

## Supplemental Material For:

### Analysis of medieval burials from Ibiza reveals genetic and pathogenic diversity during the Islamic period

Ricardo Rodríguez-Varela<sup>1,2\*</sup>; Zoé Pochon<sup>1,2</sup>; Alex Mas-Sandoval<sup>3</sup>; Reyhan Yaka<sup>1,2</sup>; Cesar A. Fortes-Lima<sup>4,5</sup>; Almudena García Rubio<sup>6</sup>; Nicholas Márquez-Grant<sup>7,8</sup>; Juanjo Mari<sup>9</sup>; Glenda Graziani<sup>10</sup>; Antoni Ferrer Abárzuza<sup>11</sup>; Mário Vicente<sup>1,2</sup>; Lander Lorca-Francisco<sup>1</sup>; Anna Linderholm<sup>1,12</sup>; Vendela K. Lagerholm<sup>1,2</sup>; Lara R. Arauna<sup>13</sup>; Patxi Pérez-Ramallo<sup>14,15,16</sup>; Maja Krzewińska<sup>1,2</sup>; Carina M. Schlebusch<sup>17,18,5</sup> Anders Götherström<sup>1,2,\*</sup>

#### Affiliations

<sup>1</sup> Centre for Palaeogenetics, Stockholm, Sweden

<sup>2</sup> Department of Archaeology and Classical Studies, Stockholm University, Sweden

<sup>3</sup> Departamento de Historia, Geografía y Comunicación, Universidad de Burgos, Spain

<sup>4</sup> McKusick-Nathans Institute and Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

<sup>5</sup> Human Evolution, Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

<sup>6</sup> Aranzadi Science Society, San Sebastián, Spain

<sup>7</sup> Cranfield Forensic Institute, Cranfield University, UK

<sup>8</sup> School of Anthropology and Museum Ethnography, University of Oxford, UK

<sup>9</sup> Independent researcher, Ibiza, Spain

<sup>10</sup> Departament de Ciències de l'Antiguitat i Edat Mitjana. Universitat Autònoma de Barcelona, Spain

<sup>11</sup> Consell Insular de Formentera, Balearic Islands, Spain

<sup>12</sup> Department of Geological Sciences, Stockholm University, Stockholm, 10691, Sweden

<sup>13</sup> Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals (BEECA), Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

<sup>14</sup> Social History of Capitalism research group (SHOC), Vrije Universiteit Brussel (VUB), Brussels, Belgium

<sup>15</sup> Archaeology, Environmental Changes & Geo-Chemistry research group (AMGC), Vrije Universiteit Brussel (VUB), Brussels, Belgium

<sup>16</sup> Department of Coevolution of Land Use and Urbanisation, Max Planck Institute of Geoanthropology, Jena, Germany

<sup>17</sup> Palaeo-Research Institute, University of Johannesburg, Johannesburg, South Africa

<sup>18</sup> Center for the Human Past, Department of Organismal Biology, Uppsala, Sweden

\*Corresponding authors (ricardo.rodriguez.varela@arklab.su.se; anders.gotherstrom@arklab.su.se)

## Supplementary Discussion

### Radiocarbon dates and Marine reservoir

To assess potential marine offsets, we estimated marine dietary protein for all twelve ( $n = 12$ ) directly dated individuals using the Bayesian mixing package *simmr* (R 4.5.0) <sup>1</sup>. Mixtures were the dentine-collagen  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  values for each individual; source baselines followed our previous approach: terrestrial herbivores ( $n = 32$ ) from Gandía late-medieval herbivores <sup>2</sup>, Zorita rabbits <sup>3</sup>, and three medieval Ibiza fauna <sup>4</sup>;  $\delta^{13}\text{C} = -18.3 \pm 2.2\text{‰}$ ,  $\delta^{15}\text{N} = 6.3 \pm 2.3\text{‰}$ ) and a well-dated Mediterranean fish assemblage from Provence ( $n = 127$ ;  $\delta^{13}\text{C} = -9.5 \pm 2.7\text{‰}$ ,  $\delta^{15}\text{N} = 10.6 \pm 2.1\text{‰}$ ). Across individuals, the marine-fish fraction ranges 17.7–34.7 % (mean across individuals 21.8, SD 1.4; median 20.0). We report per-individual posterior summaries in Supplementary Data 2.

All radiocarbon determinations were then calibrated in OxCal v4.4.4 <sup>5</sup> under two models per individual: a purely terrestrial IntCal20 calibration <sup>6</sup> and, a mixed IntCal20–Marine20 calibration that incorporates each person's *simmr* *p*marine (mean  $\pm$  SD) together with a regional marine offset  $\Delta R = -42 \pm 30$  14C yr from Banyuls <sup>7</sup>. Complete 68.3% and 95.4% Highest Posterior Densities (HPD) for both models are given in Supplementary Data 2 (see also Supplementary Fig. 1).

Under the mixed model, s.197 is robustly post-1115 CE (HPD68.3% = 1165–1280 CE), whereas s.117 straddles the 1115 threshold (HPD68.3 = 1051–1264 CE): compatible with, but not conclusive for, a post-1115 placement. Most other individuals likewise straddle 1115 at 68.3%, and s.313 remains earlier (HPD68.3% = 896–1046 CE). Because the dated material is permanent molar dentine, formed decades before death in mature adulthood, dates of death would be expected to fall later than the dentine formation windows, which further supports a post-1115 placement for s.197 and increases the plausibility for s.117, while we acknowledge remaining uncertainty at 95.4%. Limitations include non-fully local/coeval fauna, wide diet posteriors for some individuals, and the use of dentine (formation  $\neq$  death); we handle these by reporting both terrestrial and mixed calibrations and by framing all results probabilistically (HPDs). These results are overall consistent with the independent archaeological span proposed for the cemetery (ca. 950–1150 CE) <sup>4</sup>, with one clearly early burial (s.313) and one clearly late burial (s.197), while the remainder fall broadly within the 11th–12th centuries when marine-diet uncertainty is considered.

### Metagenomics

#### Investigation of specific pathogens with scores lower than 8

Individual s.197 yielded a *Mycobacterium leprae* hit with an authentication score of 7 in one of six sequencing runs of the same library. Upon reviewing the KrakenUniq outputs from each run (Supplementary Data 7), we suspect that although all libraries were double-indexed, this signal likely results from index hopping. The affected run was sequenced at higher depth and included the capture-enriched library of sample s.313, which harboured a confirmed *M. leprae* infection. Another individual from the same run, despite very low coverage, also showed *M. leprae* reads, further supporting the likelihood of low-level cross-contamination.

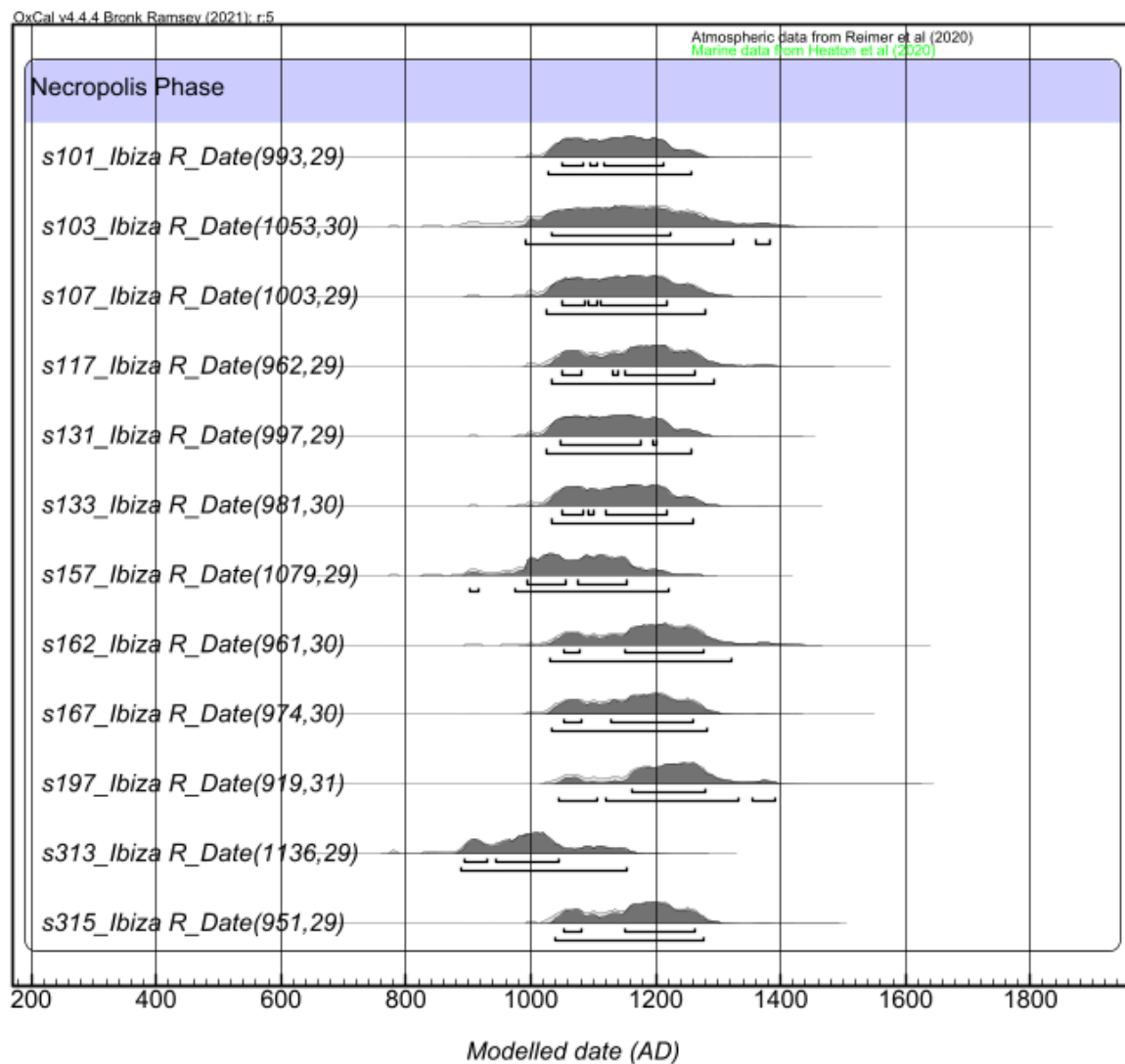
Only one library was prepared for s.197 and sequenced across six lanes. Notably, only the lane containing the s.313 capture-enriched sample showed detectable *M. leprae* reads in s.197, as well as in other unrelated, low-coverage samples sequenced in the same lane. The contamination appears to be limited in the case of s.197, involving approximately 1,000 to 2,000 *M. leprae* reads.

This case highlights the importance of exercising caution when pooling capture-enriched and shotgun libraries for sequencing, particularly in studies focused on low-abundance pathogen detection.

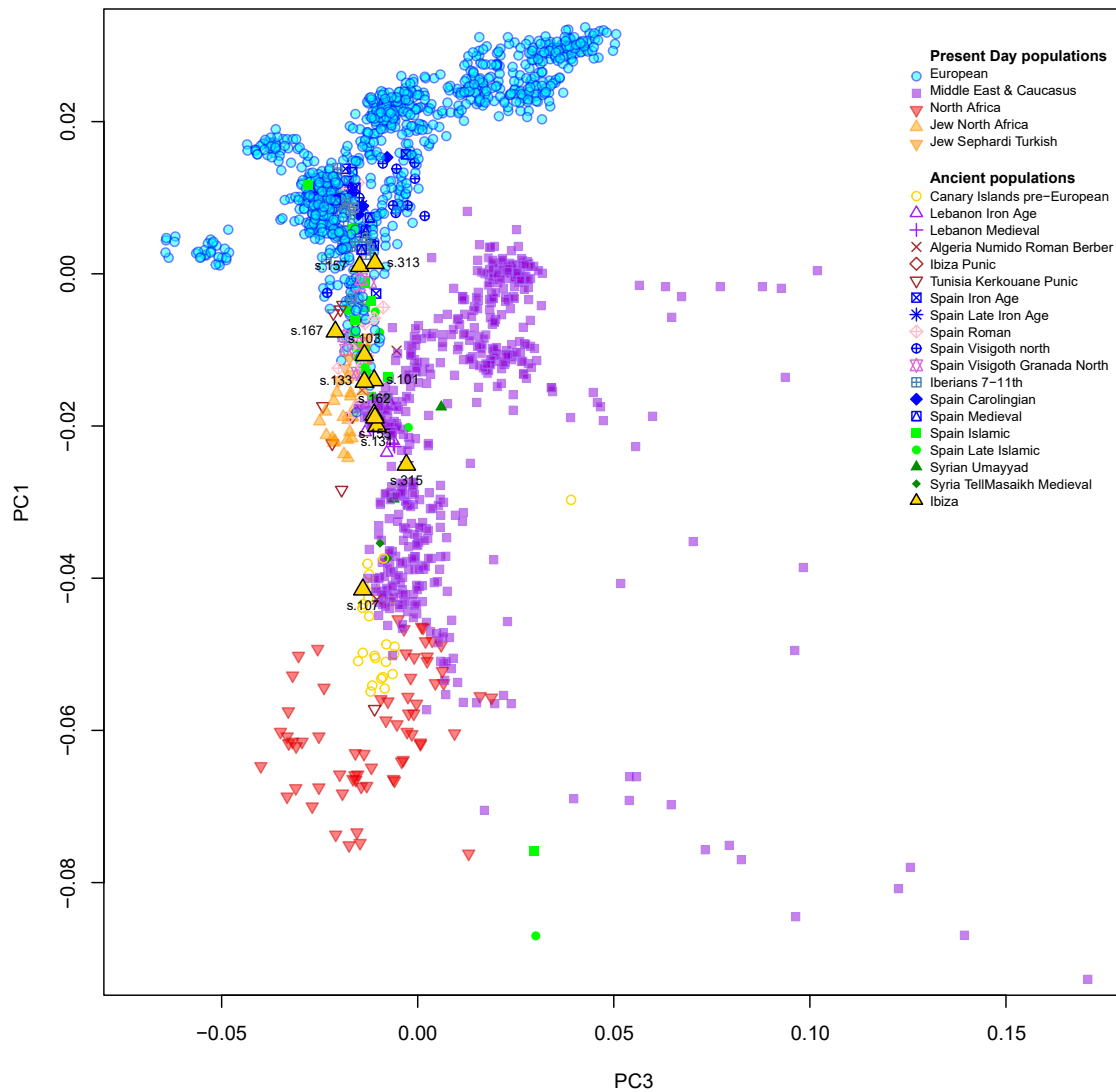
#### *Mycobacterium leprae* reference genome dataset

The dataset of all published ancient *Mycobacterium leprae* genomes was extracted based on the AncientMetagenomeDir database using AMDirT<sup>8,9</sup>. Ancient genomes of 42 ancient samples with a depth of  $\geq 3$  were included in the dataset<sup>10–17</sup>. Similarly, modern genomes of 139 samples with a depth of  $\geq 3$  were included in the dataset<sup>18–24</sup>.

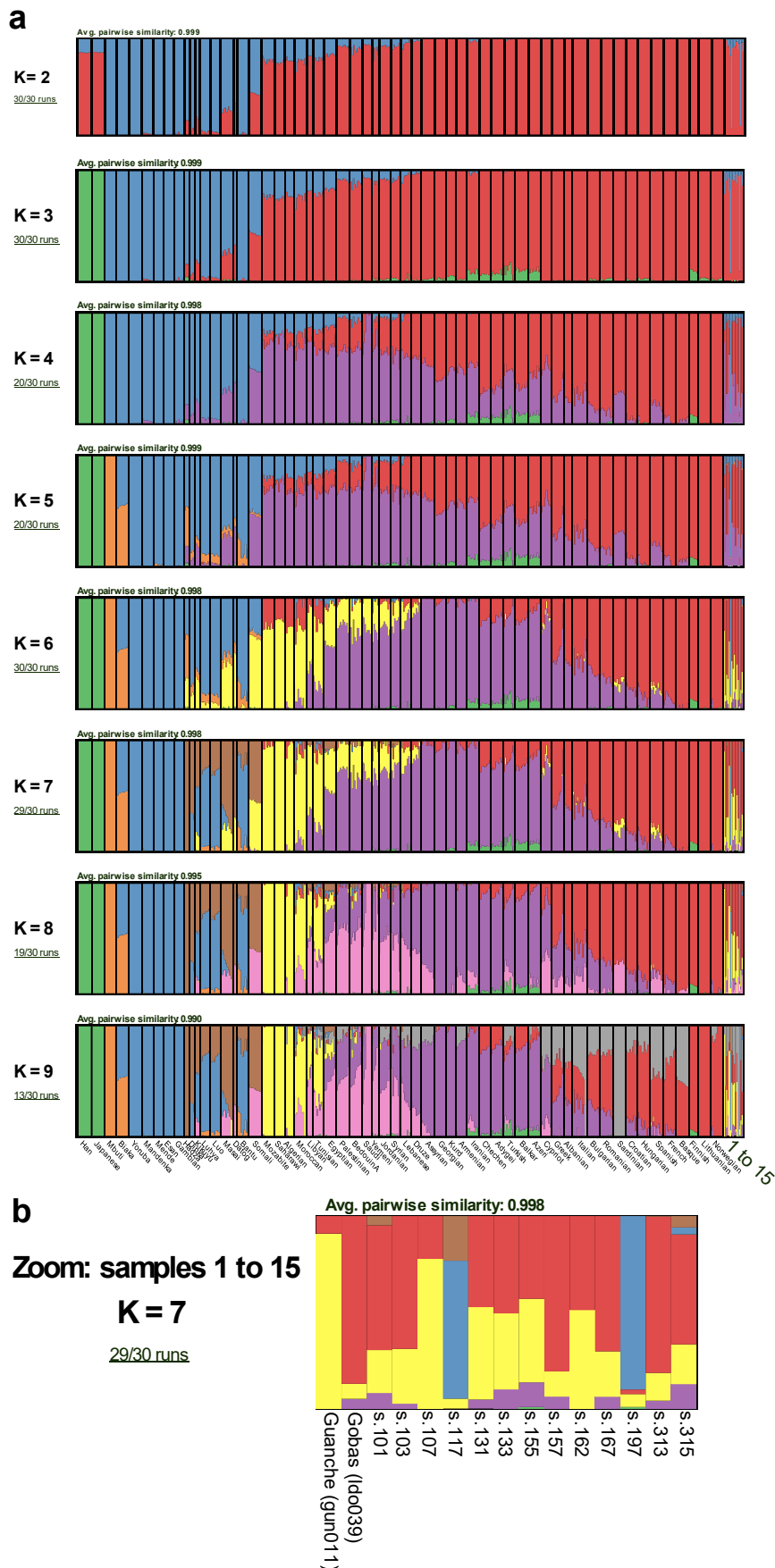
### Supplementary Figures



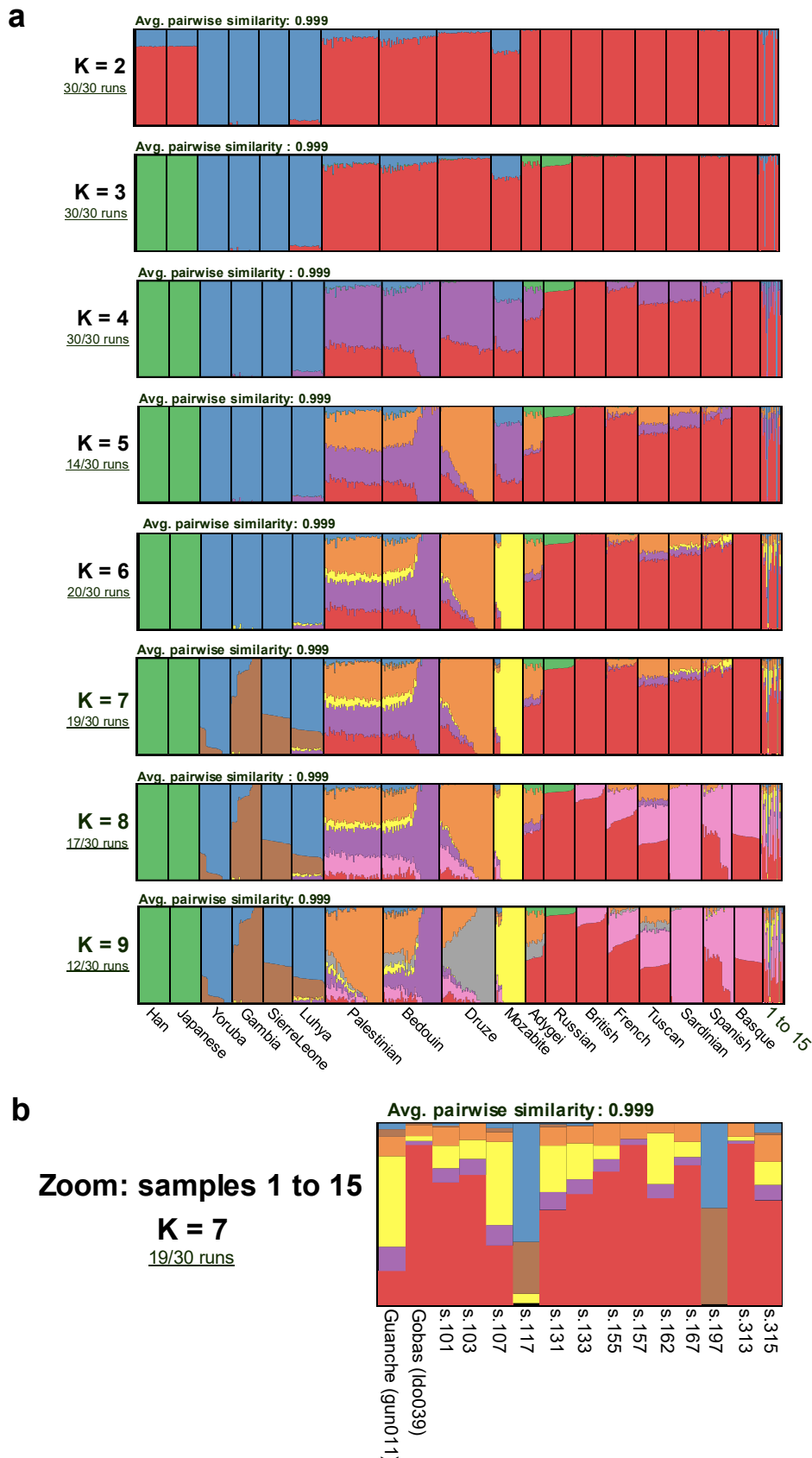
**Supplementary Fig. 1. Calibrated 14C probability distributions for all dated individuals.** Black bars mark 95.4 % credible intervals; shaded areas indicate posterior density.



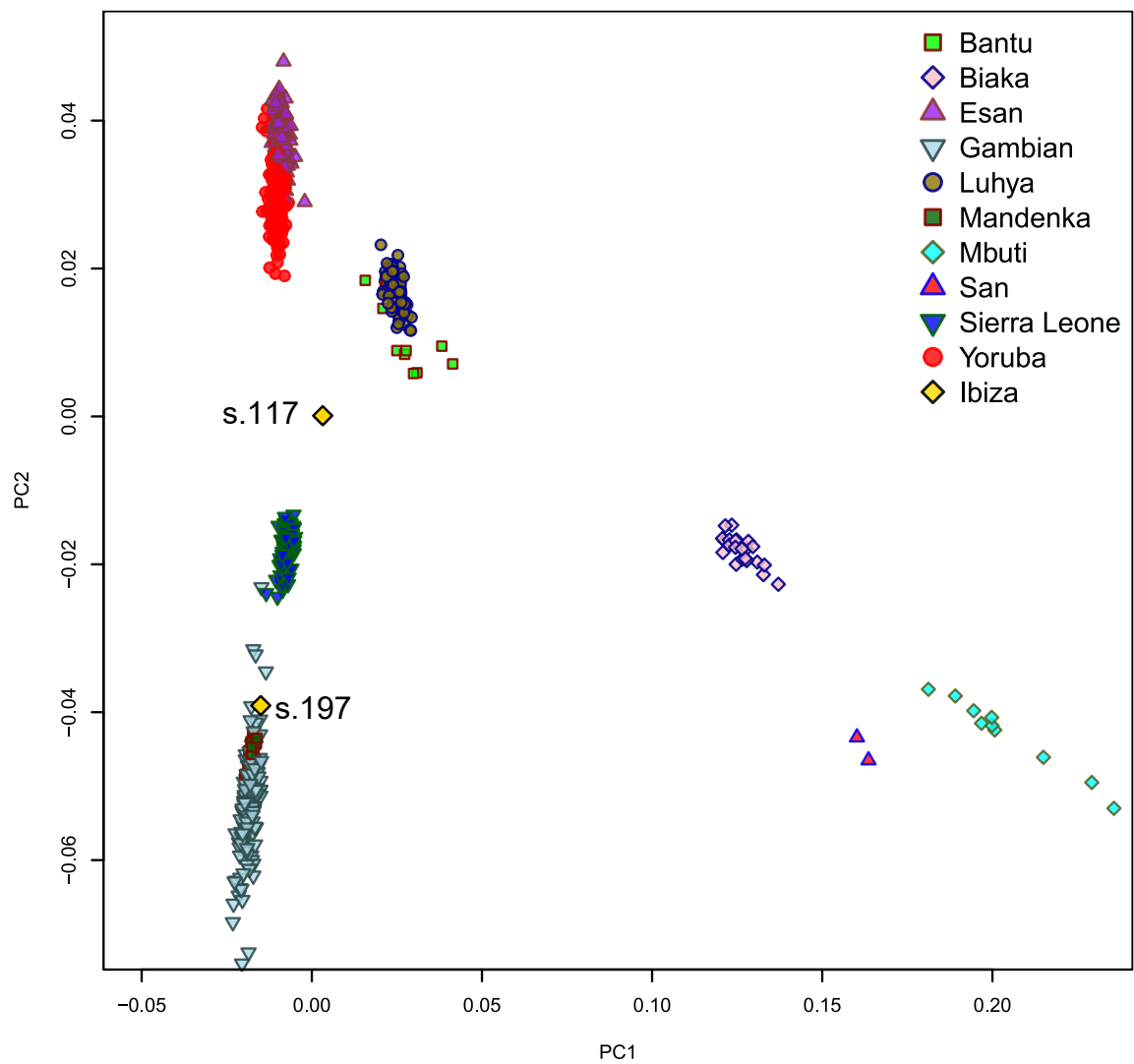
**Supplementary Fig. 2. Principal Component Analysis (PCA) of Eurasian populations with ancient DNA individuals projected.** Ancient individuals (excluding the two Sub-Saharan African samples) are projected onto PC1 and PC3, based on modern Eurasian genetic variation from the Human Origins (HO) dataset. Ancient genotypes are pseudo-haploid.



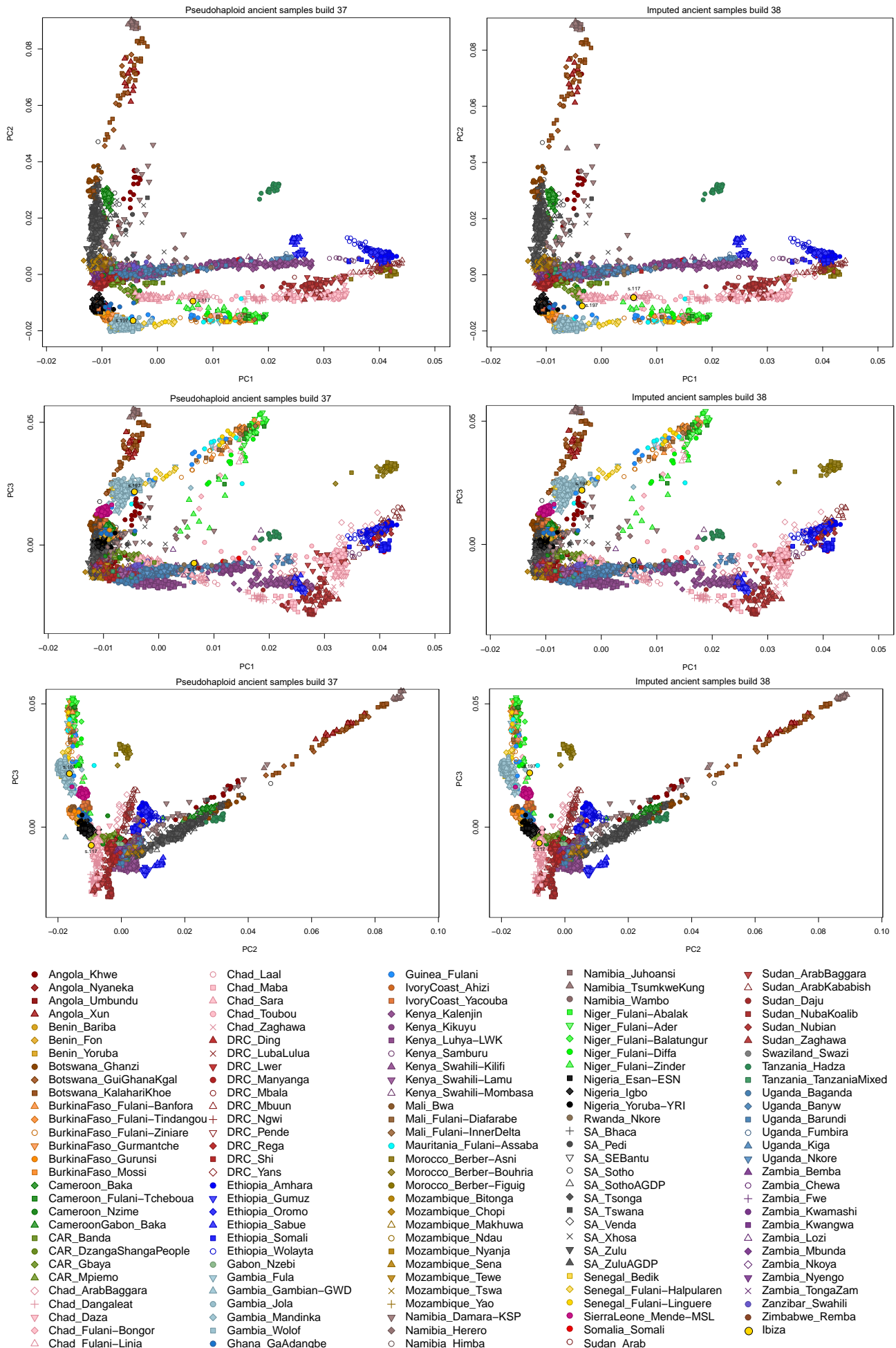
**Supplementary Fig. 3. Unsupervised ADMIXTURE Human Origins pseudohaploid dataset. a.** K=2 to k=9. 1 to 15: Guanche (gun011); Gobas (ldo039); s.101; s.103; s.107; s.117; s.131; s.133; s.155; s.157; s.162; s.167; s.197; s.313; s.315. **b.** Inset: zoom-in on the ancient individuals at K = 7.



**Supplementary Fig. 4. Unsupervised ADMIXTURE. 1000 Genomes Project plus HDGP and imputed diploid dataset. a.** K=2 to K=9. 1 to 15: Guanche (gun011); Gobas (ldo039); s.101; s.103; s.107; s.117; s.131; s.133; s.155; s.157; s.162; s.167; s.197; s.313; s.315. **b.** Inset: zoom-in on the ancient individuals at K = 7.

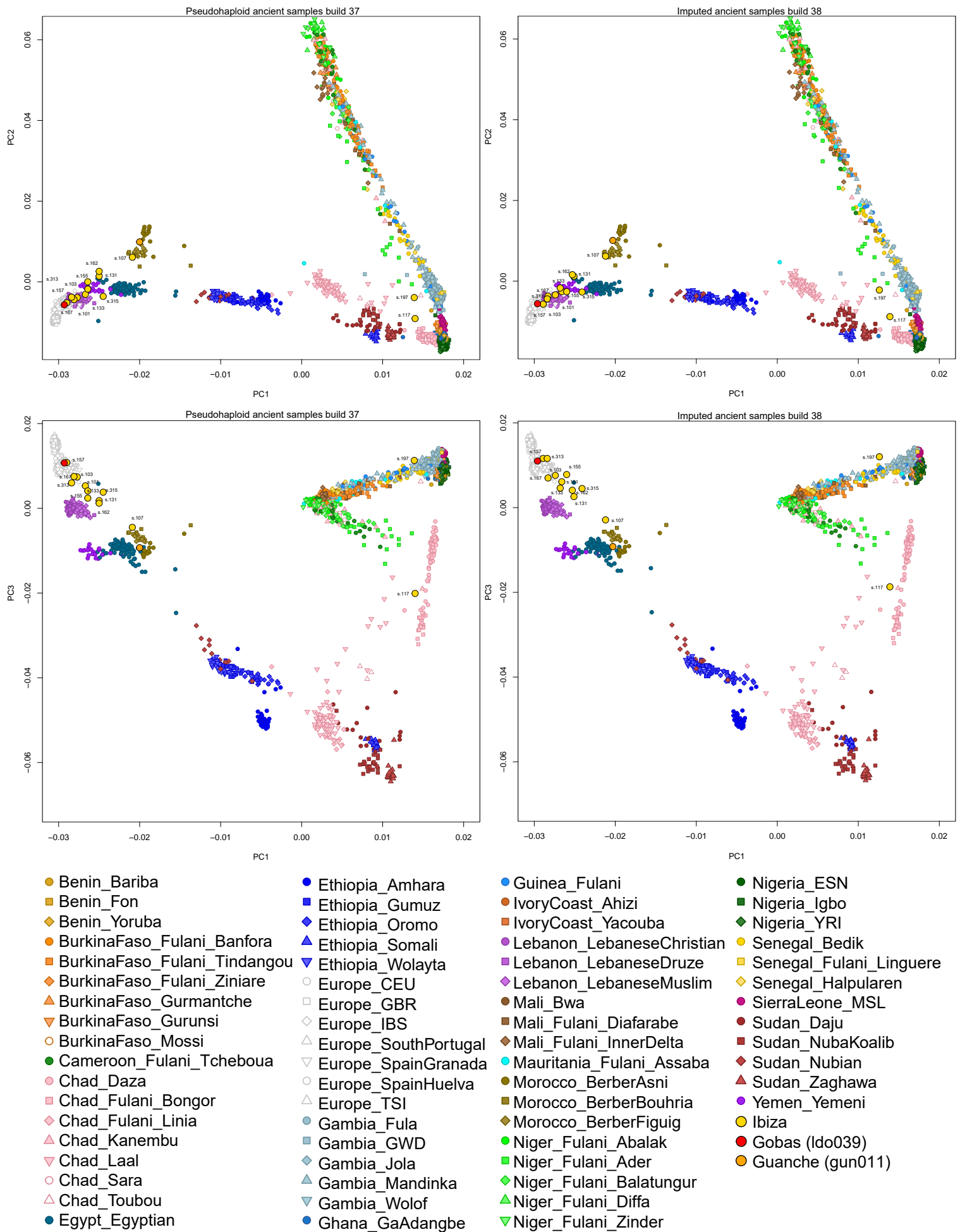


**Supplementary Fig. 5. First two principal components of the two ancient Sub-Saharan individuals projected onto modern sub-Saharan Africans included in the 1K-HGDP dataset.**

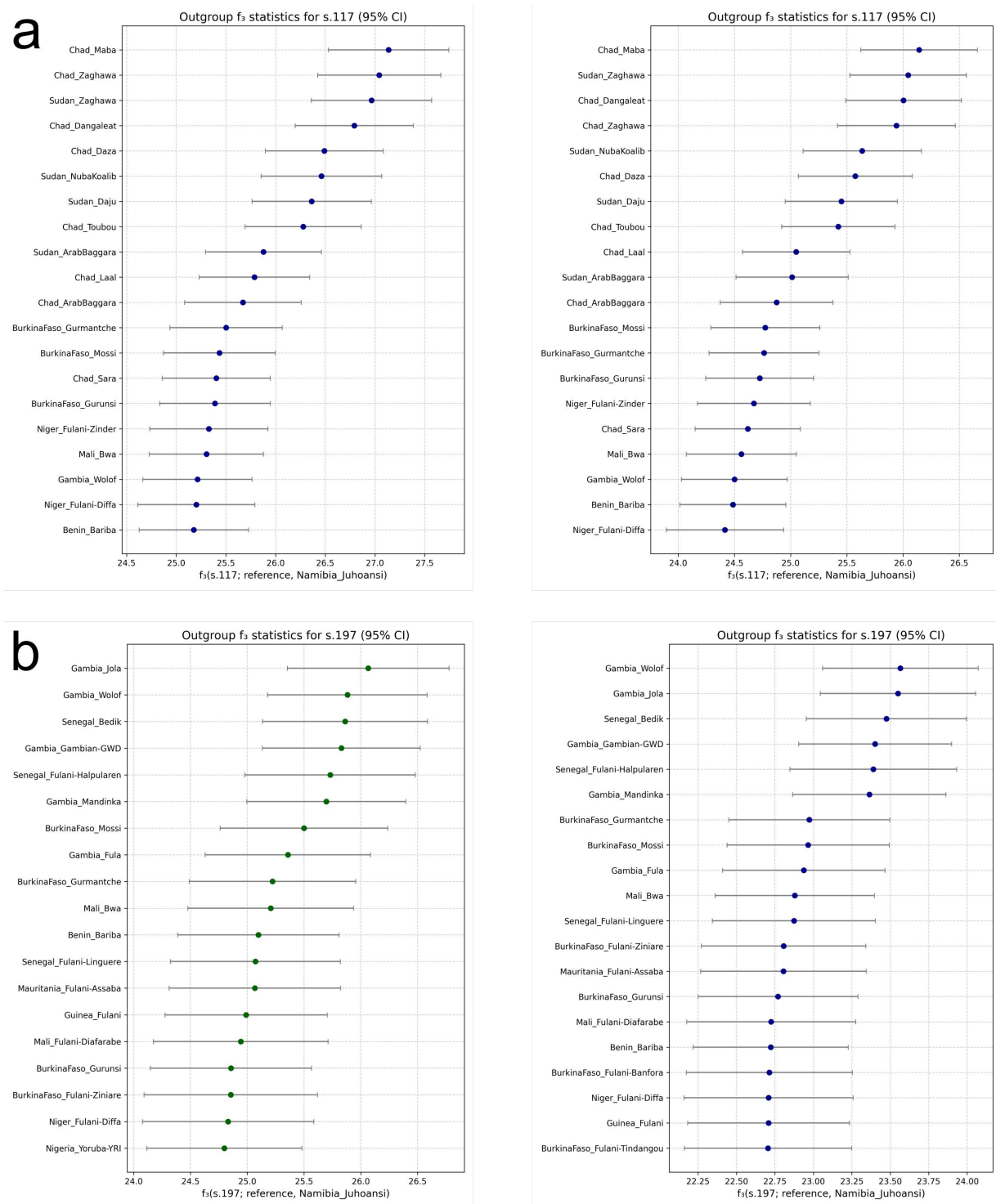


**Supplementary Fig. 6. Principal Component Analysis (PCA) showing the projection of ancient pseudo-haploid (left) and imputed diploid (right) Ibiza Sub-Saharan individuals onto modern African genetic variation with reference AfriCal dataset.**

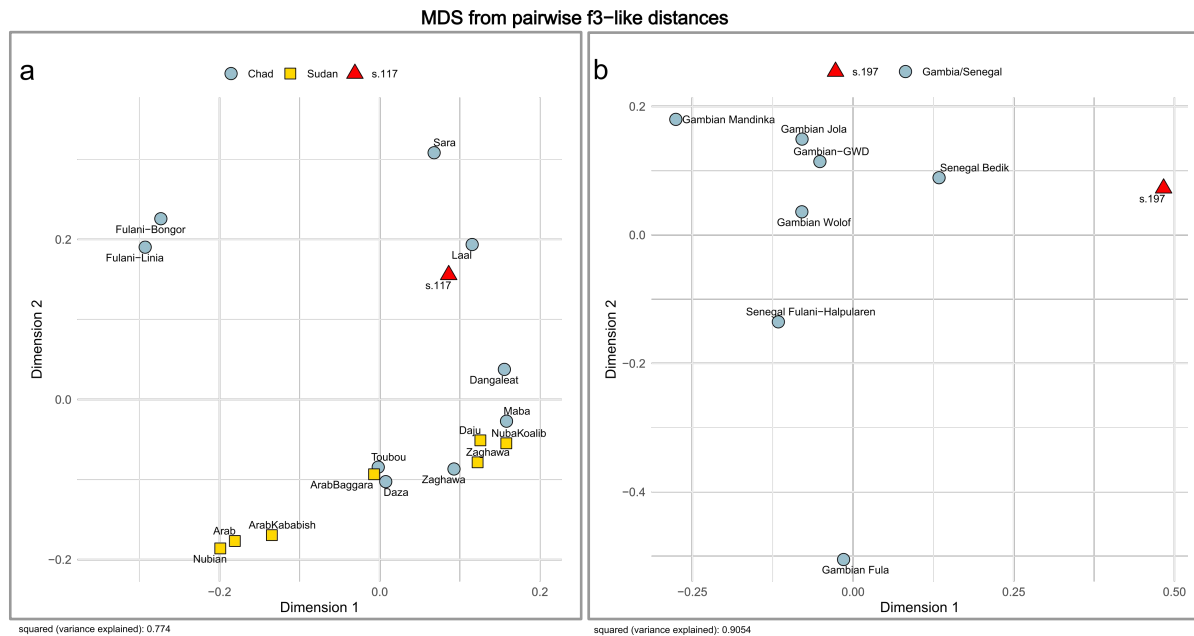




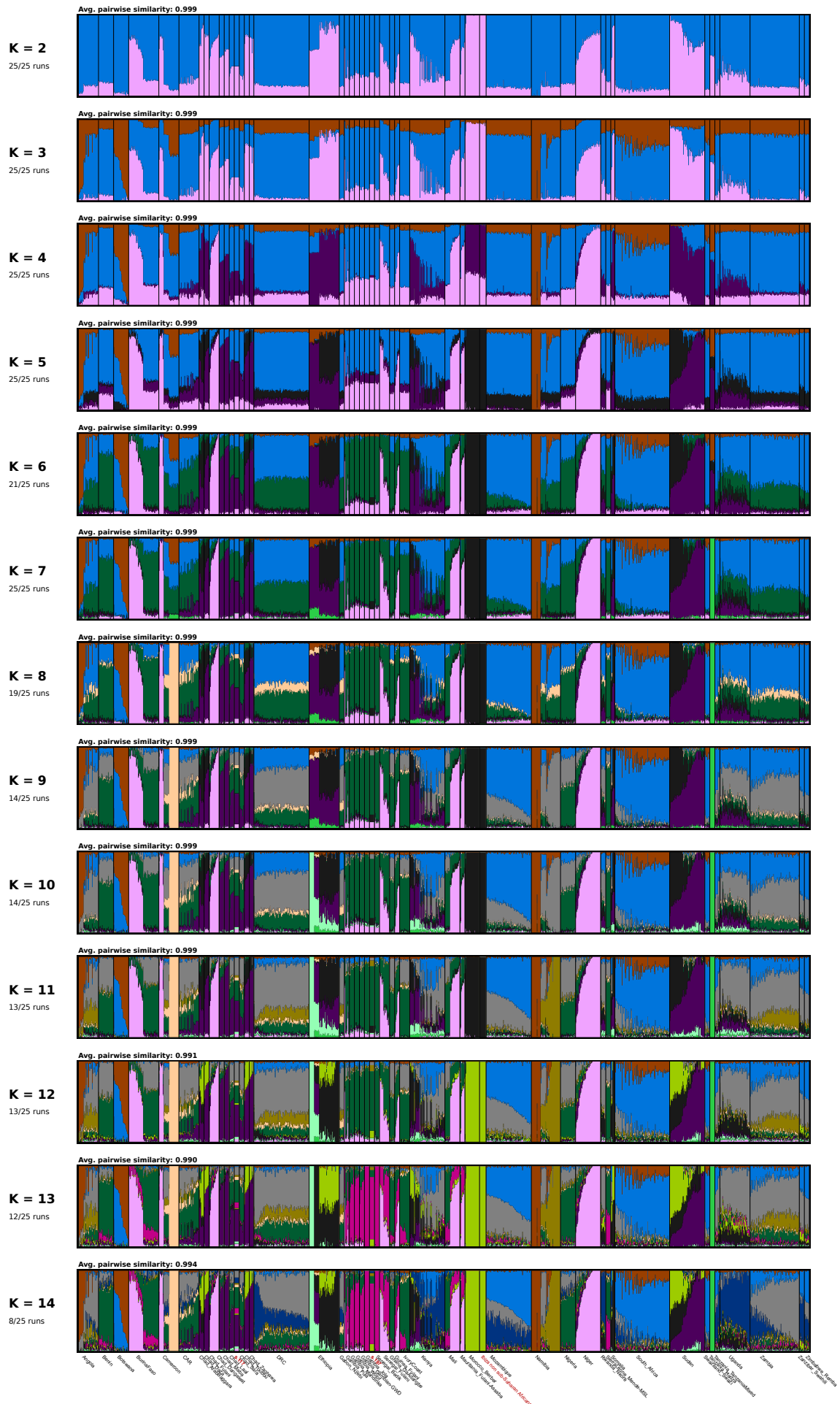
**Supplementary Fig. 7. Principal Component Analysis (PCA) showing the projection of ancient pseudo-haploid (left) and imputed diploid (right) Ibiza Sub-Saharan individuals onto modern African genetic variation with reference Africa2 dataset.**



**Supplementary Fig. 8. Outgroup  $f_3$ -statistics measuring shared genetic drift between sub-Saharan African ancient Ibiza individuals and modern African populations.** Each dot represents an  $f_3$ -statistic of the form  $f_3(\text{Ibiza}; \text{Modern Population}, \text{Outgroup})$ , with higher values indicating greater shared genetic ancestry. Error bars represent 95% confidence intervals. The outgroup used is the Namibian Ju'hoansi. Results are shown for two datasets: one using pseudohaploid ancient genotypes and the other using imputed diploid ancient genotypes. a. Individual s.117. b. Individual s.197.

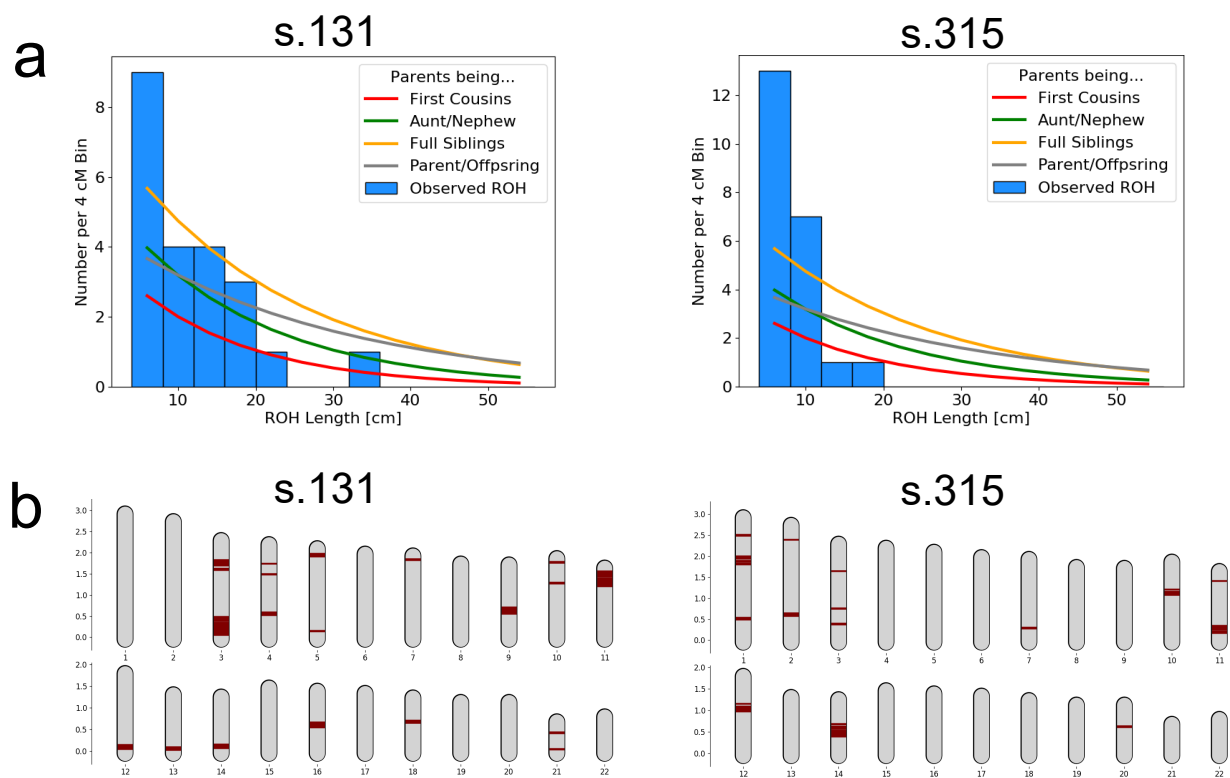


**Supplementary Fig. 9. Multidimensional scaling (MDS) plots based on pairwise outgroup  $f_3$ -statistics.** The outgroup used is the Namibian Ju'hoansi. **a.** MDS of individual s.117 and populations from Chad and Sudan, showing clustering of s.117 with groups from southern Chad (e.g., Laal). **b.** MDS of individual s.197 and populations from the Senegambia region, indicating affinities with several West African groups.

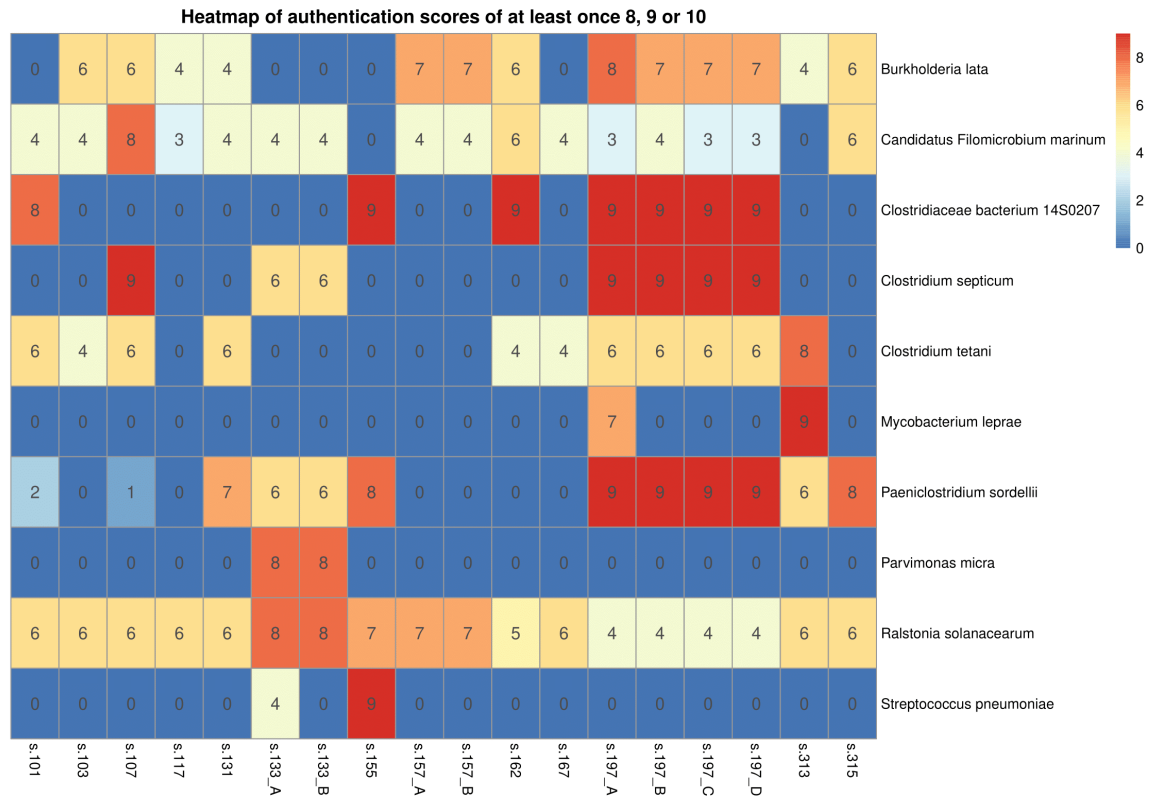


**Supplementary Fig. 10. Unsupervised ADMIXTURE analysis using Africa Dataset 1.** For visualization purposes, the two Sub-Saharan African ancient samples, s.117 and s.197, were replicated 10 times in the plot.

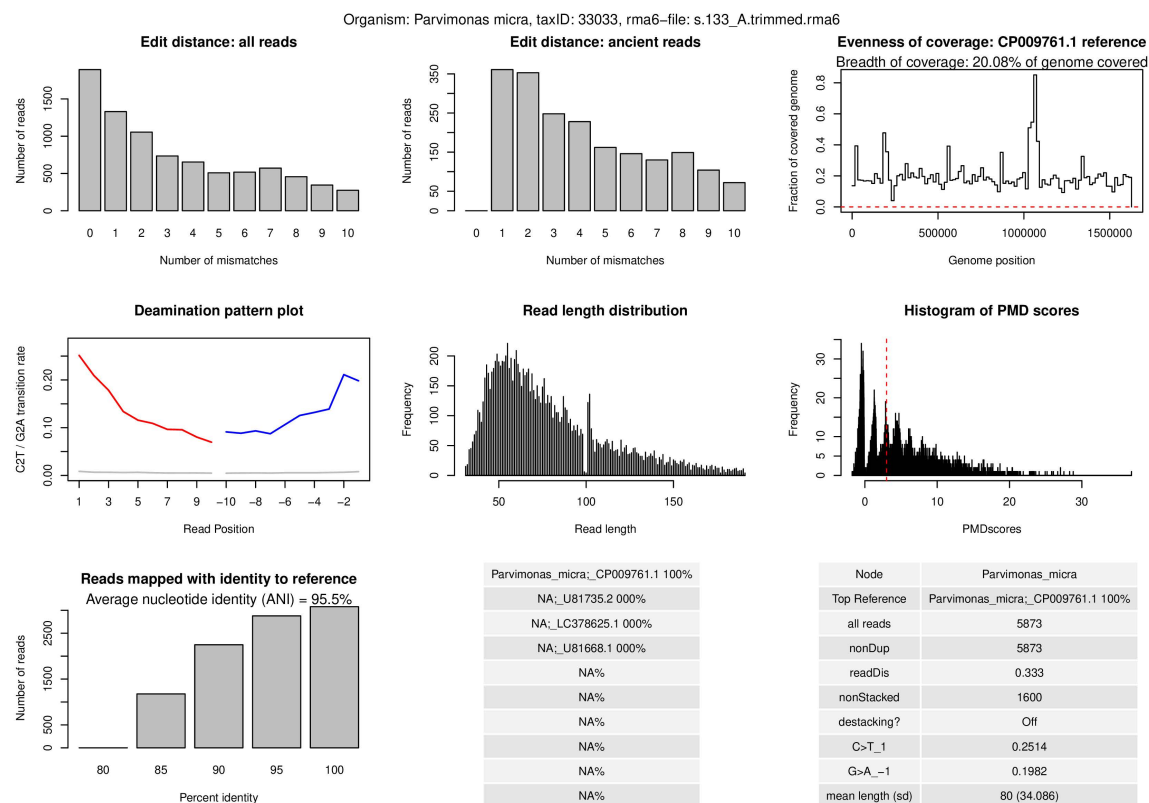




**Supplementary Fig. 12. Runs of homozygosity in the individual with the highest total ROH.**  
**a.** ROH lengths and expected density distributions for different degrees of parental relatedness. **b.** Chromosomal positions of the ROH fragments in this individual.

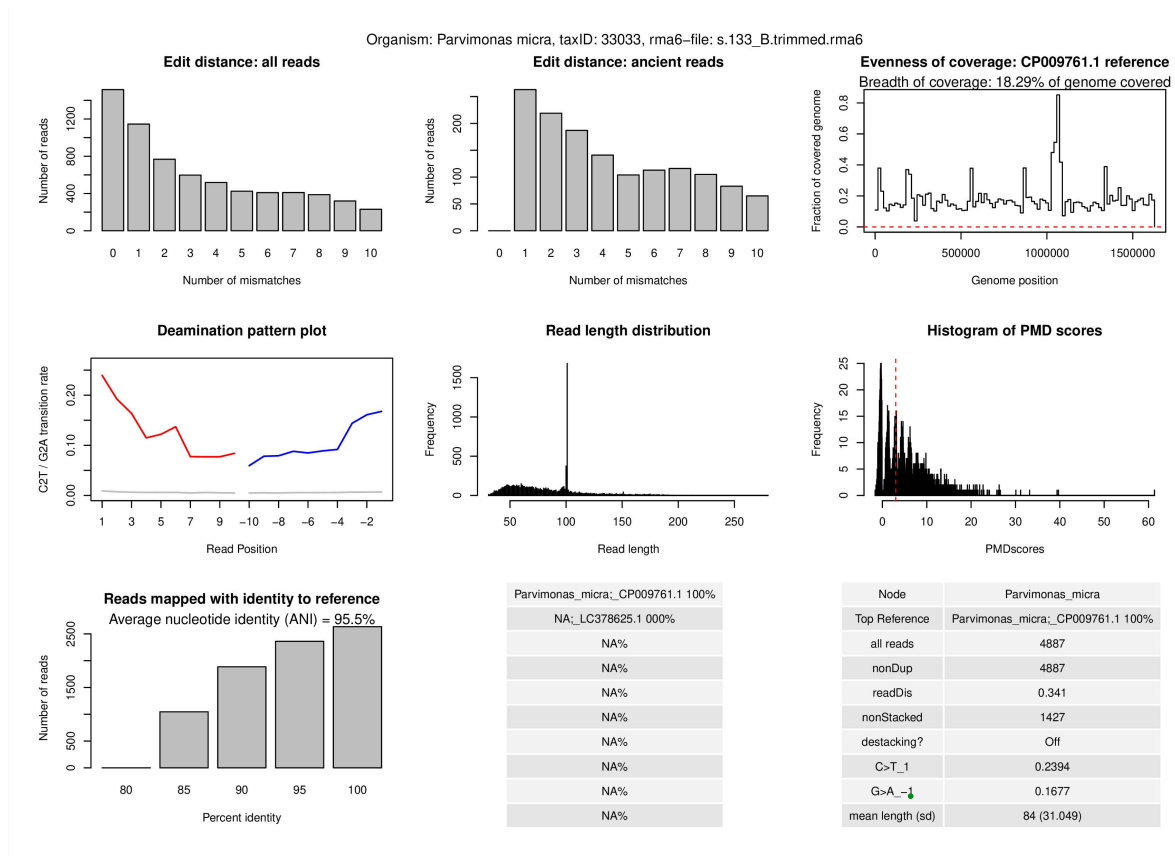


**Supplementary Fig. 13. Heatmap showing authentication scores  $\geq 8$  per microbe and per individual.** Scores reflect nine evaluation criteria: edit distance, ancient-read edit distance, terminal deamination (both ends), average read length, PMD score, average nucleotide identity (ANI), number of reads, and coverage evenness (which contributes two points). Values range from 1 to 10.

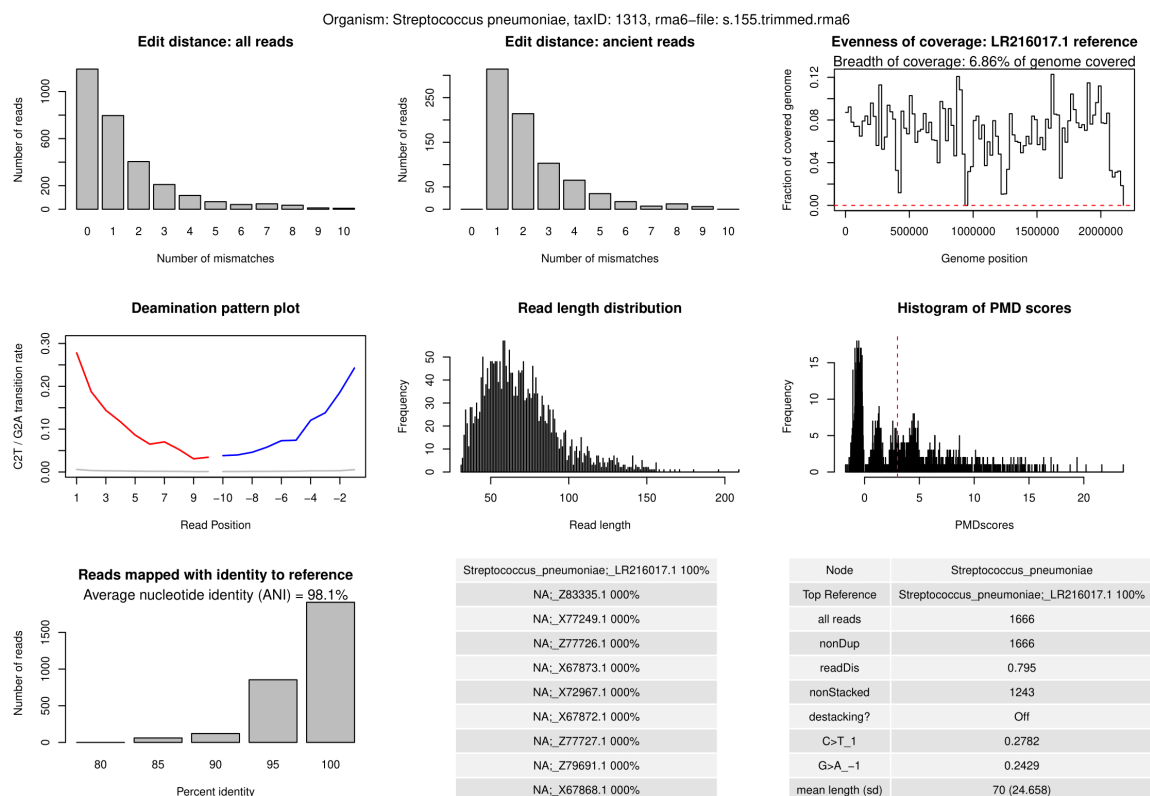


**Supplementary Fig. 14. Authentication plots for *Parvimonas micra* in individual s.133\_A.** The panels display edit distance distributions (all vs ancient reads), genome coverage breadth and evenness, terminal deamination rates, read length distribution, PMD score histogram, and percent identity of mapped reads. A summary table includes top reference matches and mapping statistics. The overall authentication score is 8.

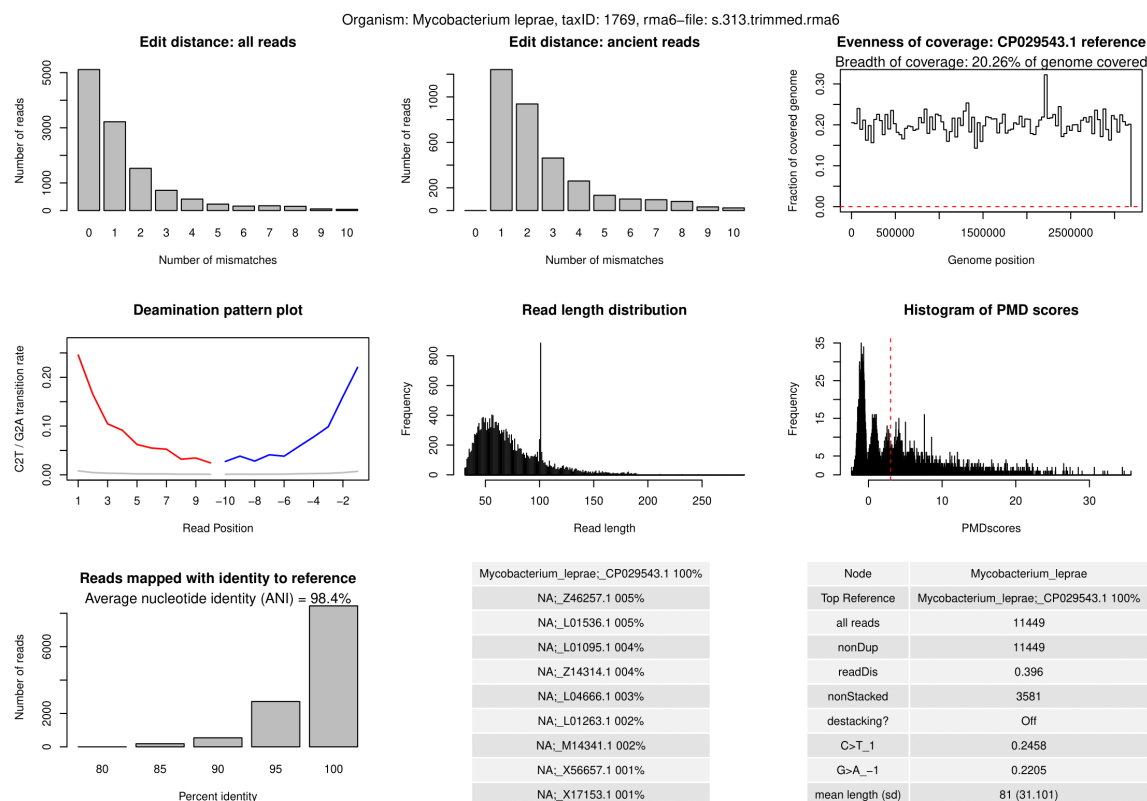




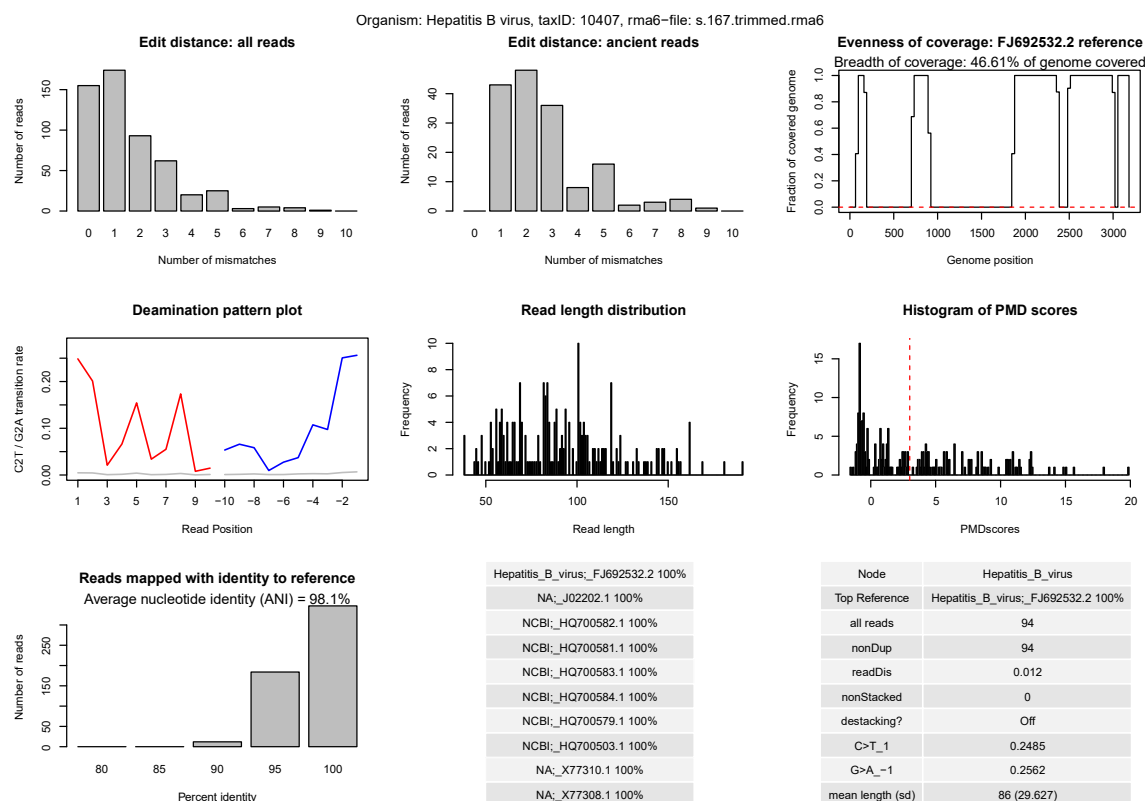
**Supplementary Fig. 15. Authentication plots for *Parvimonas micra* in individual s.133\_B.** The panels display edit distance distributions (all vs ancient reads), genome coverage breadth and evenness, terminal deamination rates, read length distribution, PMD score histogram, and percent identity of mapped reads. A summary table includes top reference matches and mapping statistics. The overall authentication score is 8.



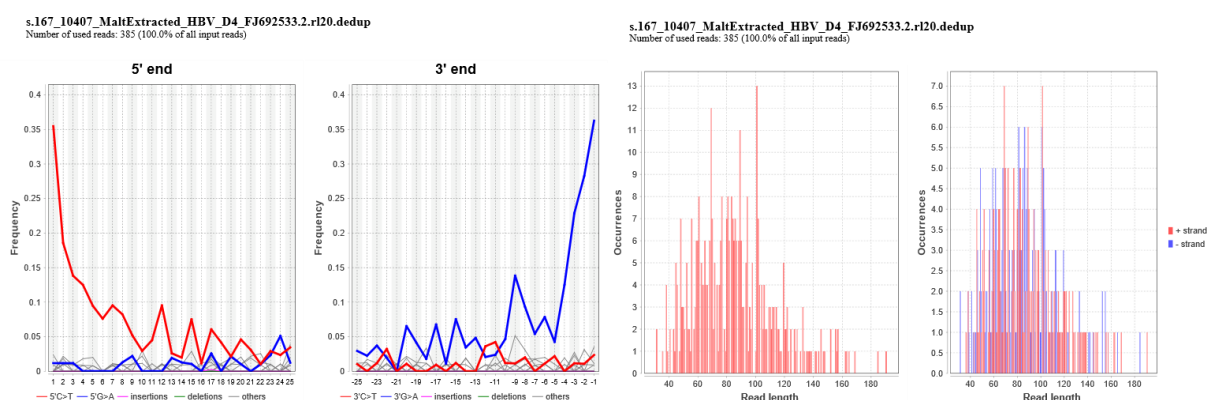
**Supplementary Fig. 16. Authentication plots for *Streptococcus pneumoniae* in individual s.155.** The panels display edit distance distributions (all vs ancient reads), genome coverage breadth and evenness, terminal deamination rates, read length distribution, PMD score histogram, and percent identity of mapped reads. A summary table includes top reference matches and mapping statistics. The overall authentication score is 9.



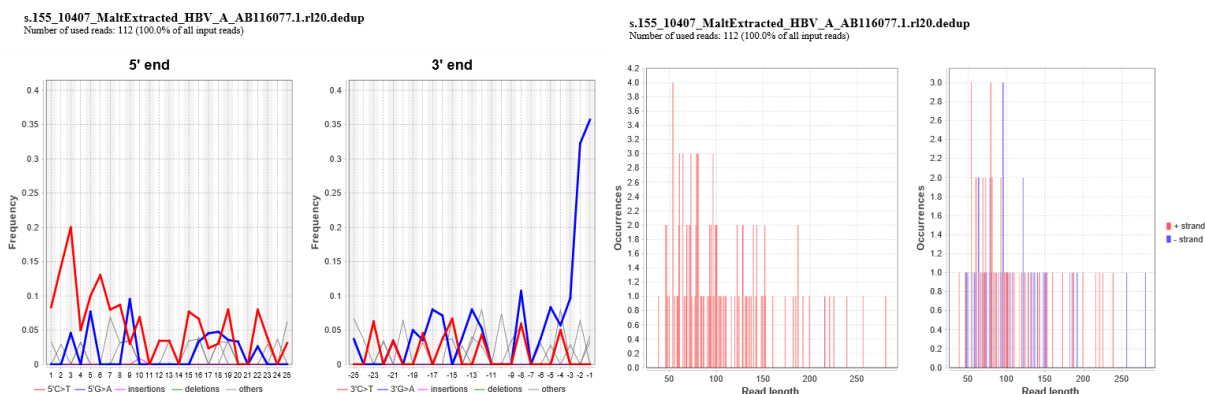
**Supplementary Fig. 17. Authentication plots for *Mycobacterium leprae* in individual s.313.** The panels display edit distance distributions (all vs ancient reads), genome coverage breadth and evenness, terminal deamination rates, read length distribution, PMD score histogram, and percent identity of mapped reads. A summary table includes top reference matches and mapping statistics. The overall authentication score is 9.



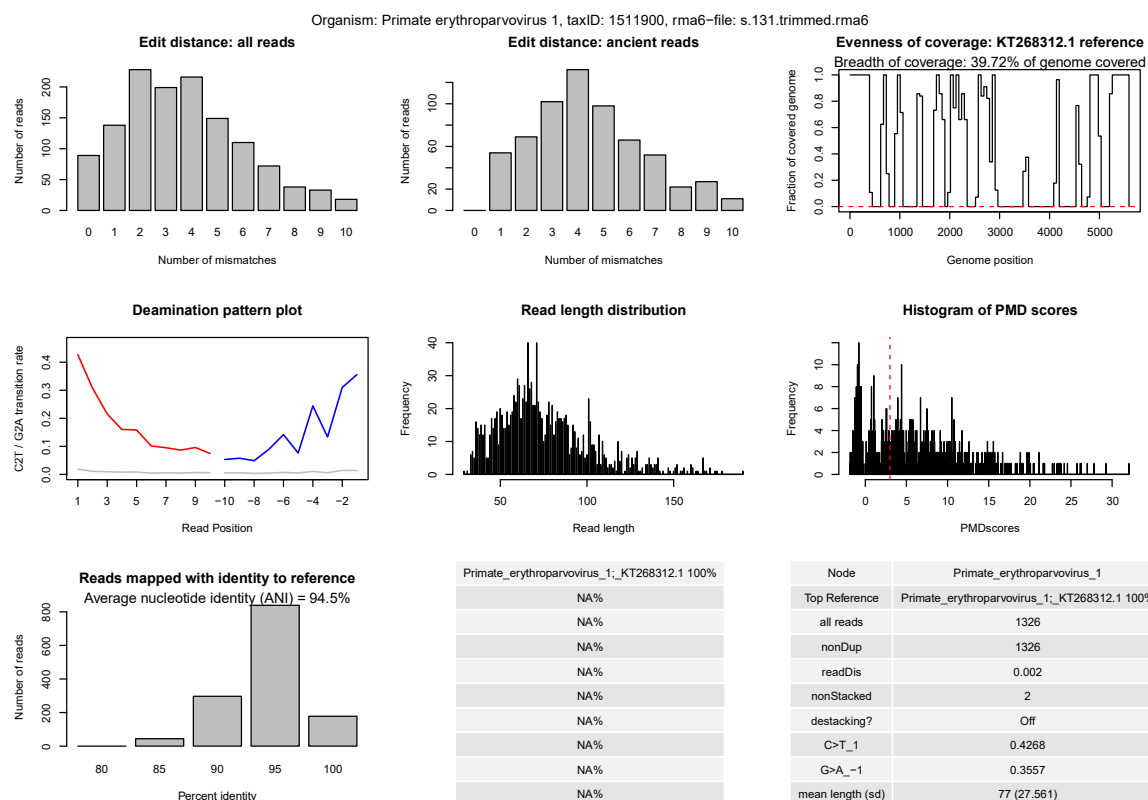
**Supplementary Fig. 18. Authentication plots for Hepatitis B virus in individual s.167.** The panels display edit distance distributions (all vs ancient reads), genome coverage breadth and evenness, terminal deamination rates, read length distribution, PMD score histogram, and percent identity of mapped reads. A summary table includes top reference matches and mapping statistics. The overall authentication score is 4.



**Supplementary Fig. 19. Deamination plots and length plots of the reads assigned to HBV in s.167 and extracted with MaltExtract.** The closest reference genome according to BLAST was used for mapping (FJ692533.2).

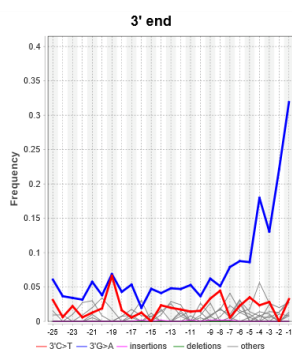
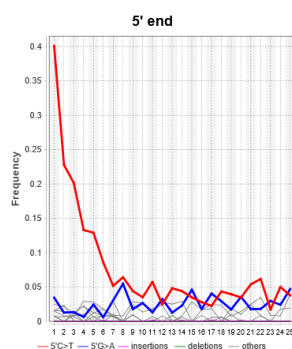


**Supplementary Fig. 20. Deamination plots and length plots of the reads assigned to HBV in s.155 and extracted with MaltExtract.** The closest reference genome according to BLAST was used for mapping (AB116077.1).

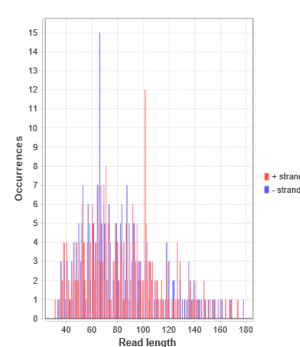
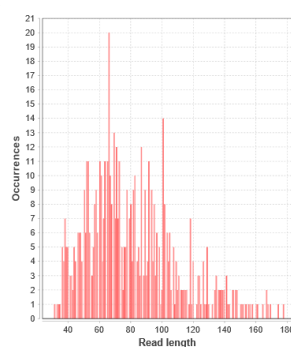


**Supplementary Fig. 21. Authentication plots for human parvovirus B19, alias primate erythroparvovirus 1, in individual s.131.** The panels display edit distance distributions (all vs ancient reads), genome coverage breadth and evenness, terminal deamination rates, read length distribution, PMD score histogram, and percent identity of mapped reads. A summary table includes top reference matches and mapping statistics. The overall authentication score is 4.

s.131\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 582 (100.0% of all input reads)

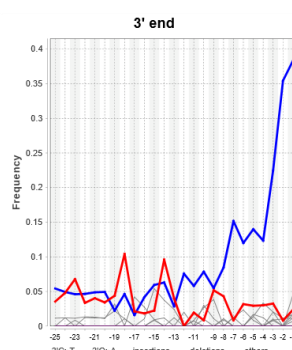
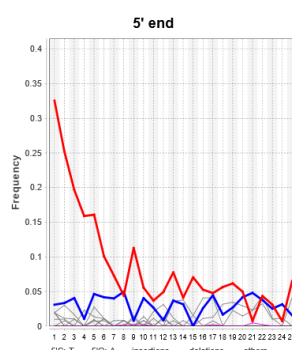


s.131\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 582 (100.0% of all input reads)

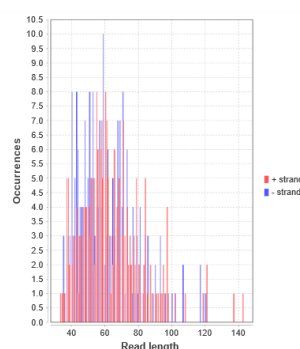
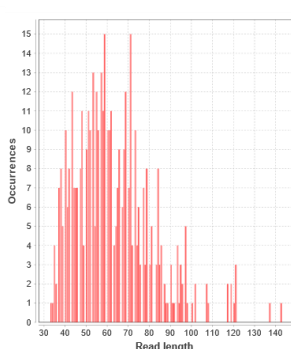


**Supplementary Fig. 22. Deamination plots and length plots of the reads assigned to B19V in s.131 and extracted with MaltExtract.** The closest complete reference genome according to BLAST was used for mapping (ON023027.1).

s.133\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 434 (100.0% of all input reads)

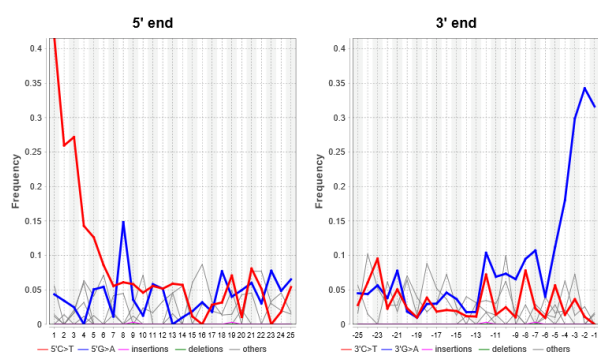


s.133\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 434 (100.0% of all input reads)

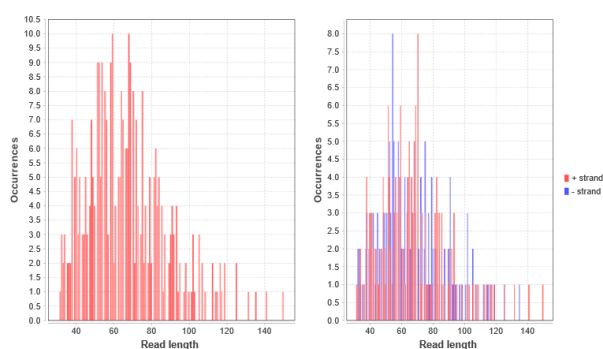


**Supplementary Fig. 23. Deamination plots and length plots of the reads assigned to B19V in s.133 and extracted with MaltExtract.** The reference genome ON023027.1 was used for mapping.

s.157\_merged\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 330 (100.0% of all input reads)

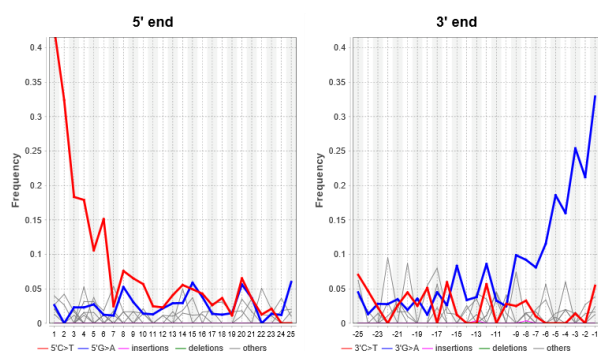


s.157\_merged\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 330 (100.0% of all input reads)

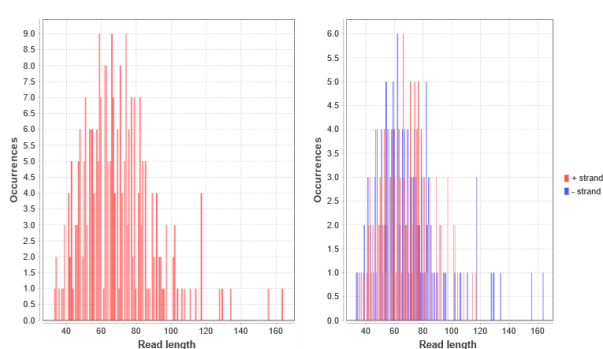


**Supplementary Fig. 24. Deamination plots and length plots of the reads assigned to B19V in s.157 and extracted with MaltExtract. The reference genome ON023027.1 was used for mapping.**

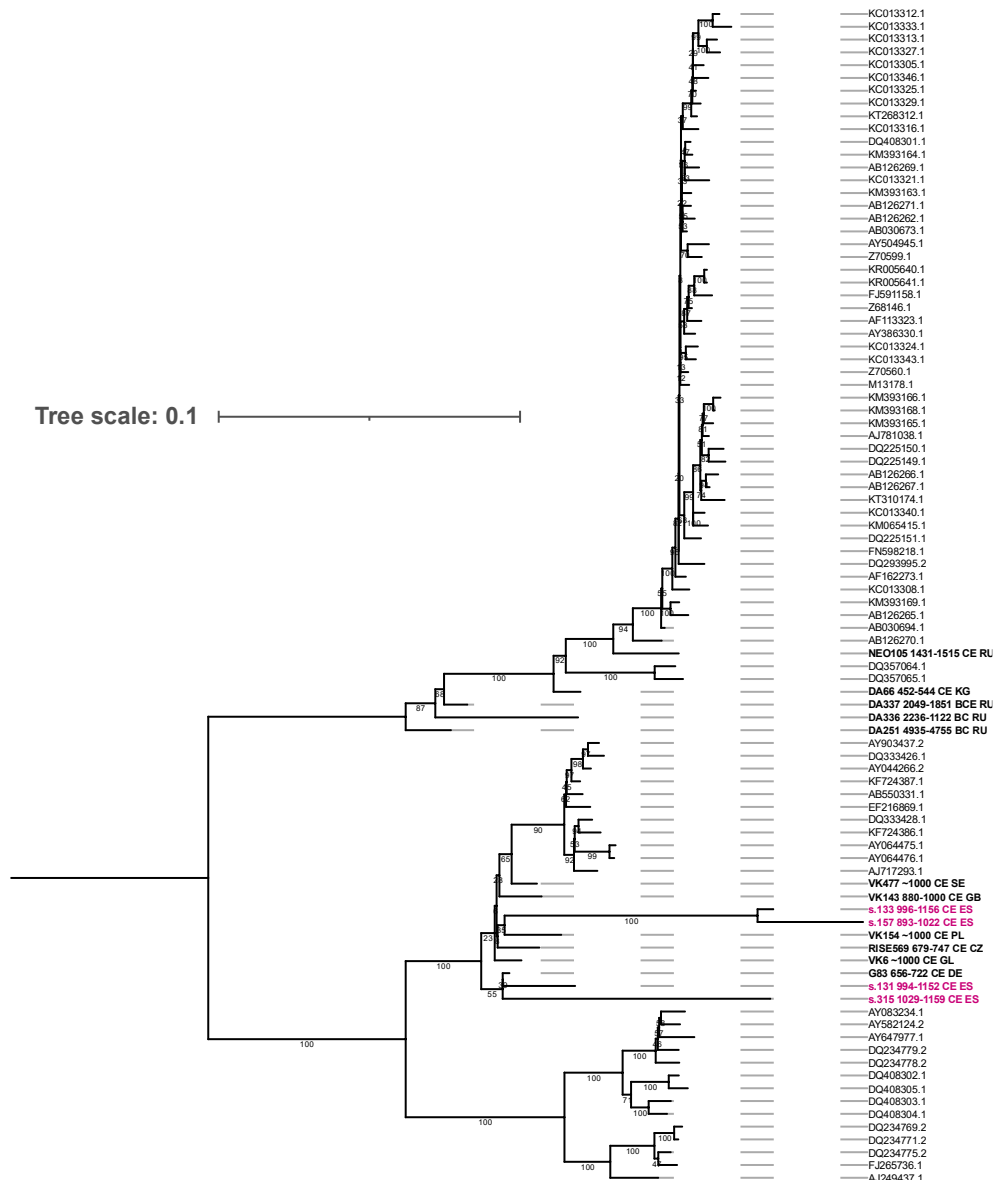
s.315\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 275 (100.0% of all input reads)



s.315\_1511900\_MaltExtracted\_B19V\_gen2\_ON023027.r120.dedup  
Number of used reads: 275 (100.0% of all input reads)

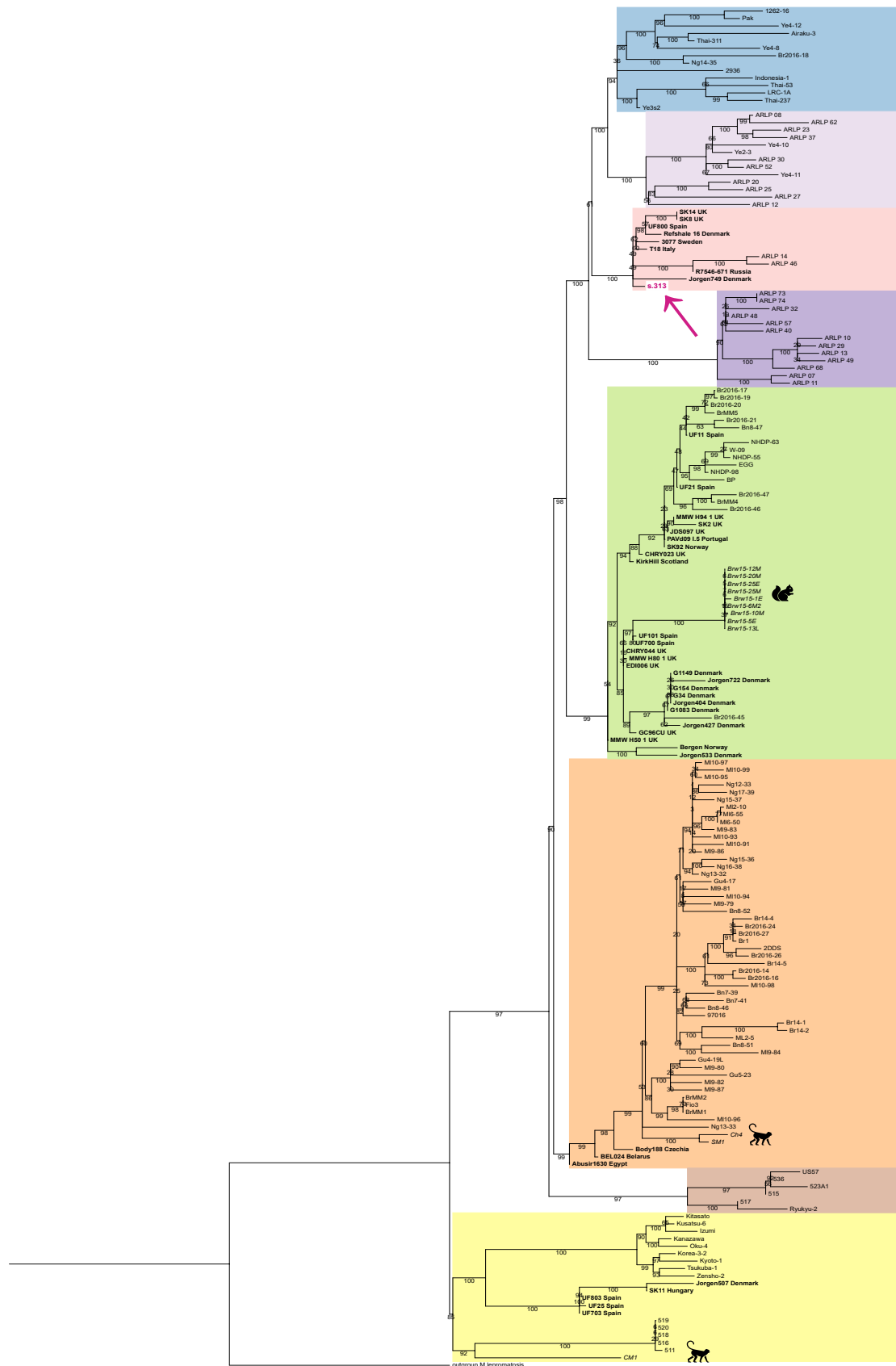


**Supplementary Fig. 25. Deamination plots and length plots of the reads assigned to B19V in s.315 and extracted with MaltExtract. The reference genome ON023027.1 was used for mapping.**



**Supplementar Fig. S26. Whole-genome maximum-likelihood (ML) phylogeny of human parvovirus B19.** The Ibiza samples (s.131, s.133, s.157 and s.315) are highlighted in red and fall within genotype 2. Ancient genomes are shown in bold with sample ID and country; modern samples are labelled with their GenBank accession numbers. Bootstrap values indicate branch support. The tree was reconstructed using the TIM3+F+R3 model and 100 bootstrap replicates and was midpoint-rooted.





**Supplementary Fig. 27. Maximum-likelihood (ML) phylogeny of *Mycobacterium leprae*.** The Ibiza genome (s.313) is highlighted in red and falls within clade 2F. Ancient genomes are shown in bold with sample ID and country; those from animals are italicised; modern samples are labelled with the sample ID only. Bootstrap values indicate branch support. The tree was inferred from high-quality informative SNPs ( $\geq 3X$  coverage, genotype quality  $\geq 30$ , base frequency  $\geq 90\%$ ) retained at positions covered in  $\geq 90\%$  of the samples. *Mycobacterium lepromatosis* was used as the outgroup. The phylogeny was reconstructed using the GTR+F model and 100 bootstrap replicates.



**Supplementary Fig. 28. Maximum Parsimony (MP) phylogeny of *Mycobacterium leprae*.**

The Ibiza genome (s.313) is highlighted in red and falls within clade 2F. Ancient genomes are shown in bold with sample ID and country; those from animals are italicised; modern samples are labelled with the sample ID only. Bootstrap values indicate branch support. The tree was inferred from high-quality informative SNPs ( $\geq 3X$  coverage, genotype quality  $\geq 30$ , base frequency  $\geq 90\%$ ) retained at positions covered in  $\geq 80\%$  of the samples. *Mycobacterium lepromatosis* was used as the outgroup. The phylogeny was reconstructed with the Subtree-Pruning-Regrafting method with 500 bootstrap replicates.

## Supplementary References

1. Govan, E., Jackson, A. L., Inger, R., Bearhop, S. & Parnell, A. C. *simmr*: a package for fitting stable isotope mixing models in R. R package version (2023).
2. Alexander, M. M., Gerrard, C. M., Gutiérrez, A. & Millard, A. R. Diet, society and economy in late medieval Spain: stable isotope evidence from Muslims and Christians from Gandía, Valencia. *Am. J. Phys. Anthropol.* **156**, 263–273 (2015).
3. Pérez-Ramallo, P. *et al.* Unravelling social status in the first medieval military order of the Iberian Peninsula using isotope analysis. *Sci. Rep* **14**, 11074 (2024).
4. Dury, G. *et al.* The Islamic cemetery at 33 Bartomeu Vicent Ramon, Ibiza: investigating diet and mobility through light stable isotopes in bone collagen and tooth enamel. *Archaeol. Anthropol. Sci* **11**, 3913–3930 (2019).
5. Bronk Ramsey, C. *OxCal v4.4 Manual*. (Oxford Radiocarbon Accelerator Unit, University of Oxford, Oxford, 2021).
6. Reimer, P. J. *et al.* The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP. *Radiocarbon* **62**, 725–757 (2020).
7. Siani, G. *et al.* Radiocarbon reservoir ages in the Mediterranean Sea and Black Sea. *Radiocarbon* **42**, 271–280 (2000).
8. Fellows Yates, B. J. A. *et al.* Community-curated and standardised metadata of published ancient metagenomic samples with AncientMetagenomeDir. *Sci. Data* **8**, **31**, (2021).
9. Borry, M. *et al.* Facilitating accessible, rapid, and appropriate processing of ancient metagenomic data with AMDiT. *F1000Research* **12**, 926 (2024).
10. Schuenemann, V. J. *et al.* Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science* **341**, 179–183 (2013).
11. Schuenemann, V. J. *et al.* Ancient genomes reveal a high diversity of *Mycobacterium leprae* in medieval Europe. *PLoS pathog.* **14**, 1006997 (2018).
12. Krause-Kyora, B. *et al.* Ancient DNA study reveals HLA susceptibility locus for leprosy in medieval Europeans. *Nat. Commun.* **9**, 1569 (2018).
13. Mendum, T. A. *et al.* *Mycobacterium leprae* genomes from a British medieval leprosy hospital: towards understanding an ancient epidemic. *BMC genomics* **15**, 1–8 (2014).
14. Fotakis, A. K. *et al.* Multi-omic detection of *Mycobacterium leprae* in archaeological human dental calculus. *Philos. Trans. R. Soc. B* **375**, 20190584 (2020).
15. Neukamm, J. *et al.* 2000-year-old pathogen genomes reconstructed from metagenomic analysis of Egyptian mummified individuals. *BMC bio.* **18**, 1–18 (2020).

16. Pfrengle, S. *et al.* Mycobacterium leprae diversity and population dynamics in medieval Europe from novel ancient genomes. *BMC biology* **19**, 1–18 (2021).
17. Urban, C. *et al.* Ancient Mycobacterium leprae genome reveals medieval English red squirrels as animal leprosy host. *Curr. Biol.* **34**, 2221–2230 (2024).
18. Avanzi, C. *et al.* Red squirrels in the British Isles are infected with leprosy bacilli. *Science* **354**, 744–747 (2016).
19. Avanzi, C. *et al.* Transmission of drug-resistant leprosy in Guinea-Conakry detected using molecular epidemiological approaches. *Clin. Infect. Dis.* **63**, 1482–1484 (2016).
20. Benjak, A. *et al.* Phylogenomics and antimicrobial resistance of the leprosy bacillus Mycobacterium leprae. *Nat. commun.* **9**, 352 (2018).
21. Blevins, K. E. *et al.* Evolutionary history of Mycobacterium leprae in the Pacific Islands. *Philos. Trans. R. Soc. B* **375**, 20190582 (2020).
22. Honap, T. P. *et al.* Mycobacterium leprae genomes from naturally infected nonhuman primates. *PLoS Negl. Trop. Dis.* **12**, 0006190 (2018).
23. Singh, P. *et al.* Insight into the evolution and origin of leprosy bacilli from the genome sequence of Mycobacterium lepromatosis. *Proc. Natl Acad. Sci. USA* **112**, 4459–4464 (2015).
24. Truman, R. W. *et al.* Probable zoonotic leprosy in the southern United States. *N. Engl. J. Med.* **364**, 1626–1633 (2011).