	<b>0.022810540</b>	0.921151200
	<i>Pedobacter</i>	1	23	1656	715	9663	19	12077	1.75	12098	1.75	62.80	0.500000000	<b>0.001704170</b>	0.075530930	0.062391830	0.161709000
<i>Firmicutes</i>	<i>Prevotella</i>	17	48	418	184	3712	43	4422	0.64	4466	0.63	61.00	0.099233850	0.305674400	0.081571890	<b>0.016625350</b>	0.631959800
	<i>Blautia</i>	15	72	276	124	3640	111	4238	0.61	4350	0.59	68.20	0.608740100	0.084446270	0.428864300	<b>0.045772080</b>	0.245663300
	<i>Butyrivibrio</i>	0	8	325	127	3082	6	3548	0.51	3555	0.51	58.20	0.395839900	<b>0.037614750</b>	0.109365700	0.273060200	0.580047400
	<i>Clostridium</i>	4	148	2671	1304	29948	135	34210	4.95	34350	4.93	65.80	0.074740200	0.296300200	0.190844300	<b>0.015163840</b>	0.360285900
	<i>Enterococcus</i>	1	156	553	298	7427	20	8455	1.22	8476	1.20	74.60	<b>0.003765520</b>	0.333696100	<b>0.022915900</b>	0.274913800	-
	<i>Eubacterium</i>	0	40	697	296	9029	51	10113	1.46	10165	1.46	65.80	0.106194500	<b>0.004360991</b>	0.056443760	<b>0.024325710</b>	0.246347100
	<i>Lactobacillus</i>	0	143	0	259	6728	14	7144	1.03	7159	1.02	69.00	0.157217100	-	<b>0.011884140</b>	0.903384000	-
	<i>Lactococcus</i>	12	118	595	361	6219	14	7319	1.06	7334	1.04	69.60	0.463765800	<b>0.006836895</b>	<b>0.033206520</b>	<b>0.000000013</b>	-
	<i>Mogibacterium</i>	0	7	176	76	2116	5	2380	0.34	2385	0.34	68.60	0.499997800	0.228221200	<b>0.019403830</b>	<b>0.008329356</b>	0.188080100
	<i>Oribacterium</i>	1	12	170	72	2018	2	2275	0.33	2277	0.33	64.60	0.167621400	<b>0.050161290</b>	0.349725600	<b>0.000162173</b>	-
	<i>Oscillibacter</i>	0	7	165	91	1787	12	2062	0.30	2074	0.30	63.00	0.120518600	<b>0.009668940</b>	0.349725600	0.095004240	0.072742730
	<i>Ruminococcus</i>	18	42	1119	464	9229	100	10972	1.59	11074	1.57	67.40	0.271433200	<b>0.048583730</b>	0.738132500	<b>0.001960024</b>	0.300862800
	<i>Sarcina</i>	0	3	508	202	327	2	1042	0.15	1044	0.15	67.00	-	0.413740500	<b>0.018963150</b>	<b>0.025439260</b>	-
<i>Proteobacteria</i>	<i>Enterobacter</i>	0	145	41220	14798	181051	205	237419	34.34	237658	34.45	71.80	<b>0.004146302</b>	<b>0.000377506</b>	<b>0.000009640</b>	<b>0.001112472</b>	0.132896300
	<i>Klebsiella</i>	14	21	3485	1217	15633	72	20442	2.96	20517	2.96	70.00	0.529037900	<b>0.001153996</b>	<b>0.001264576</b>	<b>0.000436448</b>	0.921151200
	<i>Sphingobium</i>	0	36	1023	458	8330	84	9931	1.44	10016	1.43	65.80	0.126796900	0.132801100	<b>0.045433740</b>	<b>0.000753091</b>	0.602973600
	<i>Variovorax</i>	0	112	2099	878	16970	315	20374	2.95	20692	2.90	69.60	0.259295700	<b>0.005879342</b>	<b>0.057844450</b>	<b>0.020613870</b>	0.171048100
<b>Total</b>		<b>141</b>	<b>1394</b>	<b>98083</b>	<b>37974</b>	<b>6E+05</b>	<b>1679</b>	<b>691303</b>	<b>100.00</b>	<b>693082</b>	<b>100.00</b>	<b>66.83</b>	-	-	-	-	-

1400

1401

1402

1403

1404 **Table 2. DNA sequence reads for twelve authenticated pathogenic taxa detected in the BRS palaeo-faecal specimen.**

Phylum	Genus	Pathogenic clinical manifestations	<i>E-LPC</i> reads	BWA aligned (mapped) sequence reads						Mean read length (bp)	C-T <i>p</i> -value				
				BRS1 SC1	BRS2	BRS3	BRS4	BRS5 SC2	Total		BRS1 SC1	BRS2	BRS3	BRS4	BRS5 SC2
<i>Actinobacteria</i>	<i>Mycobacterium</i>	Tuberculosis, leprosy, atypical infections	701	397	4535	1853	59096	487	67069	67.20	0.0707802	<b>0.0017862</b>	<b>0.0037356</b>	<b>0.0100530</b>	0.2953064
<i>Bacteroidetes</i>	<i>Weeksella</i>	Pneumonia, sepsis, peritonitis, urinary tract infections	1	0	461	175	2552	0	3189	45.40	–	<b>0.0069370</b>	<b>0.0484237</b>	0.1960360	–
<i>Firmicutes</i>	<i>Listeria</i>	Listeriosis (convulsions, septicaemia and meningitis)	0	92	453	233	5989	5	6772	72.80	<b>0.0000005</b>	0.3494000	0.0975000	0.8800000	–
	<i>Aeromonas</i>	Gastroenteritis and wound infections	0	13	129	289	2033	29	2493	67.20	0.2264910	<b>0.0206110</b>	0.4139431	<b>0.0060921</b>	0.4998525
<i>Proteobacteria</i>	<i>Achromobacter</i>	Skin, soft tissue and respiratory infections	21	150	36089	12869	160087	311	209527	72.40	<b>0.0011408</b>	<b>0.0017379</b>	<b>0.0000094</b>	<b>0.0015504</b>	0.1190260
	<i>Bordetella</i>	Pertussis (whooping cough, respiratory infection)	18	123	2027	683	16466	196	19513	69.60	<b>0.0204611</b>	0.1325985	<b>0.0082522</b>	<b>0.0077918</b>	0.2557714
	<i>Burkholderia</i>	Pulmonary infections (pneumonia, melioidosis)	12	48	869	287	6965	156	8337	62.00	0.2711432	<b>0.0462580</b>	0.2140947	0.3505477	0.1594083
	<i>Comamonas</i>	Acute appendicitis, urinary tract infection	8	79	3656	1467	40820	987	47017	73.40	0.6379285	<b>0.0059303</b>	<b>0.0225115</b>	<b>0.0322217</b>	0.6381353
	<i>Delftia</i>	Endocarditis, skin and ocular infections	226	594	921	348	10552	692	13333	74.80	0.2733595	<b>0.0035008</b>	0.5813131	0.0712462	0.0567663
	<i>Lysobacter</i>	Opportunistic pathogen, various symptoms	5	37	9616	3170	45813	95	58736	66.20	0.2478419	<b>0.0013587</b>	<b>0.0003318</b>	<b>0.0064930</b>	0.8210803
	<i>Shigella</i>	Shigellosis (dysentery, seizures, mucosal ulceration)	2	253	35046	12402	137623	762	186088	76.80	0.5381291	<b>0.0002471</b>	<b>0.0000002</b>	<b>0.0002182</b>	0.2954809
	<i>Vibrio</i>	Cholera (gastroenteritis, diarrhoea, septicaemia)	0	7	419	147	2333	21	2927	62.80	–	0.4420414	0.1131828	<b>0.0005438</b>	–
<b>Total</b>			<b>994</b>	<b>1793</b>	<b>94221</b>	<b>33923</b>	<b>490329</b>	<b>3741</b>	<b>625001</b>	<b>67.55</b>	–	–	–	–	–

1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421

1422 SUPPLEMENTARY MATERIALS

1423

1424 Supplementary Figures

1425

1426

1427

1428

1429

1430

1431

1432

1433

1434

1435

1436

1437

1438

1439

1440

1441

1442

1443

1444

1445

1446

1447

1448

1449

1450

1451

1452

1453

1454

1455

1456

1457

1458

1459

1460

1461

1462

1463

1464

1465

1466

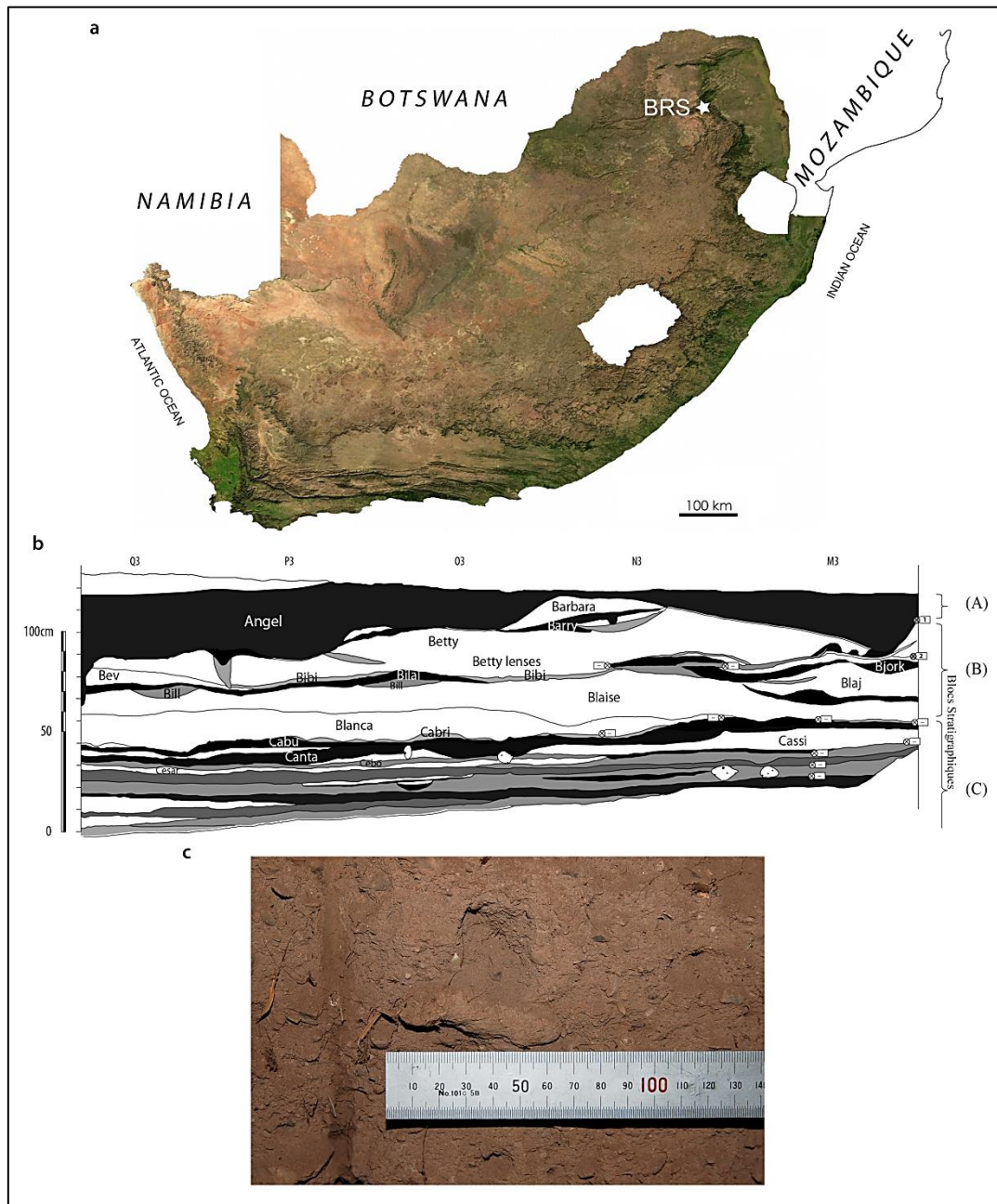
1467

1468

1469

1470

1471



**Fig. S1.** The provenience of the BRS palaeo-faecal specimen. The location of Bushman Rock Shelter (BRS) in Limpopo Province, South Africa (A), stratigraphic profile of the Iron Age occupation level from which the specimen derives and which comprise the upper layer of the rock-shelter (Layer 1) situated in excavation 'Block A' and designated 'Angel' (B) and the palaeo-faecal specimen *in situ* in the exposed (excavated) section prior to removal from the deposit (C).



1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523



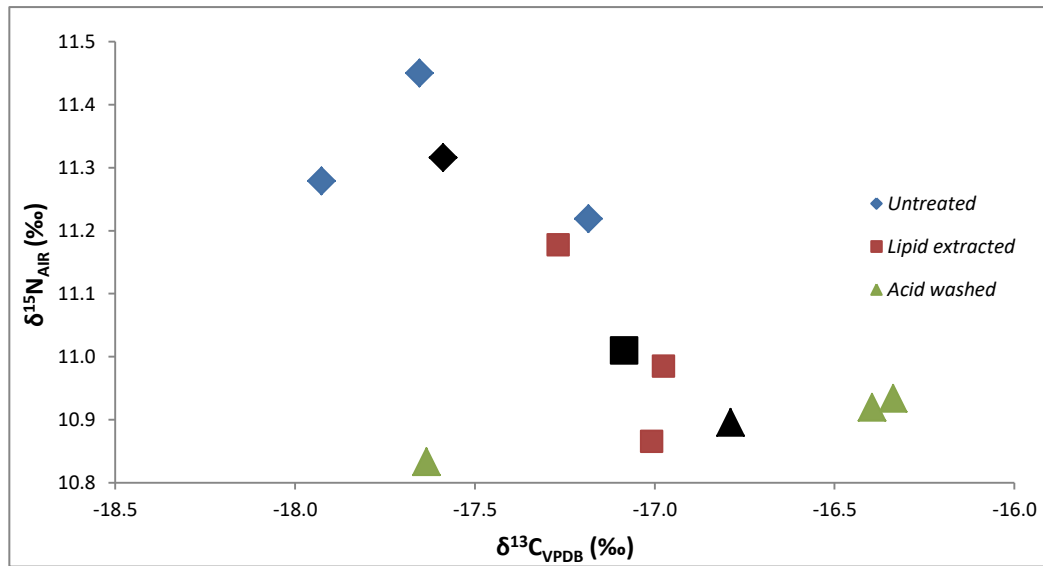
**Fig. S2.** Processing (sub-sampling and aDNA extraction) protocol applied to the BRS palaeo-faecal specimen. The frozen specimen was first removed from the sealed packaging (A), after which the outer surface or cortex (~5mm) was removed with a scalpel (B) and subsequently sub-sampled for radiocarbon ( $^{14}\text{C}$ ) dating and isotopic and microscopic intestinal parasitic analyses, preserving one-sixth of the specimen (at  $-20^{\circ}\text{C}$ ) as a voucher sample (C). Sub-sampling and aDNA extraction and library preparation was performed in the ‘clean’ ancient DNA laboratories at the Centre for GeoGenetics, University of Copenhagen (Denmark) (D).



1567 **Fig. S3.** Dot-plot based on the alignment of high-quality DNA sequence reads indicating the occurrence of  
 1568 statistically-significant C-T  $p$ -values calculated for commensal and pathogenic taxa detected in the BRS specimen  
 1569 (BRS2, BRS3 and BRS4) and the sedimentary controls (BRS1 or 'SC1' and BRS5 or 'SC2'). Circle sizes and  
 1570 colours represent mapped read counts and  $p$ -value significance, respectively (see legend and scale).

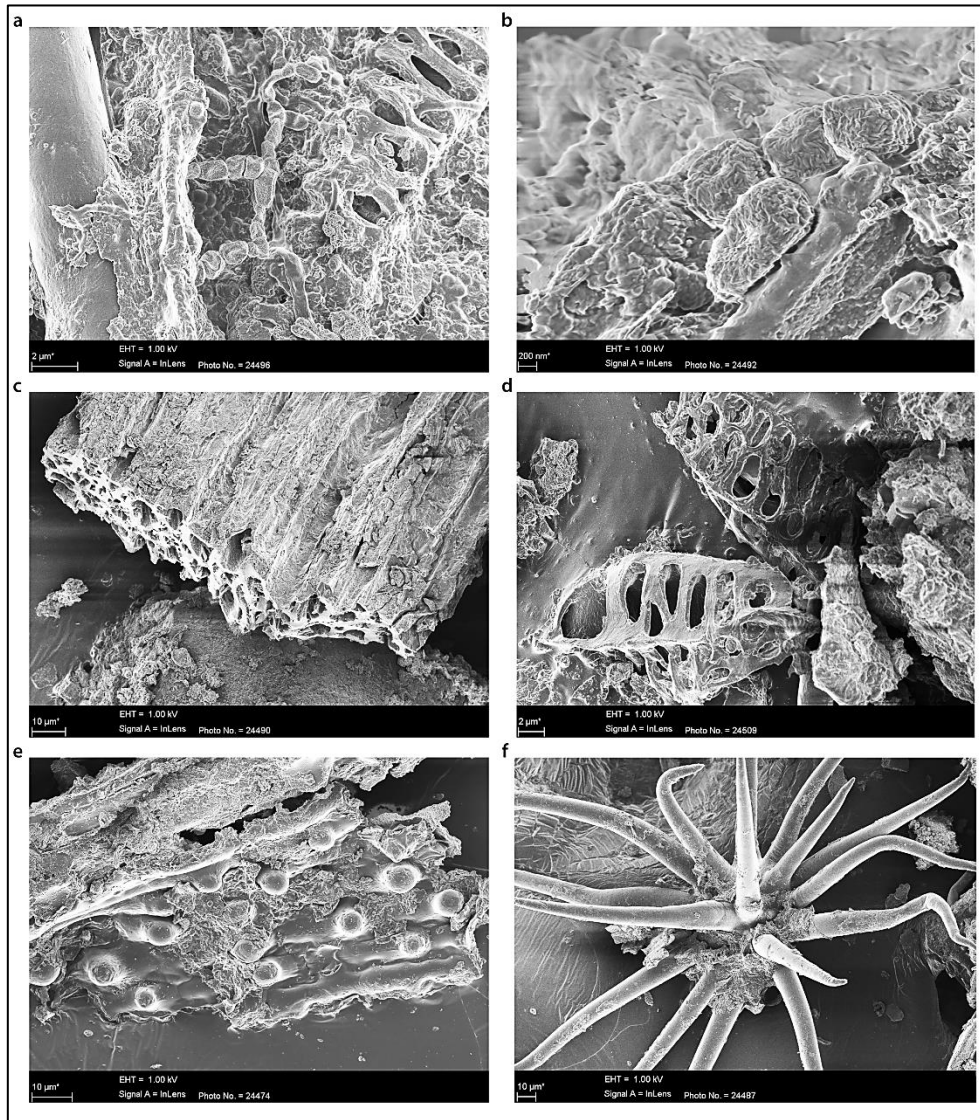
1571  
 1572  
 1573  
 1574  
 1575

1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627

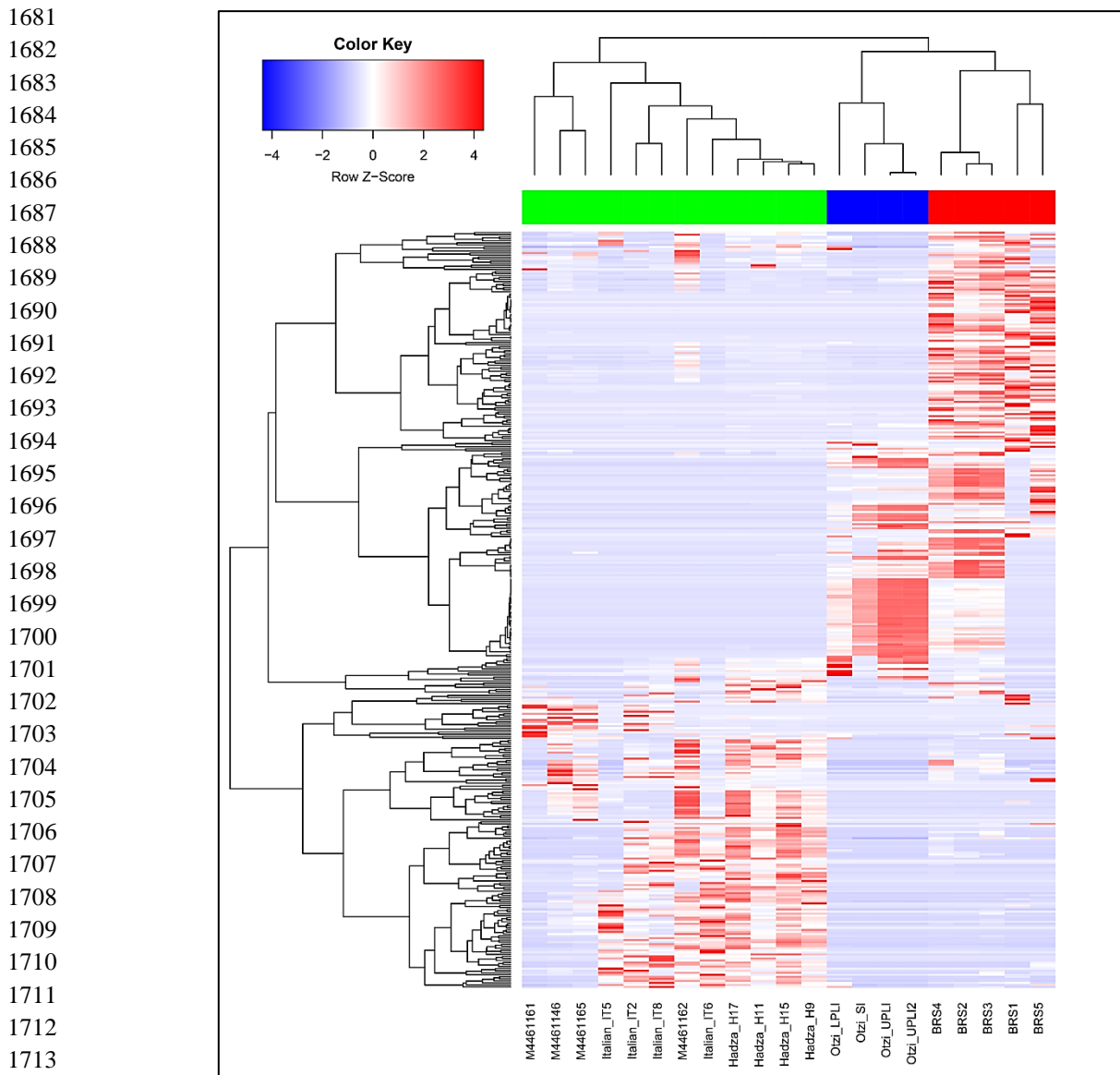


**Fig. S4.** Biplot of  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  stable isotope ratios obtained for the BRS specimen. The various pre-treatments (*i.e.*, the acid wash and lipid extraction) had some effect on both  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  values. The lipid extraction (2:1 chloroform/ethanol) removed traces of lipids with carbon isotope values becoming less negative, but nevertheless suggesting a mixed C4- (mostly) and C3-based meal. The results for the solvent residue reflect a geological signal most likely from the shelter's sediments. Average values for 'untreated', 'lipid extracted' and 'acid washed' are indicated in corresponding black markers.

1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680



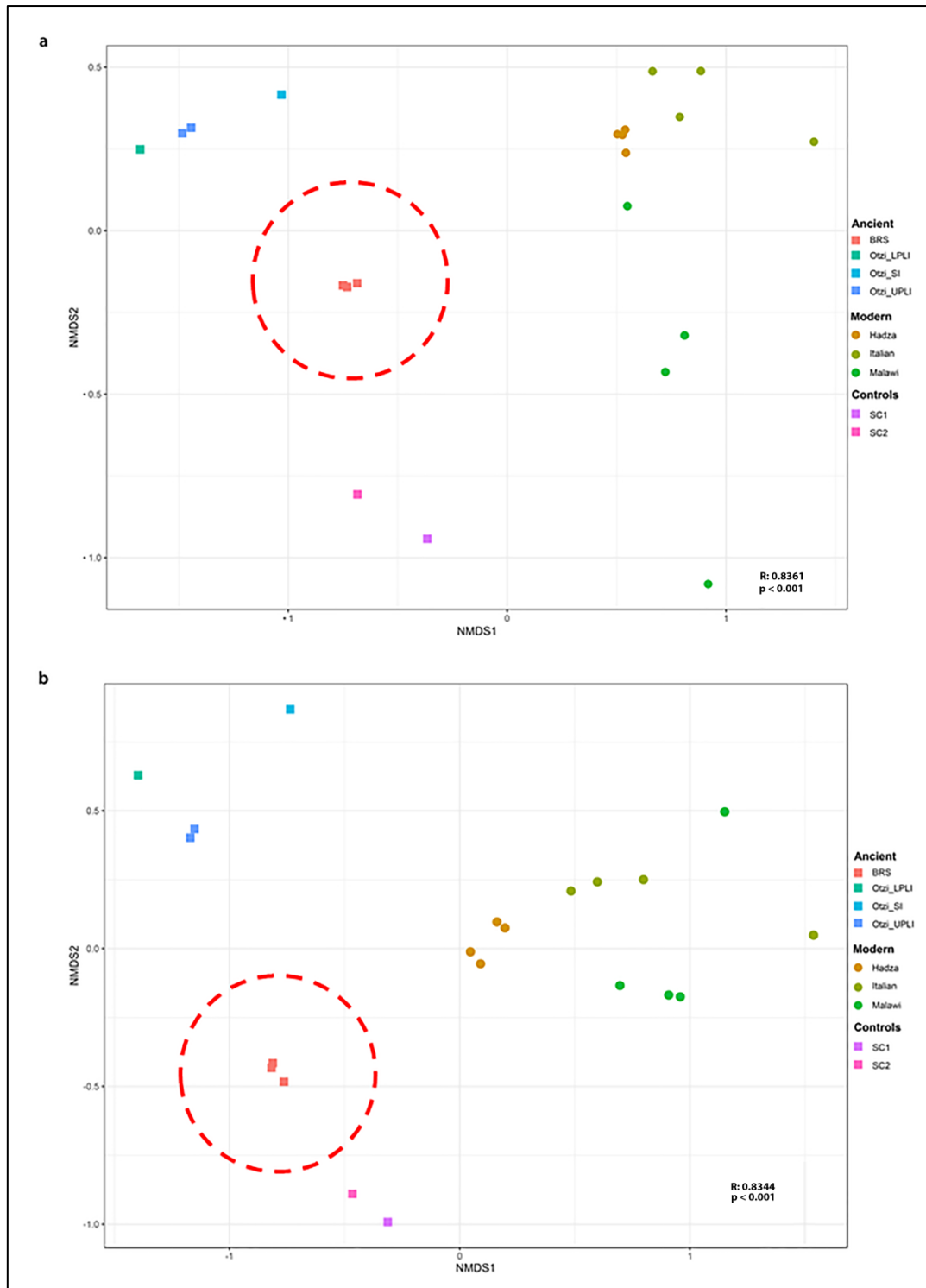
**Fig. S5.** Although scanning electron microscopy (SEM) analyses did not result in the detection of parasitic remains (e.g., eggs), it did result in the detection of desiccated bacterial cells (**A** and **B**), degraded plant fragments (**C**, **D** and **E**) and unidentified saprophytic organisms (**F**).



1715 **Fig. S6.** Heat-map comparing differences in general taxonomic community structure for BRS (1, 2, 3, 4 and 5) with  
1716 the modern (Italian), ethnographic (Hadza and Malawian) and ancient (Ötzi) IM datasets. Hierarchical clustering  
1717 using complete linkage based on Spearman's correlation, produced a clear separation between ancient (BRS and  
1718 Ötzi) and modern (Italian, Hadza and Malawian) populations. ANOSIM analysis revealed significant differences  
1719 between the ancient and modern IM samples ( $R = 0.6111$ ;  $p = < 0.05$ ) in the taxonomic categories of 731 taxa. Taxa  
1720 were filtered for occurrence of  $>3$  in at least 20% of the samples.

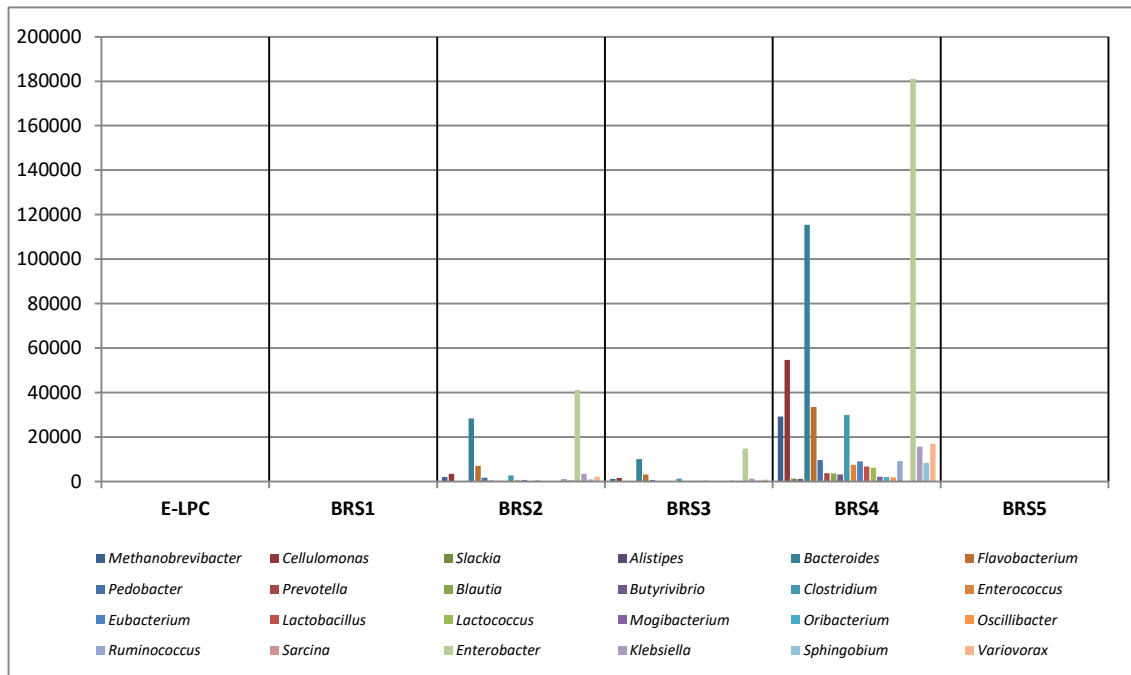
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732

1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784



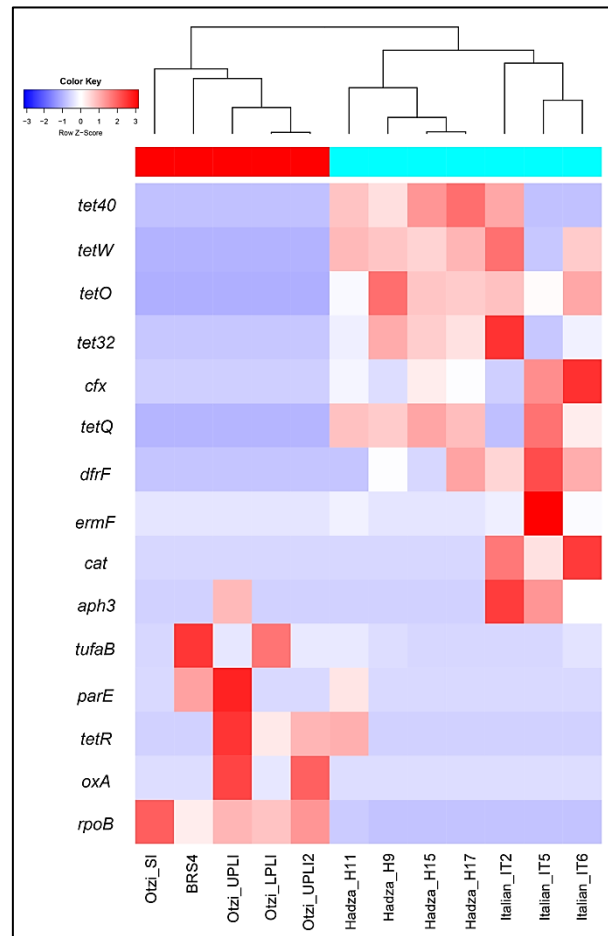
**Fig. S7.** Weighted (A) and un-weighted (B) Bray-Curtis non-metric multi-dimensional scaling (NMDS) plots comparing the use of ‘relative abundance’ (indicated in A) and ‘presence-absence’ data (indicated in B) as measures of taxonomic representation in the BRS specimen and in the modern (Italian), ethnographic (Hadza and Malawian) and ancient (Ötzi) IM datasets. Weighted (A) (as shown Fig. 1C and based on the ‘relative abundance’ of identified IM taxa) (ANOSIM  $R = 0.6111$ ;  $p < 0.001$ ) and un-weighted (B) Bray-Curtis analysis (based on ‘presence-absence’ of identified IM taxa) exhibit corresponding differences between the ancient and modern IM cohorts (ANOSIM  $R = 0.8166$ ;  $p < 0.001$ ).

1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836



**Fig. S8.** Bar-plot demonstrating that the surrounding sedimentary matrix (*i.e.*, samples BRS1 ‘SC1’ and BRS5 ‘SC2’) and the DNA extraction ( $n = 1$ ) and library preparation ( $n = 1$ ) negative controls (‘E-LPCs’) are not significant sources of the twenty-four authenticated ancient microbial taxa identified in the palaeo-faecal specimen (*i.e.*, BRS2, BRS3 and BRS4). Reverse contamination, *i.e.*, from the palaeo-faecal specimen into the surrounding sediment, is most likely responsible for the incidence of the very low numbers of reads for IM-specific taxa in the surrounding sedimentary matrix (*i.e.*, BRS1 and BRS5).

1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888



**Fig. S9.** Heat-map indicating the presence of fifteen functional ARGs identified in the analysed datasets, four of which occurs in the BRS IM, including the prokaryotic protein synthesis elongation factor Tu (EF-Tu) (*tufA* and *tufB*), fluoro-quinolone-resistant DNA topoisomerase (*parE*) and daptomycin-resistant *rpoB*. ARG categories were filtered for occurrence of >5 in at least 20% of samples.



1889 **Supplementary Tables**

1890

1891 **Table S1.** Mapped DNA sequence reads for environmental- and subsistence-related taxa detected in this study.  
1892 Statistically-significant (*i.e.*, verified ancient) C-T *p*-values are indicated in bold text. Analyses were performed  
1893 using high-quality filtered read alignments against NCBI reference genomes. DNA damage estimation analyses was  
1894 performed using PMDtools ('C-T *p*-values').

1895

1896 **Table S2.** Information concerning the two direct radiocarbon (<sup>14</sup>C) Accelerator Mass Spectrometry (AMS) dates  
1897 generated from two sub-samples taken from within the BRS palaeo-faecal specimen.

1898

1899 **Table S3a.** Processing protocol and results for isotope analyses of samples derived from the BRS specimen  
1900 indicating the relative proportions of C3 and C4 dietary contributions.

1901

1902 **Table S3b.** Results for isotope analyses (Merck standard) of samples derived from the BRS specimen indicating the  
1903 relative proportions of C3 and C4 dietary contributions.

1904

1905 **Table S3c.** Results for isotope analyses (DL-Valine standard) of samples derived from the BRS specimen  
1906 indicating the relative proportions of C3 and C4 dietary contributions.

1907

1908 **Table S4.** Assignment of abundance of bacterial taxonomic categories to the BRS and ancient (Ötzi) and modern  
1909 (Italian, Hadza and Malawian) comparative samples based on *p*-value (*p* = <0.05) designation. Group significance  
1910 assignments were obtained via Qiime v1.9.1.

1911

1912 **Table S5.** Information concerning DNA sequence read-length distribution for taxa identified in this study. Read-  
1913 length distributions were calculated from BWA alignments and BAM files.

1914

1915 **Table S6.** Relative abundance of eighteen significant KEGG pathways detected in BRS and the ancient (Ötzi) and  
1916 modern (Italian, Hadza and Malawian) comparative IM cohorts.

1917

1918 **Table S7.** Enrichment and depletion of KO metabolic gene categories in the BRS and ancient (Ötzi) and modern  
1919 (Italian, Hadza and Malawian) comparative IM sample cohorts based on *p*-value (*p* = <0.05) designation. Group  
1920 significance analyses were performed using Qiime v1.9.1.

1921

1922 **Table S8.** Enrichment and depletion of KO metabolic gene categories in the ancient (BRS and Ötzi) and modern  
1923 (Italian, Hadza and Malawian) comparative IM sample cohorts based on false discovery rate (FDR) corrected *p*-  
1924 values (*q* = <0.05).

1925

1926 **Table S9.** Enrichment and depletion of KO metabolic gene categories in the ancient and modern comparative IM  
1927 sample cohort as calculated for the twenty-four authenticated ancient IM taxa.

1928

1929 **Table S10.** Comparison of relative abundance of antibiotic resistance genes in the BRS and the ancient (Ötzi) and  
1930 modern (Italian, Hadza and Malawian) comparative IM cohorts.

1931

1932 **Table S11.** Raw and filtered high-quality sequence read counts as related to the BRS and the ancient (Ötzi) and  
1933 modern (Italian, Hadza and Malawian) comparative IM datasets.

1934

1935 **Table S12.** Information concerning the comparative NCBI genomes used during this study.