



## OPEN Intra-individual variability in ancient *plasmodium* DNA recovery highlights need for enhanced sampling

Alejandro Llanos-Lizcano<sup>1,3,7</sup>, Michelle Hämmerle<sup>1,2,7</sup>, Alessandra Sperduti<sup>4,5</sup>, Susanna Sawyer<sup>1,2</sup>, Brina Zagorc<sup>1,2</sup>, Kadir Toykan Özdoğan<sup>6</sup>, Meriam Guellil<sup>1,2</sup>, Olivia Cheronet<sup>1,2</sup>, Martin Kuhlwilm<sup>1,2</sup>, Ron Pinhasi<sup>1,2</sup>✉ & Pere Gelabert<sup>1,2</sup>✉

Malaria has been a leading cause of death in human populations for centuries and remains a major public health challenge in African countries, especially affecting children. Among the five *Plasmodium* species infecting humans, *Plasmodium falciparum* is the most lethal. Ancient DNA research has provided key insights into the origins, evolution, and virulence of pathogens that affect humans. However, extensive screening of ancient skeletal remains for *Plasmodium* DNA has shown that such genomic material is rare, with no studies so far addressing potential intra-individual variability. Consequently, the pool of ancient mitochondrial DNA (mtDNA) or genomic sequences for *P. falciparum* is extremely limited, with fewer than 20 ancient sequences available for genetic analysis, and no complete *P. falciparum* mtDNA from Classical antiquity published to date. To investigate intra-individual diversity and genetic origins of *P. falciparum* from the Roman period, we generated 39 sequencing libraries from multiple teeth and two from the femur of a Roman malaria-infected individual. The results revealed considerable variability in *P. falciparum* recovery across different dental samples within the individual, while the femur samples showed no preservation of *Plasmodium* DNA. The reconstructed 43-fold *P. falciparum* mtDNA genome supports the hypothesis of an Indian origin for European *P. falciparum* and suggests mtDNA continuity in Europe over the past 2000 years.

Malaria is an infectious disease caused by various *Plasmodium* species, whereby at least five infect humans. *Plasmodium vivax* has the broadest distribution, while *Plasmodium falciparum* is responsible for most malaria-associated deaths. Both *P. vivax* and *P. falciparum* are believed to have originated in Africa around 50,000–60,000 years ago<sup>1,2</sup> and spread worldwide with complex patterns of migration, probably following human migrations<sup>1,3–8</sup>. Until the 20th century, malaria was globally spread and its widest-known geographical distribution included most of Eurasia and North America<sup>9,10</sup>. Its present-day distribution predominantly spans warmer climates near the equator, covering Africa, parts of the Middle East, Southeast Asia, China, and the Americas. It is estimated that in 2022, approximately 608,000 people died from malaria<sup>11</sup>. Due to this extensive geographic reach and clinical, malaria remains one of the most significant health threats to humans.

It has been hypothesised that both *P. vivax* and *P. falciparum* may have reached Europe during the Neolithic, about 8,500 years ago, due to a combination of favourable parameters, including climatic conditions, increased human population densities, and the presence of a capable *Anopheles* mosquito vector species<sup>4</sup>. However, this claim lacks archaeological or genetic evidence and remains contested<sup>2</sup>. The current oldest evidence of *P. vivax* comes from a German Middle Neolithic individual and the oldest detected case of *P. falciparum* infection dates back to the Iron Age in Austria<sup>5</sup>. The ancient cases detected so far coincided with *Anopheles* spp.—endemic regions<sup>4</sup>. Scholars addressing the effects of this endemic disease on societies in antiquity have stressed the dramatic political and economic consequences of malaria, an issue sometimes neglected by historiography<sup>12–14</sup>.

<sup>1</sup>Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria. <sup>2</sup>Human Evolution and Archeological Sciences (HEAS), University of Vienna, Vienna, Austria. <sup>3</sup>Facultad de Química y Farmacia, Universidad del Atlántico, Barranquilla, Colombia. <sup>4</sup>Museo delle Civiltà, Roma, Italy. <sup>5</sup>Dipartimento di Archeologia, Asia, Africa e Mediterraneo, Università L'Orientale, Napoli, Italy. <sup>6</sup>Department of History and Art History, Utrecht University, Utrecht, Netherlands. <sup>7</sup>Alejandro Llanos-Lizcano and Michelle Hämmerle have contributed equally to this work. ✉email: ron.pinhasi@univie.ac.at; pere.gelabert@univie.ac.at

Malaria remained endemic in Europe until the 1970s, extending from the Baltic Sea to the Mediterranean<sup>15</sup>. Currently, less than 20 ancient strains of *P. falciparum*, *P. vivax* and *P. malariae* have been published.

The identification and sequencing of ancient malaria strains present several challenges. Firstly, osteological lesions are not distinctively indicative of infection. Some are broadly suggestive of underlying infections/diseases but can stem from other similarly-presenting pathologies. One example is recent research suggesting the presence of porous lesions with a higher prevalence in malaria-endemic areas<sup>16</sup>. However, it is impossible to confirm the aetiology of these lesions in archaeological skeletal collections. Consequently, like with *Hepatitis B Virus (HBV)*, *Mycobacterium tuberculosis*, *Yersinia pestis*, and other pathogens, the detection of malaria often depends on extensive random sampling of ancient individuals for genetic studies. Secondly, the success in recovering ancient *Plasmodium* genomes from human remains is contingent on the survival of *Plasmodium* DNA in bone and dental tissues. Some studies have suggested the presence of *P. falciparum* DNA in ancient individuals, such as a 5th-century common era (CE) infant from Lugnano in Teverina, Italy<sup>17</sup>, and ancient Egyptian mummies<sup>18</sup>. However, these studies could not conclusively prove the presence of the parasites, an issue that can now potentially be resolved with next-generation sequencing technologies. A recent study comparing the reliability of antigen detection and DNA sequencing in identifying pathogens in ancient skeletal remains found that, despite a limited sample size, paleogenomics methods are the most dependable for this purpose<sup>19</sup>. When soft tissue is available, microscopy remains an alternative as it is possible to observe the parasite inside the infected erythrocytes<sup>20</sup>. A recent study that screened more than 10,000 ancient individuals could only recover less than 20 strains suitable from mtDNA analyses<sup>5</sup>.

Paleogenomics also helps to determine malaria's historical impact on societies, revealing selective pressures and human adaptation to the parasite. Recent research has shown that the significant occurrence of an enzyme deficiency (G6PD) providing resistance against *Plasmodium* infection<sup>21</sup> among present-day Arab populations likely traces back 6000 years before present (BP), aligning with the onset of the Bronze Age in Eastern Arabia. This indicates a possible link between the spread of agriculture and increased resistance to *Plasmodium* infection<sup>22</sup>. However, the global effect of malaria in ancient populations as a selective pressure is contested<sup>23</sup>.

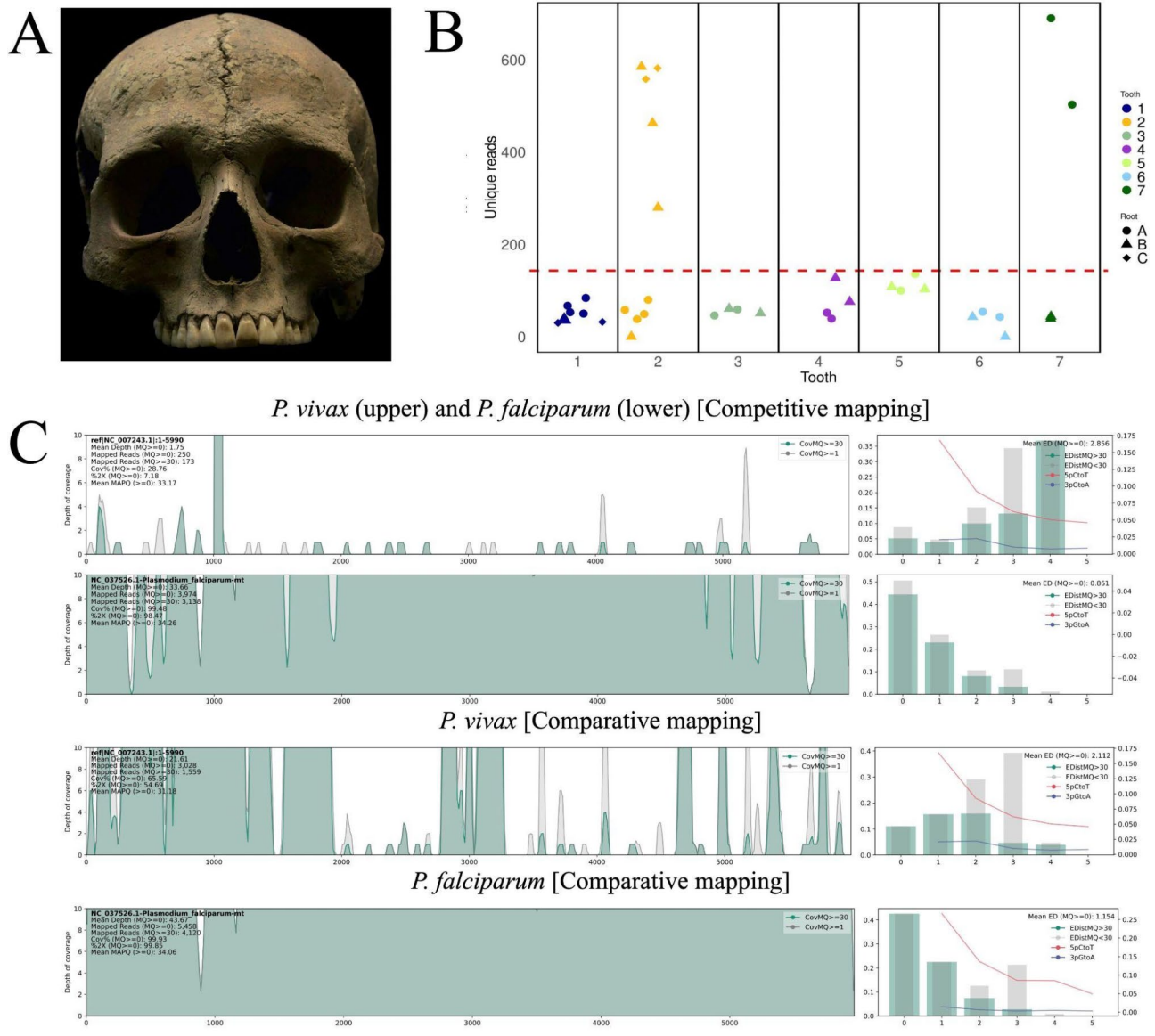
Ancient *P. falciparum* detections are currently restricted to Eurasia. All known ancient samples show genetic affinities to modern Indian strains, which is evidenced by both nuclear and mitochondrial data<sup>5,6,24,25</sup>. Ancient strains of *P. vivax* from Eurasia spanning almost 1,000 years show similarity to known modern American strains suggesting a likely European colonial origin of these lineages<sup>5,6</sup>. However, data on ancient *Plasmodium* species remains sparse, and there is a need to increase the number of *Plasmodium* strains to enhance our understanding of *Plasmodium* dispersal patterns and affinities in historical contexts, specially within periods linked to extensive mobility through the Mediterranean such as the Roman Period<sup>26</sup>.

To date, no aDNA studies have assessed the differential preservation of *Plasmodium* aDNA across skeletal elements of the same individual. Here, we introduce the first Roman-era full mitochondrial genome sequence (43-fold coverage) of *Plasmodium falciparum*, derived from the individual known as Velia-186 (LV13), previously confirmed to be infected with the pathogen<sup>27</sup>. We demonstrate extensive differences in *Plasmodium* reads yields from different teeth of the same individual suggesting a differential and randomized preservation pattern, which required broad sampling at individual basis to overcome this limitation and recover substantial *Plasmodium* data.

## Results

In an initial screening, we produced sequencing libraries from both teeth and parts of the femur of individual Velia-186 (Velia, Porta Marina necropolis, I-II cent. CE, Fig. 1A). We targeted teeth based on previous evidence that pathogens are generally well-preserved in dental tissues<sup>28</sup>, and the femoral diaphysis and head, as *P. falciparum* gametocytes commonly mature in human bone marrow<sup>29</sup>. As one of the libraries from the lower left premolar yielded just over 100 unique reads for *P. falciparum* (Table S1C), we decided to test the possibility of differential recovery across the individual's dentition. For that purpose, 38 DNA libraries from seven teeth were generated by sampling at least two roots per tooth (see Table S1), to assess if different roots from the same tooth would yield different amounts of *Plasmodium*. The libraries were enriched with in-solution baits covering both *P. falciparum* and *P. vivax* mitochondrial genomes and sequenced on an Illumina NextSeq 550, generating paired-end reads with a length of 150 bp. Following preprocessing, quality control, and collapsing, a total of 946 million reads were recovered, with an average of 22 million reads per library (standard deviation (SD) = 1.167 million). The data was then merged for subsequent analysis. To assess whether the recovered reads originated from a co-infection of *P. falciparum* and *P. vivax*, we performed competitive mapping, which maps reads to different reference genomes simultaneously to ascertain which genome the read fits best. Competitive mappings for *P. falciparum* and *P. vivax* resulted in 33.66-fold (Fig. 1C, Table S1) and 1.75-fold (Fig. 1C, Table S1) depth of coverage, respectively, indicating that the vast majority of reads stem from *P. falciparum* and not *P. vivax*. We then performed comparative mapping, which maps to each genome individually to obtain the coverage per species. This mapping resulted in a mean depth of 21.61 for *P. vivax*. However, only 22% of the genome had coverage due to uneven mapping, which indicates mismapping. On the other hand, the mapping to *P. falciparum* showed an even distribution of coverage with an average depth of 43.67-fold (Fig. 1C, Table S1). Based on these results, we deduce that the Velia-186 individual was only infected with *P. falciparum* and exclude the possibility of a co-infection with *P. vivax* as in Ebro-1944 or STR samples. Hence, below, we focus on reads mapped to the *P. falciparum* mtDNA genome.

In our study, following the identification of only *P. falciparum* infection, we proceeded to analyse the *P. falciparum* DNA across various sequencing libraries. This analysis was centred on aligned reads with mapping scores above 30 (Table S1). Notably, we found no reads aligning to *P. falciparum* on the external surfaces of the first right inferior premolar nor in the femoral diaphysis or head (Table S1). On average, each library from the seven teeth produced 108 reads (with a standard deviation of 180 across all 38 libraries), which shows



**Fig. 1.** (A) Velia-186 skull (male 20–25 years old). Photo by the Museum of Civilizations. (B) The number of unique reads mapping to the reference genome K1 [NC\_037526] was recovered from each sequencing library. The red line denotes the average amount of reads recovered per library. Samples 2, 3 (Upper right second molar), and 8 (Lower right first molar) contribute the majority of reads. (C) Mapping plots to both *P. vivax* and *P. falciparum*. Reads with an MQ of or above 30 are depicted in green. The coverage is shown across the whole mitochondrial genome. The bar plot on the right illustrates the edit distance and the percentage of C → T mutations at the 5' end and the G → A mutations at the 3' end.

significant diversity. Intriguingly, a mere seven out of the 38 libraries contributed to 72% of the total unique fragments recovered (see Fig. 1B). Moreover, there is a clear uneven preservation pattern between the human and *P. falciparum* ancient DNA (aDNA), where different teeth yielded more unique reads for one or the other (see Fig. 1B and Fig. S2). Additionally, we inspected misincorporation patterns characteristic of aDNA, finding a clear deamination pattern in both 5' (22%) and 3' (18%) ends, corroborating an ancient origin of the sequencing data (Fig. 1C). Furthermore, the read-length distribution showed the typical small fragment lengths associated with ancient DNA (Table S1).

We analysed a potential correlation between the dental samples DNA yield and their respective unique Q30 mapped reads but found no significant correlation ( $R^2 = 0.02$  and  $p\text{-value} = 0.9$ ) (Table S3). Further, considering the non-normal distribution of reads, we investigated whether there was greater diversity within the different roots of one tooth or between the teeth. This revealed no significant variations in their medians ( $\chi^2(6) = 11.06$ ,  $p\text{-value} = 0.08$ ) (Table S3), suggesting that homogeneity of sequencing yields was observed between the dental samples.

We combined the 5458 mapped reads from all seven teeth, yielding a mitochondrial genome of *P. falciparum* with an average depth of 43-fold. Next, we used BLASTn<sup>30</sup> to identify any potential sources of contamination. BLASTn classified 4120 filtered reads that we later classified at the family/order level with MEGAN using an LCA algorithm<sup>31</sup>. From these, 3,913 were successfully classified, all in the order of Haemosporida, which confirmed that only *P. falciparum* is present and no contamination from the soil was affecting our samples. Following these results, we decided to use the 5458 *P. falciparum* reads for the downstream analyses.

We used the 5458 *P. falciparum* reads to generate a consensus sequence with ANGSD (version 0.941) and obtained a mitochondrial genome with 99.1% of the sequence covered. The consensus sequence has three gaps, with a size of 59 (858–917), 20 (1151–1171), and 9 (3495–3504) base pairs, respectively. The SNPs in our consensus sequence were inspected individually in IGV (v2\_16.0) to verify and validate their presence. Our high-coverage genome consensus sequence has seven mutations compared to the reference strain mitochondrial genome (74, A -> T; 276 G -> A; 725, C -> T; 772, T -> C; 2172, T -> C; 2763, C -> T; 3938, A -> T) all supported by at least ten reads (see Table S2). Four of the seven mutations are located in coding regions of the mtDNA (Table S2). However, none of these are non-synonymous using the Apollo<sup>32</sup> annotation tool accessed through PlasmoDB (v. 68)<sup>33</sup>. More details about the mutations are presented in Supplementary Table S2.

To elucidate how our *P. falciparum* consensus sequence compares to other *P. falciparum* genomes, we downloaded 345 mitochondrial genomes of *P. falciparum* strains from NCBI, representing all presently known *P. falciparum*. (see Table S4). First, we observed that only two substitutions, at positions 2,172 (T>C) and 3938 (A>T), were not present in any *P. falciparum* strain of the dataset. Next, we performed a maximum likelihood tree using a multiple sequence alignment, including our ancient consensus sequence ( $n = 345$ ). Although some clades cluster geographically, we observed that the current mtDNA genetic diversity does not exclusively reflect a geographical distribution, as previously reported<sup>34</sup>. In the phylogenetic tree, Velia-186 clusters exclusively with strains that are currently found in India as well as some ancient strains from Spain, France, Austria, Germany, Belgium, and Taiwan dated to different periods, with a 97 bootstrap support (Fig. 2 and S1). Out of the seven SNPs described above, mutations 276 (G>A), 725 (C>T) and 2763 (C>T) have been observed in present-day strains from different locations in India<sup>35</sup>. These SNPs are characteristic of the Indian subclade called PfIndia, described by Tyagi et al., 2014<sup>35</sup>, which further supports the phylogenetic clustering of the Velia-186 sequence close to the Indian strains. Further, the three mutations also exist in Ebro-1944, GOE016, LIP011, STR025, STR016 and MOB025 indicating genetic similarity between ancient European samples.

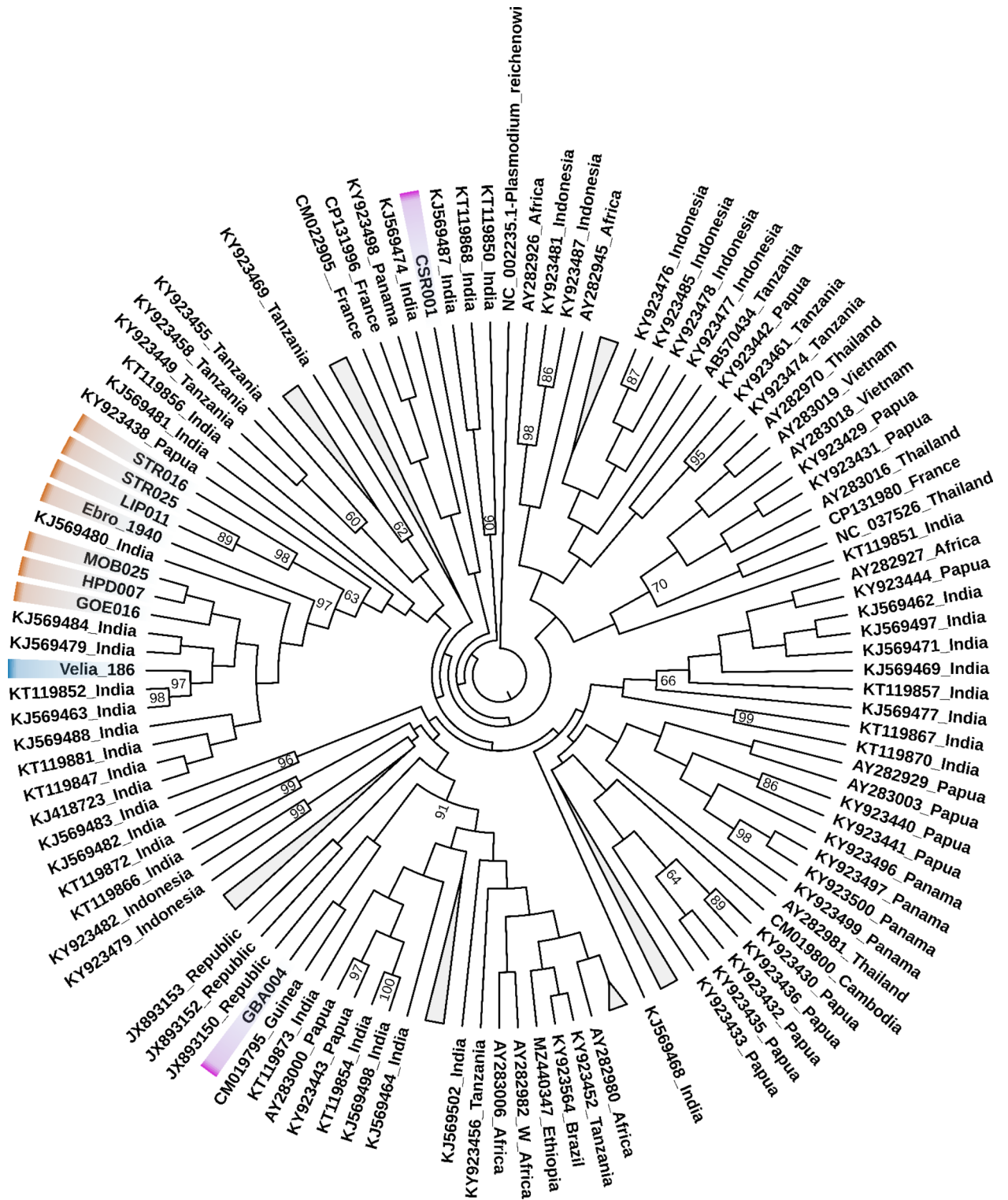
We tested the capacity to recover nuclear DNA from the seven best libraries by shotgun sequencing (Table S1D). The overall endogenous content is very low (below 0.01%). Out of 50,000,000 sequenced reads, we only recovered 1931 unique ones belonging to *P. falciparum*. Assuming the current complexity would enable it, we estimated it would require more than 16,000 million sequencing reads to target a 1X *P. falciparum* genome, the prohibitive cost of which makes it impossible.

We also aligned the sequencing reads against the human genome (hg19), including the mtDNA sequence (rCRS). We could recover 2,141,320 reads by combining all the libraries; 53,022 reads are aligned to the human rCRS genome. We observe differences in human DNA preservation per tooth, but these do not coincide with the preservation of *Plasmodium* DNA, as most reads are from tooth 6 (Fig. S2). The recovered reads also show consistent signals of deamination due to age (Fig. S3). We recovered an 1169X mtDNA human sequence and identified that the individual Velia 186 belongs to haplogroup T2b7a (88%), which is already identified in MBA individuals from the Levant<sup>36</sup>.

## Discussion

Implementing next-generation sequencing (NGS) in paleogenomics has enabled the detailed study of microbial agents' genomic diversity, evolutionary trajectories, and pathogenicity, moving beyond merely detecting their presence in ancient remains<sup>28</sup>. Detecting *Plasmodium* infections in ancient human remains is still<sup>19</sup> particularly challenging due to limited genomic material and degraded samples<sup>27</sup>, making it difficult to obtain comprehensive mtDNA or genomic data for pre-20th century *P. falciparum*. Building on prior findings that Velia-186 was infected with *Plasmodium*<sup>27</sup>, this study aimed to sequence the complete mtDNA to infer its genetic affinities and assess the differential preservation of *Plasmodium* DNA across skeletal elements. Here, we show that these challenges persist beyond identifying infected individuals through extensive sampling, highlighting the importance of selecting appropriate skeletal elements for analysis. Based on our results, we recommend sampling multiple teeth per individual to improve the likelihood of successful *P. falciparum* recovery in case of poor DNA recovery after initial sampling and identification. Our results highlight how the processing of single samples do not reflect the overall preservation for target genomes across an individual. This study thus provides valuable insights into refining sampling strategies to optimize recovery of *P. falciparum* from ancient skeletal remains.

After analysing 41 libraries from the same individual, we have uncovered significant intra-individual variability in the presence of *Plasmodium* within a single individual or even tooth. This finding underscores the importance of employing a sampling strategy that includes multiple samples from the same individual for effective pathogen detection rather than relying solely on a single sample. Notably, After multiple sampling from dental pieces, most *Plasmodium* reads were concentrated in only seven libraries out of a total of 39. This study highlights the need for a larger sample to potentially unveil the finer preservation patterns of *Plasmodium* DNA. Additionally, our observations indicate that the quantity of *Plasmodium* DNA detected does not appear to correlate with the amount of host DNA present. For instance, tooth 6 (Lower molar left 3) yielded predominantly human DNA but minimal *Plasmodium* DNA (it yielded 44% of the entire human reads but below the 2% of *Plasmodium* ones). Our results suggest that the varying levels of *Plasmodium* DNA in our sample set may be due to differences in DNA preservation between teeth or to variations in parasite levels in the bloodstream peri-mortem, which could affect the likelihood of obtaining sufficient material from each tooth<sup>38</sup>. This underscores



**Fig. 2.** *P. falciparum* mtDNA maximum likelihood phylogeny. It is observable that the *P. falciparum* from Velia-186 (in blue) clusters with Indian strains and previously published ancient *P. falciparum* strains (in orange). Additional ancient strains of *P. falciparum* are in purple. The tree was visualised with TreeViewer<sup>37</sup>. We present a rooted phylogeny without branch lengths, and numbers represent bootstrap values.

the challenge of recovering nuclear DNA from *Plasmodium* in aDNA samples, suggesting that future sequencing efforts might require extensive sampling to be successful.

Our findings provide definitive evidence for the continuity of *P. falciparum* in Europe during ancient times<sup>5</sup>. Remarkably, the mitochondrial genome of Velia-186, originating from an individual nearly two millennia old,

shows a close phylogenetic relationship with almost all other ancient *P. falciparum* strains, especially from Europe and Asia spanning from the Neolithic to colonial times. This connection suggests the genetic continuity of the parasite in Europe over the past 2000 years. In a previous study on the same individual based on only a few sequenced fragments and low depth of coverage, Marciniak et al. (2016) reported 21 mutations between the ancient genome and the mitochondrial *P. falciparum* reference genome. Our high-coverage consensus sequence has seven mutations compared to the reference mitochondrial genome. When comparing our high-coverage SNP call with previously reported mutations for LV-13<sup>27</sup> we could only verify the presence of two of the mutations (2763, C → T; 3938, A → T). Additionally, compared to LV-13, this genome carries five novel mutations (74, A → T; 276 G → A; 725, C → T; 772, T → C; 2172, T → C), which are all supported by at least ten reads (see Table S2). The previously reported higher number of mutations, reported for LV-13, is probably the result of the low coverage of degraded DNA, which could have impacted the SNP call. Remarkably the report of two undescribed mutations, evidences that the diversity of past strains is still not fully described. Although we lack nuclear data, the published evidence strongly indicates that the recovered strain belongs to the endemic ancient *P. falciparum* of Eurasia. Future extensive sampling of the region over time could illuminate regional differences, enhance our understanding of the pathogen's fine structure in antiquity, and relate it to human movements and health conditions.

## Materials and methods

### Archaeological context

We sampled seven dental pieces from Individual Velia-186. Several bioarchaeological analyses have been carried out on individuals from this cemetery<sup>39–41</sup>, and *P. falciparum* was identified genetically in the individual we selected for downstream analyses<sup>27</sup>. This sample is from the 1st–2nd centuries CE.

Velia is located on a peninsula on the Tyrrhenian coast 112 km southeast of Naples and was incorporated into the Roman territory in the third century BCE. It became a port city utilised for the shipment of goods, boat maintenance, fish processing, and arboriculture. Subsistence strategies also included cultivation in the hinterland and well-watered intramural areas. The cemetery of Porta Marina (1st–2nd centuries CE) was investigated by Fiammenghi (2003) and led to the identification of approximately 330 burials (mostly inhumations)<sup>42–44</sup>. The human skeletal material is entrusted for anthropological study to the Bioarchaeology Service of the Museum of Civilizations based in Rome to reconstruct the funerary rituals and describe the demographic and bio-social characteristics of the ancient inhabitants of Velia within an interpretative framework guided by historical-archaeological evidence.

### Experimental model and subject details

Velia-186 is represented by a complete and well-preserved skeleton. Morphological and morphometric analyses led to a diagnosis of a young adult male (20–25 years). The values of oxygen and strontium isotopes indicate that the individual was likely born and raised in Velia<sup>45,46</sup>. The skeleton shows a robust morphology with mildly developed muscle attachments. The estimated living height (by Pearson regression formulas on the femur maximum length) is 166.9 cm, slightly above the average of the VELIA male series (164.7 ± 4.8 cm). The individual presents *cribra orbitalia* on both orbital roofs. The lesions (Type 3, according to Stuart-Macadam 1991<sup>47</sup> are in a remodelling phase<sup>40</sup>.

### Laboratory procedures

100 mg was ground from multiple locations of different teeth and bone material including dentition pieces and other anatomical locations (Table S1A and S1C), and all procedures were done with permission from Museo delle Civiltà. DNA was extracted from bone powder in the Ancient DNA (aDNA) Laboratory at the University of Vienna following detailed protocols adapted to aDNA<sup>48</sup>, and single-stranded libraries were prepared<sup>49</sup>. Libraries were enriched with an in-solution capture designed by Daicel Arbor Biosciences (<https://arborbiosci.com/products/targeted-ngs/mybaits-custom-kits/mybaits-custom-dna-seq/>) following the manufacturer manual. The solution kit included baits targeting the mtDNA sequences of both *P. falciparum* and *P. vivax*. Libraries were pooled in 20 µl and sequenced 2 × 150 PE on NextSeq 550 at the Polo d'Innovazione di Genomica Genetica e Biologia (Siena, Italy).

### Bioinformatics

Sequenced reads were clipped with cutadapt 4.5<sup>50</sup>, removing Illumina adapters and base quality < 30 and read length < 30, and later collapsed using PEAR version 0.9.11<sup>51</sup>, requiring a minimum overlap of 11 bp and a minimum length of 30 bp. Filtered collapsed reads were aligned with a competitive mapping of *P. vivax*, *P. falciparum* and human mitochondrion (Salvador 1 [ID: NC\_007243.1], K1 [ID: NC\_037526] and H. sapiens [ID: NC\_012920.1], respectively) using Burrows-Wheeler Aligner (BWA) 0.7.17 aln<sup>52</sup>, whereby aln stands for the alignment mode and enables mapping of short DNA sequences. Also we disabled seeding to allow a higher sensitivity of aDNA reads<sup>53</sup>, gap open penalty of 2 and edit distance of 0.04. Next, the mapped reads were deduplicated using the MarkDuplicates program in Picard-tools 3.1.1<sup>54</sup> and low mapping-quality reads were removed using the flag -q 30 in Samtools 1.9<sup>55</sup>. We calculated the deamination pattern using MapDamage<sup>56</sup>; mapping plots were created using aDNA-BAMplotter<sup>57</sup>. We used this competitive mapping between *P. falciparum* and *P. vivax* to determine the possible presence of co-infection. We used Samtools 1.9<sup>52</sup> to convert the bam files into fasta files that were screened with BLASTn<sup>30</sup> 2.15.0 using the whole NCBI nt database (2023-08-17). The results were uploaded to MEGAN 6.25.9<sup>31</sup> and processed with the LCA algorithm.

A consensus sequence was obtained using ANGSD (v0.941)<sup>58</sup>, which used the majority of calls, and excluding sites with coverage lower than 5, all the SNPs were fixed. The detected mutations were examined visually with IGV<sup>59</sup>. The consensus sequence was aligned with multiple *P. falciparum* sequences<sup>6–8,35,60–63</sup> with MAFFT

(v7.520)<sup>64</sup>. For the phylogeny, the alignment was filtered by positions shared by 90% coverage to eliminate incorrectly assembled regions from the analysed sequences, which left 5967 sites, including 230 variable sites. The TPM2u + F + I substitution model was chosen with IQTree2 (v 2.2.6) using the ModelFinder algorithm. Due to alignments containing the whole sequence with variant and invariant sites, no ascertainment bias corrections were needed. We constructed a Maximum likelihood tree with 1000 nonparametric bootstrap replicates. The tree was visualized via *TreeViewer*<sup>37</sup>.

### Human mtDNA analyses

We aligned the reads against the human genome (hg19), including the mtDNA sequence (rCRS), using the same procedures previously described. We recovered the consensus sequence using ANGSD<sup>58</sup> and identified the haplogroup using Haplogrep 3.2.2.1<sup>65</sup>.

### Data availability

Sequencing data and the filtered sequences are available at the European Nucleotide Archive (ENA) under the accession number PRJEB72667.

Received: 15 August 2024; Accepted: 30 December 2024

Published online: 04 January 2025

### References

- Hupalo, D. N. et al. Population genomics studies identify signatures of global dispersal and drug resistance in *Plasmodium Vivax*. *Nat. Genet.* **48**, 953–958 (2016).
- Otto, T. D. et al. Genomes of all known members of a *Plasmodium* Subgenus reveal paths to virulent human malaria. *Nat. Microbiol.* **3**, 687–697 (2018).
- Carter, R. Speculations on the origins of *Plasmodium Vivax* malaria. *Trends Parasitol.* **19**, 214–219 (2003).
- Sallares, R., Bouwman, A. & Anderung, C. The spread of malaria to Southern Europe in antiquity: New approaches to old problems. *Med. Hist.* **48**, 311–328 (2004).
- Michel, M. et al. Ancient plasmodium genomes shed light on the history of human malaria. *Nature* **631**, 125–133 (2024).
- Gelabert, P. et al. Mitochondrial DNA from the eradicated European *Plasmodium Vivax* and *P. Falciparum* from 70-year-old slides from the Ebro Delta in Spain. *Proc. Natl. Acad. Sci. U S A.* **113**, 11495–11500 (2016).
- Rodrigues, P. T. et al. Human migration and the spread of malaria parasites to the New World. *Sci. Rep.* **8**, 1993 (2018).
- Joy, D. A. et al. Early origin and recent expansion of *Plasmodium Falciparum*. *Science* **300**, 318–321 (2003).
- Piperaki, E. T., Manguin, S. & Dev, V. in *Malaria Eradication in the European World: Historical Perspective and Imminent Threats*. 315–335 (eds Manguin, S. & Dev, V.) (Towards Malaria Elimination, 2018).
- Arrow, K., Panosian, C. & Gelband, H. Committee on the Economics of Antimalarial Drugs. of Malaria Drugs in an Age of &#8230.
- World Health Organization. World Malaria Report 2023. (2023).
- Jones, W. H. S. *Malaria. A Neglected Factor in the History of Greece and Rome* (Macmillan, 1907).
- Smith-Guzmán, N. E. The skeletal manifestation of malaria: an epidemiological approach using documented skeletal collections. *Am. J. Phys. Anthropol.* **158**, 624–635 (2015).
- Gowland, R. L. & Western, A. G. Morbidity in the marshes: using spatial epidemiology to investigate skeletal evidence for Malaria in Anglo-Saxon England (AD 410–1050). *Am. J. Phys. Anthropol.* **147**, 301–311 (2012).
- Piperaki, E. T. Malaria eradication in the European World: historical perspective and imminent threats. in *Towards Malaria Elimination* (eds Manguin, S. & Dev, V.) (IntechOpen, Rijeka, (2018).
- Schats, R. Developing an archaeology of malaria. A critical review of current approaches and a discussion on ways forward. *Int. J. Paleopathol.* **41**, 32–42 (2023).
- Sallares, R. & Gomzi, S. Biomolecular archaeology of malaria. *Anc. Biomol.* **3**, 195–213 (2001).
- Lalremruata, A. et al. Molecular identification of *falciparum* malaria and human tuberculosis co-infections in mummies from the Fayum depression (Lower Egypt). *PLoS One.* **8**, e60307 (2013).
- Loufouma Mbouaka, A. et al. The elusive parasite: Comparing macroscopic, immunological, and genomic approaches to identifying malaria in human skeletal remains from Sayala, Egypt (third to sixth centuries AD). *Archaeol. Anthropol. Sci.* **13**, 115 (2021).
- Maixner, F. et al. Microscopic Evidence of Malaria Infection in visceral tissue from Medici Family, Italy. *Emerg. Infect. Dis.* **29**, 1280–1283 (2023).
- Tishkoff, S. A. et al. Haplotype diversity and linkage disequilibrium at human G6PD: Recent origin of alleles that confer malarial resistance. *Science* **293**, 455–462 (2001).
- Martiniano, R. et al. Ancient genomes illuminate eastern arabian population history and adaptation against malaria. *Cell. Genom.* **4**, 100507 (2024).
- Gelabert, P., Olalde, I., de-Dios, T., Civit, S. & Lalueza-Fox, C. Malaria was a weak selective force in ancient europeans. *Sci. Rep.* **7**, 1377 (2017).
- van Dorp, L. et al. *Plasmodium Vivax* Malaria viewed through the Lens of an eradicated European strain. *Mol. Biol. Evol.* **37**, 773–785 (2020).
- de-Dios, T. et al. Genetic affinities of an eradicated European *Plasmodium falciparum* strain. *Microb. Genom.* **5**, (2019).
- Antonio, M. L. et al. Stable population structure in Europe since the Iron Age, despite high mobility. *Elife* **13**, (2024).
- Marciniak, S. et al. *Plasmodium Falciparum* malaria in 1st–2nd century CE southern Italy. *Curr. Biol.* **26**, R1220–R1222 (2016).
- Spyrou, M. A., Bos, K. I., Herbig, A. & Krause, J. Ancient pathogen genomics as an emerging tool for infectious disease research. *Nat. Rev. Genet.* **20**, 323–340 (2019).
- Talman, A. M., Domarle, O., McKenzie, F. E., Ariey, F. & Robert, V. Gametocytogenesis: the puberty of *Plasmodium Falciparum*. *Malar. J.* **3**, 24 (2004).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* **17**, 377–386 (2007).
- Lee, E. et al. Web Apollo: a web-based genomic annotation editing platform. *Genome Biol.* **14**, R93 (2013).
- The *Plasmodium* Genome Database Collaborative. PlasmoDB: an integrative database of the *Plasmodium falciparum* genome. Tools for accessing and analyzing finished and unfinished sequence data. The *Plasmodium* Genome Database Collaborative. *Nucleic Acids Res.* **29**, 66–69 (2001).

34. Schmedes, S. E., Patel, D., Kelley, J., Udhayakumar, V. & Talundzic, E. Using the Plasmodium mitochondrial genome for classifying mixed-species infections and inferring the geographical origin of *P. falciparum* parasites imported to the U.S. *PLoS One*. **14**, e0215754 (2019).
35. Tyagi, S. & Das, A. Mitochondrial population genomic analyses reveal population structure and demography of Indian Plasmodium falciparum. *Mitochondrion* **24**, 9–21 (2015).
36. Agranat-Tamir, L. et al. The genomic history of the Bronze Age Southern Levant. *Cell* **181**, 1146–1157e11 (2020).
37. Bianchini, G., Sánchez-Baracaldo, P. & TreeViewer Flexible, modular software to visualise and manipulate phylogenetic trees. *Ecol. Evol.* **14**, e10873 (2024).
38. Nguyen, T. N. et al. The persistence and oscillations of submicroscopic Plasmodium falciparum and Plasmodium Vivax infections over time in Vietnam: An open cohort study. *Lancet Infect. Dis.* **18**, 565–572 (2018).
39. Sperduti, A. et al. Dental twinning in the primary dentition: new archaeological cases from Italy. *Bull. Int. Association Paleodontology*. **15**, 6–20 (2021).
40. Sperduti, A., Bondioli, L. & Garnsey, P. Skeletal evidence for occupational structure at the coastal towns of Portus and Velia (1st–3rd c. AD). More than just numbers (2012).
41. Scheidel, W. *The Science of Roman History: Biology, Climate, and the Future of the Past* (Princeton University Press, 2018).
42. Fiammenghi, C. A. & La Torre, C. *Ledificio Funerario Numero 1 Dalla Necropoli Di Porta Marina Sud Di Velia* (Festschrift für Friedrich Krinzinger, 2005).
43. Fiammenghi, C. A. La Necropoli Di Elea-Velia: qualche osservazione preliminare. Elea-Velia. Le Nuove Ricerche. *Quad. Del. Centro Studi Magna Grecia*. **1**, 49–61 (2003).
44. Craig, O. E. et al. Stable isotopic evidence for diet at the imperial roman coastal site of Velia (1st and 2nd centuries AD) in Southern Italy. *Am. J. Phys. Anthropol.* **139**, 572–583 (2009).
45. Stark, R. J. et al. Imperial Roman mobility and migration at Velia (1st to 2nd c. CE) in southern Italy. *J. Archaeol. Sci. Rep.* **30**, 102217 (2020).
46. Stark, R. J. et al. Dataset of oxygen, carbon, and strontium isotope values from the Imperial Roman site of Velia (ca. 1st–2nd c. CE), Italy. *Data Brief* **38**, 107421 (2021).
47. Stuart-Macadam, P. Porotic hyperostosis: changing interpretations. *Hum. Paleopathology: Curr. Syntheses Future Options* 36–39 (1991).
48. Dabney, J. & Meyer, M. Extraction of highly degraded DNA from ancient bones and Teeth. In *Ancient DNA: Methods and Protocols* (eds Shapiro, B. et al.) 25–29 (Springer New York, New York, NY, (2019).
49. Kapp, J. D., Green, R. E. & Shapiro, B. A. Fast and efficient single-stranded genomic Library Preparation Method optimized for ancient DNA. *J. Hered.* **112**, 241–249 (2021).
50. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
51. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: A fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics* **30**, 614–620 (2014).
52. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
53. Schubert, M. et al. Improving ancient DNA read mapping against modern reference genomes. *BMC Genom.* **13**, 178 (2012).
54. Broad Institute. Picard Toolkit. Broad Institute, GitHub repository. Available at: <http://broadinstitute.github.io/picard> (accessed 20 Mar. 2024).
55. Li, H. et al. The sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
56. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: Fast approximate bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
57. Guellil, M. aDNA-BAMPlotter. (2021). <https://doi.org/10.5281/zenodo.5702679>
58. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of next generation sequencing data. *BMC Bioinform.* **15**, 356 (2014).
59. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
60. Pacheco, M. A., Cranfield, M., Cameron, K. & Escalante, A. A. Malarial parasite diversity in chimpanzees: the value of comparative approaches to ascertain the evolution of Plasmodium falciparum antigens. *Malar. J.* **12**, 328 (2013).
61. Tyagi, S., Pande, V. & Das, A. New insights into the evolutionary history of Plasmodium Falciparum from mitochondrial genome sequence analyses of Indian isolates. *Mol. Ecol.* **23**, 2975–2987 (2014).
62. Li, Y. et al. Geographical origin of Plasmodium Vivax in the Hainan Island, China: Insights from mitochondrial genome. *Malar. J.* **22**, 84 (2023).
63. Conway, D. J. et al. Origin of Plasmodium Falciparum malaria is traced by mitochondrial DNA. *Mol. Biochem. Parasitol.* **111**, 163–171 (2000).
64. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
65. Weissensteiner, H. et al. HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–63 (2016).

## Acknowledgements

The study was funded by the Austrian Science Fund (FWF) [<https://doi.org/10.55776/P36433>] and INFRAVEC ISID\_2019 to P. G., the Vienna Science and Technology Fund (WWTF) [<https://doi.org/10.47379/VRG20001>], and the Austrian Science Fund (FWF) [<https://doi.org/10.55776/TAI729>] to M.K. A. L-L. was funded by Colfuturo (Fundación para el futuro de Colombia) during his MSc. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## Author contributions

P. G. conceptualised the study. A. S. sampled and provided archeo-anthropological context. A. L-L, O. C, K.T. O, S. S. and B.Z performed the experiments, P. G., M. G., A. L-L., and M. H. analysed the data, P. G., A. L-L., R. P, A. S, M. K, M. G, S.S. and M. H. wrote the text with inputs from all collaborators.

## Declarations

## Competing interests

The authors declare no competing interests.



### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-85038-z>.

**Correspondence** and requests for materials should be addressed to R.P. or P.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025