



OPEN

## Bioarchaeological and palaeogenomic portrait of two Pompeians that died during the eruption of Vesuvius in 79 AD

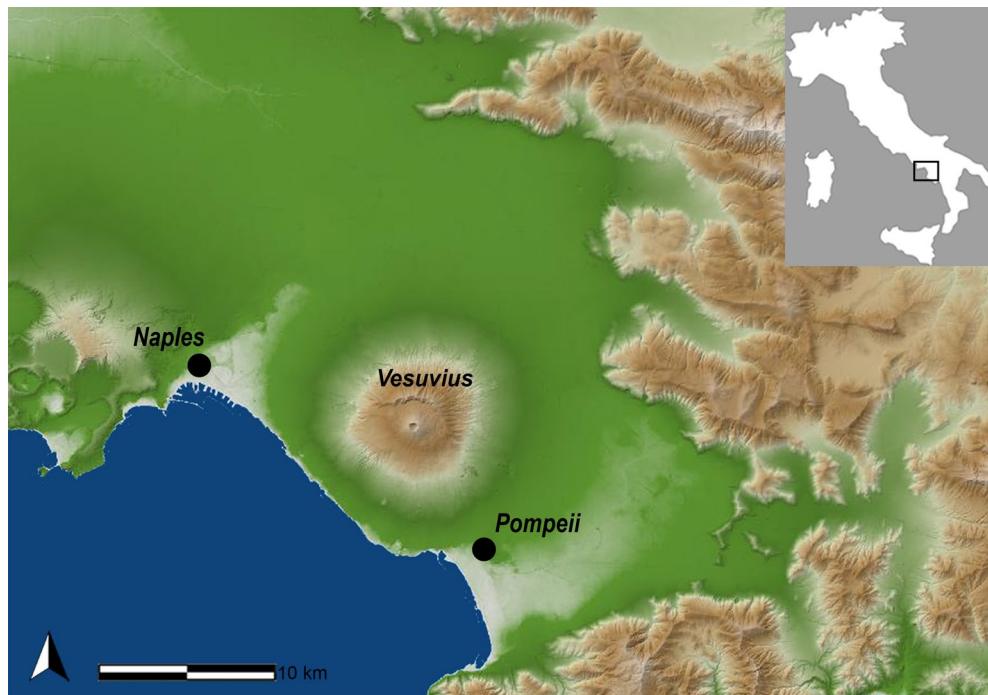
Gabriele Scorrano<sup>1,2</sup>✉, Serena Viva<sup>3</sup>, Thomaz Pinotti<sup>2,4</sup>, Pier Francesco Fabbri<sup>3</sup>✉, Olga Rickards<sup>1,6</sup> & Fabio Macciardi<sup>5,6</sup>✉

The archaeological site of Pompeii is one of the 54 UNESCO World Heritage sites in Italy, thanks to its uniqueness: the town was completely destroyed and buried by a Vesuvius' eruption in 79 AD. In this work, we present a multidisciplinary approach with bioarchaeological and palaeogenomic analyses of two Pompeian human remains from the *Casa del Fabbro*. We have been able to characterize the genetic profile of the first Pompeian genome, which has strong affinities with the surrounding central Italian population from the Roman Imperial Age. Our findings suggest that, despite the extensive connection between Rome and other Mediterranean populations, a noticeable degree of genetic homogeneity exists in the Italian peninsula at that time. Moreover, palaeopathological analyses identified the presence of spinal tuberculosis and we further investigated the presence of ancient DNA from *Mycobacterium tuberculosis*. In conclusion, our study demonstrates the power of a combined approach to investigate ancient humans and confirms the possibility to retrieve ancient DNA from Pompeii human remains. Our initial findings provide a foundation to promote an intensive and extensive paleogenetic analysis in order to reconstruct the genetic history of population from Pompeii, a unique archaeological site.

The archaeological site of Pompeii is one of the 54 UNESCO World Heritage sites in Italy. Pompeii was a Roman Imperial Age port city located south of Naples in Central Italy (Fig. 1) until it was completely destroyed and buried by the ashes of the Mount Vesuvius' eruption in 79 AD<sup>1,2</sup>. According to Gaius Plinius Caecilius Secundus (better known as Pliny the Younger: a lawyer, author, and magistrate of Ancient Rome), the Vesuvius' eruption occurred around 1 p.m. on the 24th of August and was visible from over 40 km away. More than 2000 individuals died as a direct consequence of the eruption<sup>1</sup>, the deadliest ever in European history. The several exceptionally well-preserved buildings found in Pompeii such as *Casa del Chirurgo* (House of the Surgeon), *Casa del Fauno* (House of Faun) and the *Casa dei Casti Amanti* (House of the Chaste Lovers) suggest that Pompeii was probably a holiday resort for wealthy Romans. However, Pompeii was also an important city for trading and business, with a population ranging between 6400 and 20,000 dwellers.

Despite continuing an intense scientific research on the site since the nineteenth century to this day, conducting both bioarchaeological and genetic studies from Pompeian human remains has been a challenge, as exposure to high temperature effectively destroys the bone matrix, altering the structure of the bioapatite<sup>3</sup> and diminishing the quality and quantity of recoverable DNA<sup>4</sup>. On the other hand, it is also possible that the pyroclastic materials that covered the remains could have shielded them from environmental factors, like the atmospheric oxygen, that degrades DNA<sup>5</sup>.

<sup>1</sup>Centre of Molecular Anthropology for Ancient DNA Studies, Department of Biology, University of Rome "Tor Vergata", 00133 Rome, Italy. <sup>2</sup>Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen, Copenhagen, Denmark. <sup>3</sup>Department of Cultural Heritage, University of Salento, 73100 Lecce, Italy. <sup>4</sup>Laboratório de Biodiversidade e Evolução Molecular (LBEM), Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. <sup>5</sup>Laboratory of Molecular Psychiatry, Department of Psychiatry and Human Behavior, University of California, Irvine, CA 92868, USA. <sup>6</sup>These authors contributed equally: Olga Rickards and Fabio Macciardi. ✉email: gabrielescor@gmail.com; pierfrancesco.fabbri@unisalento.it; fmacciardi@uci.edu



**Figure 1.** Geographic location of the Pompeii site, Campania (Italy). Map source: SINAnet ISPRA – Dem75 (QGIS 3.22 ‘Biatowieza’) <https://www.qgis.org/it/site/>.

Past studies have shown the possibility of retrieving genetic data from both human and zooarchaeological remains in Pompeii<sup>6–12</sup>, but those initial analyses were limited to short stretches of mitochondrial DNA obtained using PCR-based methods. New available methodologies, based on high-throughput shotgun sequencing, DNA capture, enrichment strategy<sup>13,14</sup>, as well as using optimal sources of ancient DNA (aDNA) from teeth and petrous bones<sup>15,16</sup>, have dramatically increased the amount of data that can be obtained from previously unsuitable samples for genetic research, and may open new avenues to substantially increase the knowledge of the genetic diversity in the ancient Pompeian population.

In this work, we present a multidisciplinary approach with bioarchaeological and palaeogenomic analyses of two human remains from the *Casa del Fabbro* (House of the Craftsman: Supplementary Fig. S1) from Pompeii. The successful recovery of aDNA from one individual enabled us to reconstruct its genetic history and to investigate the presence of blood-borne pathogens, alongside skeletal biology evidence. Furthermore, this data can also give us an overview of the genetic diversity outside of Rome during the Roman Empire.

**Individuals analysed from Casa del Fabbro.** The analysed human remains came from Room 9 of the *Casa del Fabbro* (Regio I, Insula 10, civic 7), and their position and orientation are compatible with instantaneous death due to the approach of the high-temperature volcanic ash cloud<sup>17</sup>. More than half of individuals found in Pompeii died inside their houses, indicating a collective unawareness of the possibility of a volcanic eruption or that the risk was downplayed due to the relatively common land tremors in the region<sup>2</sup>. Both skeletons have been discovered in anatomical position. They were both leaning on a low relief in a corner of what probably was the dining room, on the remnants of a *triclinium*, a sort of couch or *chaise longue* used in Roman buildings during meals. Individual A was in left lateral recumbent position with flexed limbs, with the left arm and leg on the ground and right limbs on the *triclinium*. The individual B had the arms gathered in front of the skull and legs on the ground flexed on the right side, with the back leaning against the *triclinium*.

## Results

The two *Casa del Fabbro* individuals (Supplementary Fig. S1) underwent osteological examination to establish their respective sex, estimated height, and approximate age-at-death. Individual A was a male between 35 and 40 years-old and stood 164.3 cm tall. Individual B was a female over 50 years of age who stood 153.1 cm tall. Estimated heights were obtained through averaging two methods<sup>18–21</sup> and are consistent with Roman Age height averages (male: 164.4 cm; female: 152.1 cm)<sup>22</sup> as well as Pompeii and Herculaneum height averages<sup>23,24</sup> (see Supplementary Table S1).

We extracted and sequenced DNA from a petrous bone from each Pompeian individual with identical procedures (see “Materials and methods”) and obtained 0.4 X (individual A) and 0.0013 X (individual B) average genome-wide depth of coverage (Table 1). Both individuals displayed typical signatures of aDNA<sup>25</sup> (Supplementary Fig. S2). Low rates of contamination with modern human DNA (0.8% for the mtDNA and between 0.6 and 0.8% for the X chromosome, Table 1 and Supplementary Figs. S2, S3) for individual A confirmed the authenticity

Individuals #	Sample code #	Total	Uniq	Rmdup	Clonality%	Endo%	Coverage	Number of identified SNPs	mtDNA coverage	mtDNA haplogroup	Y-chromosome coverage	Y-chromosome haplogroup	mtDNA contamination (%)	X coverage	X chromosome contamination (%)	Sex determination
Pompei casa del Fabbro individual_A	f_1R	67,481,814	22,071,517	17,340,415	21.43	37.40	0.42	450,751	41.42	HV0a (85.3%)	0.074	A-M13	0.8 (15.3; 0.2)	0.25	0.8–0.6	XY
Pompei casa del Fabbro individual_B	f_11R	156,203,139	42,007	40,123	4.48	0.017	0.0013	/	/	/	/	/	/	/	/	/

**Table 1.** Statistic parameters and contamination results of the individuals analysed. Total is the total number of reads per library; Unique is the number of sequences mapping uniquely to the human reference; Rmdup is the unique sequences without the duplicate; Endogenous (%) is the proportion of human sequences after trimming; genome-wide; we also reported the number of identified SNPs, the mitochondrial, X and Y-chromosome average depth of coverage.

of the aDNA. The low coverage obtained for individual B prevented us from reaching any further assessments of quality parameters. Herein, we report the details of the analyses only for individual A.

**Sex determination and uniparental genetic markers.** The genetic sex determination (estimated by  $R_Y$  parameter<sup>26</sup>, and by the X chromosome coverage) confirmed the morphological determination that individual A was a male (Table 1).

The mitochondrial DNA haplogroup was identified using HaploGrep2<sup>27,28</sup> (Table 1), and revealed that the individual belongs to the haplogroup clade HV0a, the main monophyletic branch of HV0 and subclade of haplogroup HV. This mitochondrial lineage is absent among published Roman Imperial individuals from Italy<sup>29</sup>. In Europe, the first evidence of the HV haplogroup is from a Magdalenian period individual from Spain<sup>30</sup> while in Italy from a Mesolithic individual from Sicily (Favignana)<sup>31</sup>. The HV haplogroup is actually associated with the early human dispersal in Eurasia after the Last Glacial Maximum (LGM)<sup>32</sup>. It is unevenly spread across Europe with highest frequencies in the Near East (~11%)<sup>33</sup>, in south Europe (from ~4% to ~11%)<sup>34</sup> and in the Balkan peninsula (~8%)<sup>35</sup>. HV0a coalesces around 12.5–11.0 kya ago<sup>34</sup> and, among the extant populations, is common in Sardinia<sup>36</sup>.

Individual A, albeit at low coverage (Table 1), was found to belong to the Y-chromosome lineage A-M13 (A1b1b2b), a rare lineage absent among ancient individuals from the Italian Peninsula<sup>29</sup>, mainly found in Eastern Africa (~40%), but with known occurrences, at much lower frequencies, in the Near East (Turkey, Yemen, Egypt, Palestine, Jordan, Oman and Saudi Arabia) and the Mediterranean islands of Sardinia, Cyprus and Lesbos<sup>37–40</sup>. Downstream of A-M13 and restricting the analysis to transversions polymorphisms, the individual can be placed at A-V5880, a sub-haplogroup that contains all A-M13 positive Sardinians from past studies<sup>39,40</sup>, and that has been dated to coalesce around 7.62 ( $\pm 0.92$ ) thousands of years ago, using Bayesian analysis<sup>40</sup>.

**Pompeian genetic structure.** To understand the relationship of the higher-coverage ancient Pompeian individual A, we assembled a dataset of relevant previously published ancient populations from Upper Palaeolithic to Medieval periods<sup>13–15,29,41–66</sup> (genotyped on the “1240K” SNP panel—detailed in Supplementary Table S2) combined with 471 present-day West Eurasian individuals (subset panel of the Human Origins)<sup>56</sup>, which was used on all subsequent analyses.

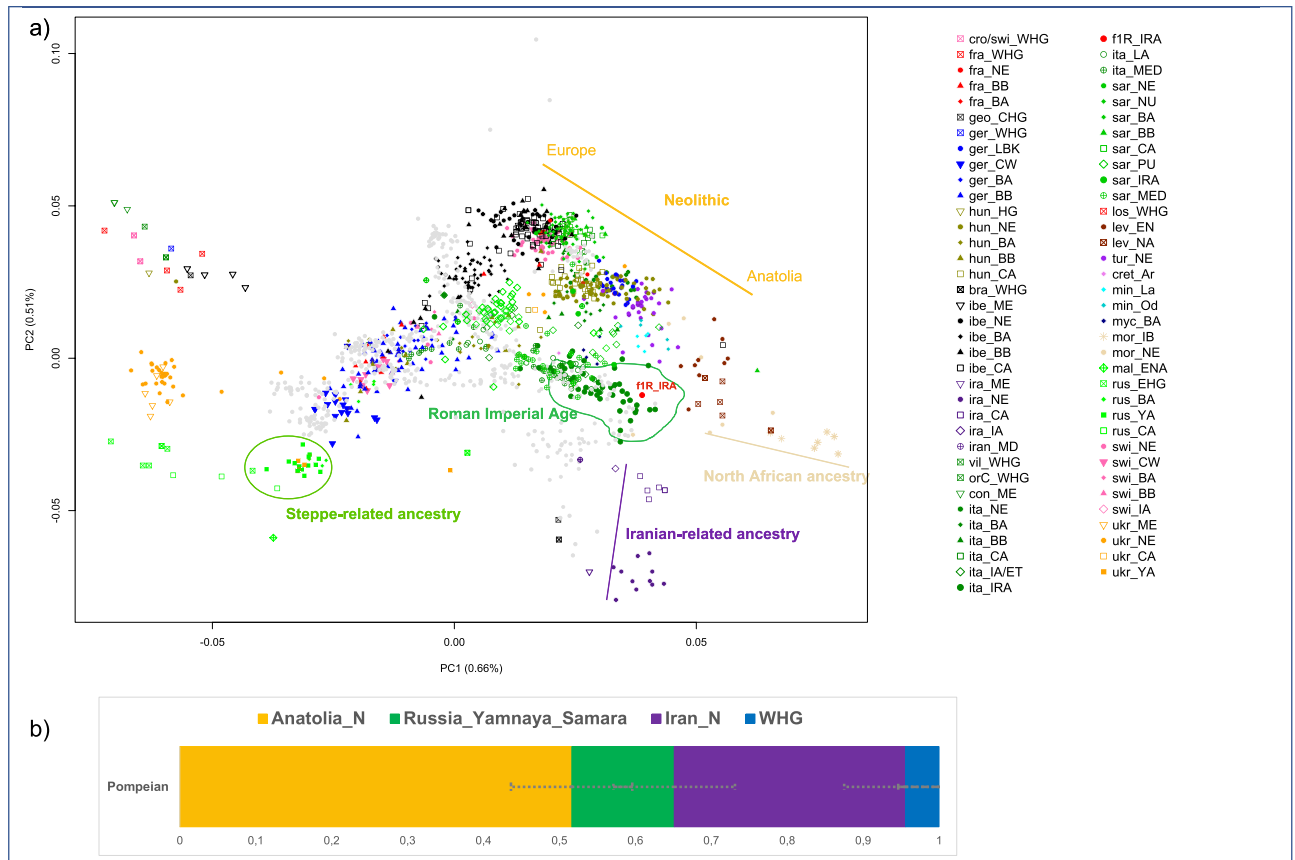
Using the EIGENSOFT package<sup>67</sup>, we performed principal component analysis (PCA), and according to the results the Pompeian individual A clustered with other Italian Imperial Roman Age individuals and is positioned close to the well-documented Neolithic cline of Anatolians to European populations<sup>58</sup> (Fig. 2a).

These results can be formally tested using a  $D$ -statistics of the form  $D(\text{Mbuti, Test; Pompeian, Italy\_IRA})$ , which gauges whether the Pompeian individual A forms a clade with other Central Italian Imperial Roman Age individuals<sup>29</sup>, to the exclusion of other test populations (Supplementary Table S4). With the exception of Bronze Age Iberia, which shares a high genetic drift with Italy Imperial Roman Age ( $Z = 3.33$ ), for all the other populations considered we cannot reject a clade-like relationship for the Pompeian and Imperial Roman Age cluster (Supplementary Table S4).

Furthermore, using statistics in the form  $D(\text{Mbuti, Test; Pompeian, Russia\_MA1\_HG})$ , we tested which other population showed high affinity to the Pompeian individual, using a 24 thousand year old individual from Russia as an outgroup (Mal'ta)<sup>65</sup>. Populations with  $Z$ -score values lower than  $-3$  represent statistically significant results of excess shared drift with Pompeian individual A. Among them, Neolithic Anatolian (Anatolia\_N) had the highest scores, with  $Z = -9.57$  (Supplementary Table S5, Fig. 3).

The position of the Pompeian individual A in the PCA (Fig. 2a) falls also close to the distribution of modern Mediterranean and Near Eastern populations, such as Greeks, Maltese, Cypriots, and Turks. Such a result allowed us to hypothesize a genetic contribution from the Near East. This hypothesis can be also supported by a cline from Neolithic Iran (Iran\_N) to Italy Imperial Roman Age (Italy\_IRA, including the Pompeian individual) passing through Chalcolithic Iran (Iran\_CA) and Iron Age Iran (Iran\_IA) (Fig. 2a). The presence of Iranian-related ancestry has been identified in Italy since the Neolithic period, with a reported increase in this component in Central Italy during the Roman Imperial Age compared to the previous Iron Age period<sup>29,66</sup>. However, when performing the same four-population test, but using the Pompeian individual instead of Imperial Age Romans from Rome, the result is statistically non-significant (Supplementary Table S6), indicating that in individual A of Pompeii no further contribution by Iranian-related ancestry occurred after the Iron Age.

From the distribution of individuals obtained with our PCA analysis (Fig. 2a), it is also possible to recognize a cline from Morocco\_Iberomaurusian to Italy Imperial Roman Age (Italy\_IRA) passing through Morocco



**Figure 2.** (a) PCA on selected 471 present-day west Eurasians (gray dots) and projected 1030 ancient individuals plotted by R version 3.6.2 (<https://www.r-project.org/>): the Pompeian individual A is in red and labelled with f1R\_IRA; (b) four-way qpAdm models with source populations Anatolia Neolithic, Russian\_Yamnaya\_Samara, Iranian\_N and Western hunter-gatherers (WHG). Error bars represent  $\pm 1$  standard errors of the proportion of each component. The complete results are reported in Supplementary Table S3.

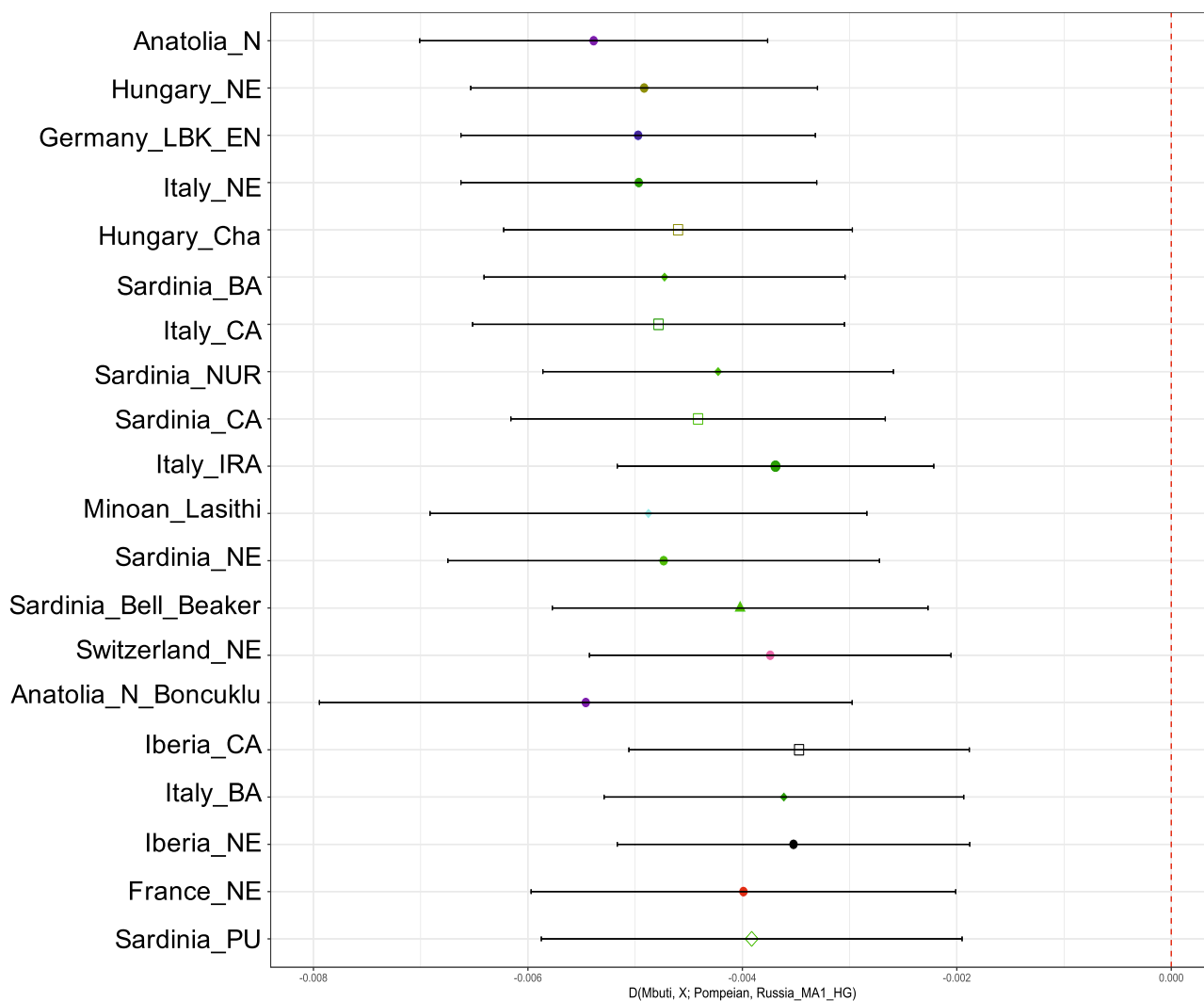
Neolithic after the Iron Age. The genetic contribution derived from a North African source is already evident in the Italian prehistory. Indeed, admixture of a North African ancestry was recognized in Sardinia since the Chalcolithic<sup>48</sup> and in central Italy since the Iron Age (Etruscan)<sup>29,66</sup> and continued into the Roman Imperial period<sup>29</sup>. Nevertheless, we did not identify any North African ancestry contribution in the Pompeian individual using *D*-statistics (Supplementary Table S6). Another significant genetic component in most post-Bronze Age European populations comes from a source ultimately deriving from the Eurasian Steppe<sup>41,68</sup>, and has been attested in the Italian Peninsula in Iron Age Italy<sup>29,69</sup>, Bronze Age Sicily<sup>45</sup> as well as in the Pompeian individual A.

To confirm these findings, we attempted to fit the Pompeian as either a three- or four-way combination of Anatolian Neolithic, Russian Yanmaya, Iranian Neolithic and Western Hunter-Gatherers (WHG) using *qpAdm*<sup>70</sup> (Supplementary Table S3). We set a minimum threshold of 100,000 SNPs and only considered results when  $p > 0.05$ .

Three-way admixtures that fit the data always included both ancestries from Anatolian and Iranian Neolithic, with varying contributions from the Steppe and WHG (Supplementary Table S3). Four-way models displayed major contributions of Anatolian Neolithic ( $51.6 \pm 7.8\%$ ) and Iran Neolithic ( $30.5 \pm 8.1\%$ ) in comparison to Steppe-related ancestry ( $13.5 \pm 8.0\%$ ) and WHG ( $4.4 \pm 5.4\%$ ) components (Fig. 2b, Supplementary Table S3).

We further attempted to investigate if the fit could be improved by including Morocco\_Iberomaurusian as a fifth source<sup>46</sup>, but no three-, four- or five-way mixture with this last component produced significant results (Supplementary Table S3).

**Mycobacterium tuberculosis.** A palaeopathological study carried out on the Pompeian individual diagnosed spinal tuberculosis (Pott's disease) (Table 2) on the basis of diagnostic morphological markers such as a large lytic destruction on the upper anterior half of the fourth lumbar vertebra L4 (Fig. 4). Moreover, the digital radiograph analysis shows erosion in the antero-superior portion of the vertebral body, with a reduced downward cortical rim and a bowl-shaped appearance (Fig. 4). All the other vertebral osteitis that can cause similar lesions (pyogenic osteomyelitis, actinomycosis, metastatic neoplasms, osteoporosis, brucellosis and extrapulmonary tuberculosis<sup>71</sup>) have been excluded for the following reasons.



**Figure 3.** Point estimates and  $\pm 3$  standard errors for the top twenty populations with significantly ( $Z < 3$ ) more allele sharing with the ancient Pompeian in comparison to 24 ky old Mal'ta based on the statistic  $D(\text{Mbuti, Test; Pompeian, Russia\_MA1\_HG})$ . All results can be found in Supplementary Table S5.

Lesion	Individual A	Actinomycosis	Tuberculosis	Brucellosis	Osteoporosis	Metastatic neoplasia	Pyogenic osteomyelitis
Spinous process and neural arch are involved <sup>71,73</sup>	No	No	No	No	Yes	Yes	Yes
Large and spheroid lesions surrounded by reactive new bone <sup>72</sup>	No	Yes	No	No	No	No	No
First affects the peduncles <sup>70</sup>	No	No	No	No	No	Yes	No
Vertebral fracture <sup>73</sup>	No	No	Possible	No	Yes	No	No
Demineralization <sup>73</sup>	No	No	No	No	Yes	No	No
Vertebral bodies in lumbar tract affected <sup>71,73,74</sup>	Yes	No	Yes	Yes	Yes	No	Yes
Collapse of vertebral bodies and angular deformity <sup>74</sup>	Yes	No	Yes	Yes	No	No	No
At least two adjacent vertebrae affected in most of cases <sup>71</sup>	nd	No	Yes	No	No	No	Yes
Sign of Pedro-i-Pons <sup>74</sup>	No	No	No	Yes	No	No	No

**Table 2.** Differential diagnosis of tuberculosis in individual A from “Casa del Fabbro”. nd: characteristic not determinable.



**Figure 4.** Photography and digital radiograph of the fourth lumbar vertebra (L4) affected by tuberculous spondylodiscitis of the individual A.

Pyogenic osteomyelitis may cause lytic lesions of vertebral bodies, with a predilection for the lumbar spine; however, the spinous process and neural arch are often involved, and lesions often show marked new bone proliferation<sup>71</sup>. Actinomycosis is characterized by large and spheroid lesions surrounded by reactive new bone; the vertebral column is affected in a way quite different from that of any other form of infection<sup>72</sup>. Metastatic neoplasia first affects the peduncles, neural arches and spinous processes with circular lytic lesions, and later the vertebral body. Osteoporosis, the most frequent demineralizing disease, can produce vertebral fracture and collapse due to bone demineralization<sup>73</sup> and normally affects aged people. The brucellar spondylitis, a consequence of brucellosis, a highly contagious zoonosis caused by ingestion of milk or infected meat, has an evolution similar to Pott's disease (extrapulmonary tuberculosis), so much so as to be called pseudo-Pott. It involves destructive vertebral lesions and formation of ossifluent abscesses. Moreover, at digital radiograph analyses, brucellar spondylitis shows disseminated decalcification, irregular erosions at the vertebral edges with the typical destruction of the upper anterior angle of one or more contiguous vertebrae, reduction of the intervertebral space up to the fusion between two contiguous vertebrae and the Pedro-i-Pons sign<sup>74</sup> (osteosclerotic semicircle around the osteolysis area of the upper antero-vertebral angle, typical of brucellar spondylitis). The digital radiography image shows no sign of Pedro-i-Pons (Fig. 4), and the case fully matches the palaeopathological and radiological criteria for tuberculosis proposed by Buikstra and Roberts<sup>74</sup>.

Extrapulmonary tuberculosis can cause characteristic skeletal changes, such as collapse of the vertebrae (Pott's disease), periosteal reactive lesions, and osteomyelitis.

Accordingly, the most probable diagnosis for Individual A from *Casa del Fabbro* is spinal TB (Pott's disease) (Table 2), the most common type of tuberculosis involving the bony elements, and one of the most common

and devastating diseases in human history<sup>75</sup>. The partial preservation of the lumbar spine does not allow us to observe if the lesions also affected the L3 which has not been recovered. Nevertheless, the degree of extension of the lesion on the upper body of L4 prompts us to think that the lower body of L3 was also probably affected. The importance of this case is also because skeletal injuries, the only events that can be identified in an archaeological context, occur only in a minority of cases. Pre-antibiotic data suggest that about 3–5% of TB cases show bone alterations<sup>74</sup>.

As it is known that the DNA from ancient pathogens can also be preserved on ancient bone material<sup>76</sup>, we attempted to recover *Mycobacterium tuberculosis* DNA from the sequenced reads of the individual palaeopathologically diagnosed with TB. For this, we used a k-mer-based competitive metagenomic classifier<sup>77</sup> on a database containing all archaea, bacterial and viral genomes, as well as the human reference. This resulted in 14,096 reads ranked at the *Mycobacterium* genus, but only 403 of those could be uniquely assigned to the *Mycobacterium tuberculosis* species complex (Supplementary Table S7). We extracted and remapped those reads to *Mycobacterium tuberculosis* human strain H37Rv and evaluated deamination patterns. When using reads from the whole *Mycobacterium* genus, we indeed have a damage pattern consistent with aDNA (Supplementary Fig. S4), but when restricting to reads classified as *Mycobacterium tuberculosis* complex the low amount of data does not provide enough resolution for authentication. This result could be due to three different, not mutually exclusive, reasons: it is already known that the petrous bone is an excellent source of endogenous aDNA but it shows little aDNA from pathogens<sup>76</sup>; in addition, *Mycobacterium tuberculosis* is known to be difficult to molecularly diagnose even on affected and symptomatic patients from buccal swabs<sup>77–80</sup>. Moreover, within the *Mycobacterium* genus, the *Mycobacterium tuberculosis* shares up to ~99% genetic sequence identity with other common soil Mycobacteria<sup>81</sup>. However, in our case, the two individuals had never been in contact with the soil during the diagenesis process, because they were entirely covered by volcanic material. This makes the finding of the *Mycobacterium tuberculosis* DNA more likely to be endogenous.

## Discussion

To our knowledge, our results represent the first successfully sequenced Pompeian human genome.

The genome-wide analyses point out that the Pompeian individual A is genetically close to the extant Mediterranean peoples, mostly to Central Italians and Sardinians. It is plausible to think that, thanks to the expansion and the increase in effective population size during the Roman Imperial Age<sup>82</sup>, the Roman genetic pool could have contributed to the nearby populations with a genetic signature that can still be recognized in the extant Mediterranean regions today. Consistent with the autosomal results that show a high affinity of individual A with Neolithic Anatolians, its Y-chromosome haplogroup is presently only found among present-day Sardinians<sup>45</sup>. This makes it likely that this male lineage arrived in the Italian peninsula through an Anatolian source during the Neolithic.

Both Y-chromosome and mtDNA lineages from the Pompeian individual were absent among published individuals in Roman Imperial age in Italy<sup>29</sup>, suggesting a high diversity during that period across the Italian Peninsula. This signal can also be seen at the genome-wide level, by comparing the estimated ancestry proportions found in individual A with those from published Roman Imperial Age individuals<sup>29</sup> (Supplementary Table S8, Supplementary Fig. S5). Actually, for some of the published individuals, it was not possible to reach a fit with the model using the same source populations (Supplementary Table S8), suggesting a different genetic composition. This genetic variability found in the Roman Imperial Age could be also supported by contacts, interactions, and migrations of people across the Mediterranean basin already identified using different methodologies<sup>83,84</sup>.

Despite this high genetic variability in the Imperial period<sup>29</sup> the Pompeian individual A shows a higher level of shared genetic drift with the central Italy Roman Imperial Age group. This result strongly suggests that the individual that we have analysed should come from the Italian peninsula. Whether this individual belongs to the local population of Pompeii or is part of the 5% of the internal migrants characterizing the imperial population of Italy<sup>85</sup> is difficult to address, but very likely he is not part of the large external migrations related to the practice of enslavement.

Finally, we used multiple lines of evidence to determine that one of the individuals was affected by tuberculosis. It is already known that tuberculosis was endemic in the Roman Imperial period thanks to the writings and ancient descriptions from Celsus (*De Medicina*, III, 22; 1st cent. AD), Galen and Celio Aureliano, and Aretheus of Cappadocia (*Signa chron.*, I, 8; 2nd cent. AD). The increased population density that characterized the beginning of the Roman era, probably due to the development of an urban Roman way of life, favoured the spread of tuberculosis across Italy<sup>86–88</sup>.

In conclusion, our study—albeit limited to one individual—confirms and demonstrates the possibility of applying palaeogenomic methods to study human remains from this unique site. Our initial findings provide a foundation to promote an intensive analysis of well-preserved Pompeian individuals. Supported by the enormous amount of archaeological information that has been collected in the past century for the city of Pompeii, their paleogenetic analyses will help us to reconstruct the lifestyle of this fascinating population of the Imperial Roman period.

## Materials and methods

**Skeletal biology.** Sex-determination was carried out using DSP (*Diagnose Sexuelle Probabiliste*)<sup>89</sup>.

Age-at-death was determined using changes to the pubic symphysis<sup>90</sup> and the auricular surface of the ilium<sup>91</sup> for both individuals. Concerning individual A, a radiographic method based on the apposition of secondary dentine was applied to the digital radiograph of the two left lower premolars<sup>22</sup>. Periapical digital radiographs were taken using a NOMAD hand-held dental X-ray device (Aribex, USA) combined with a digital sensor (DSX,

Anthos, Italy) linked to a portable pc. All radiographs were taken with a Rinn-type digital sensor holder with 0.05 s exposure time and 60 kV.

Stature was estimated using statistical methods, taking into account the maximum length of long bones (humerus, ulna, femur and tibia), according to Giannecchini and Moggi-Cecchi<sup>22</sup>, and by applying the equations proposed by Pearson<sup>76</sup> and by Trotter and Gleser for African-American<sup>19–21</sup> that provide the most consistent estimates for Italian populations of Roman Age<sup>18</sup>.

Moreover, Pott's disease, a form of osteo-articular tuberculosis (TB), was examined by morphological approach and digital radiograph of the fourth lumbar vertebra (L4). For radiographic images, a DR Fujifilm machine was used with an exposure (100 ms) at 55 kV to 100 mA.

**Ancient DNA.** *DNA extraction.* In order to obtain the largest possible amount of aDNA<sup>15</sup> we have sampled one petrous bone for each individual. The ancient DNA extraction was performed according to the Allentoft et al.<sup>41</sup> protocol. All molecular work (DNA extraction and library preparation) was conducted in dedicated ancient DNA clean laboratory facilities at the Center of Molecular Anthropology for Ancient DNA Studies, Department of Biology, University of Rome Tor Vergata<sup>4</sup>. The otic capsule was targeted and around 100–200 mg bone powder was used for DNA extraction. In order to remove the surface contaminants, the samples were pre-digested using a digestion buffer (0.46 M of EDTA pH = 8; 10 mM of TE buffer 100×; 0.14–0.22 mg/ml of Proteinase K; 0.5% of *N*-laurylsarcosine; 1/1000 vol of Phenol red) for 45 min at 37 °C. After that step, the samples were centrifuged at 2000g for 2 min and the supernatant was discarded, new digestion buffer was added for a 24-h digestion at 37 °C. The samples were then centrifuged at 2000g for 5 min and the pellet was stored for later re-extraction. The aDNA extraction was performed on the digested solution using Silica powder-based DNA extraction protocol. To each digested sample, 100 µl silica suspension and 10× volume of binding buffer (4.88 M GuHCl and 29.3% 2-propanol; 1/1000 vol. of phenol red; 24.88 mM of NaCl; 87.6 mM of Na Acetate; final pH = 4) was added and adjusted to pH 4 with 37% HCl. The solution was incubated for 1 h at room temperature after which the samples were centrifuged for 2 min at 2000g and the supernatant was discarded. The silica was re-suspended in 1 ml of binding buffer, transferred on a new 2 ml tube and the aDNA was washed using ice-cold ethanol. Finally, the DNA was eluted in 80 µl Qiagen EB buffer.

*NGS library preparation.* The blunt-end DNA libraries were built from around 20 µl on DNA extracts using Illumina specific adapters and NEBNext DNA Sample Pre Master Mix Set 2 (E6070) kit according to manufacturer's instruction with few modifications. For each step, negative library controls were included. 25 µl reactions mix were incubated at 12 °C for 20 min and 37 °C for 15 min for the end-repair step. A purification step using Qiagen MinElute spin columns was performed and DNA was eluted in 17 µl EB buffer. In the ligation step, Illumina specific adapters were prepared according to Meyer and Kircher<sup>92</sup> and added to 15 µl of purified DNA. For ligation NEB Quick Ligation module (E6056L) was followed and the mix was incubated at 20 °C for 15 min. Then, the mixture was purified using Qiagen MinElute spin columns and DNA was eluted in 20 µl EB buffer. In the last step, adapter fill-in reaction was performed incubating at 65 °C for 20 min and 80 °C for 20 min 30 µl of reaction mix. The quantification of the library was conducted using SYBER green mix according to manufacturer's instructions and using IS8 and IS7 primers. The amount of DNA library was used to assess the optimal number of PCR cycles required for DNA library indexing. The indexing was performed on 20 µl DNA library using 2X Kapa U (following the manufacturer's temperature instruction) and 1 µl of each primer (10 mM, inPE forward primer and indexed reverse primer). The indexed amplified DNA libraries were then purified using Qiagen MinElute Kit and eluted in 50 µl EB buffer. To quantify the DNA libraries an Agilent Bioanalyzer 2100 was used, and the libraries were sequenced on Illumina HiSeq 2500 using v3 chemistry and paired end (PE) 100 cycles.

*Bioinformatics and DNA authentication.* To remove the Illumina adapter sequences, we used AdapterRemoval 1.5.2<sup>93</sup>. The adapter-free reads were then mapped against human reference genome built 37 using BWA 0.6.2<sup>94</sup> and only the aligned sequences with mapping quality at least 30 were sorted using samtools 0.1.18<sup>95</sup>. PCR duplicate reads were removed by Picard MarkDuplicate <http://broadinstitute.github.io/picard/>. All mapping statistics for the analysed individuals are reported in Table 1.

The aDNA authentication was evaluated assessing the damage pattern of the DNA, estimating the contamination in X chromosome and in mtDNA. The damage is one of the most important characteristics of the aDNA and includes deamination at the 5' end of the DNA leading to the typical C to T transition as well as G to A at the 3' end. The damage evaluation was assessed using mapDamage 2.0<sup>96</sup>.

The X chromosome contamination was evaluated by ANGSD<sup>97</sup> following the commands suggested by the authors. The mitochondrial contamination was evaluated with contamMix<sup>98</sup>. ContamMix reports the contamination rate with a Bayesian estimate of the posterior probability of the contamination proportion. The method is based on consensus mitochondrial sequences analysis by comparing the mtDNA reads of ancient individuals to the reconstructed consensus and 311 present day whole mitochondrial genomes (the possible contaminants). The analysis was restricted only to individual A due to the low mtDNA coverage shown by individual B. To reconstruct the consensus sequence, the option `-doFasta` of ANGSD package<sup>97</sup> was used and each SNP was called only if it presents at least at 3× coverage. The mitochondrial genome mapping was performed against the revised Cambridge Reference Sequence (rCRS) using base and reads mapping quality > 30.

*Sex determination.* The sex determination was evaluated by calculating the  $R_Y$  parameter, applying the Skoglund et al.<sup>26</sup> python script. The  $R_Y$  parameter represents the fraction of the total number of reads aligned with the Y chromosome (ny) divided the total number of reads mapped with both sex chromosome (ny and ny):



$R_Y = n_Y / (n_X + n_Y)$ . A  $R_Y$  parameter value above 0.077 is consistent with male individuals while a value lower than 0.016 with female ones.

**Uniparental genetic markers analysis.** For the mtDNA analysis the mitochondrial consensus sequences have been used and haplogroup was assigned using the command line version of HaploGrep<sup>27,28</sup>.

Y-chromosome haplogroup placement was carried out first using *PathPhynder*<sup>99</sup>, a software designed to place ancient and/or low-coverage individuals in a high-confidence phylogeny. While *PathPhynder* was able to fit individual A in the tree, the paucity of publicly available individuals from haplogroup A hindered a higher definition. We then proceeded to use ANGSD<sup>97</sup> to call genotypes at all phylogenetically informative sites listed in ISOGG. This curated database was split into transition and transversion variants to allow the usage of different filters, taking into consideration ancient DNA damage that manifests itself mainly through transitions<sup>25</sup>. Only variants with over 95% frequency in the population of non-clonal reads in the loci were called. When calling transition variants, 7 bases were trimmed in each read end, following mapDamage 2.0<sup>96</sup>. Although haplogroup A1b1 is the sister branch of BT—a macro-haplogroup that contains all other Y-chromosome haplogroups other than A00, A0 and A1a—less than 50 SNPs are listed in ISOGG 2020 either defining or downstream to it. Therefore, we further included all SNPs listed in previous studies<sup>39,40</sup> and from ISOGG raw data (<https://ybrowse.org/gbrowse2/gff/>) belonging to haplogroup A. Those variants were filtered to include exclusively single nucleotide point mutation biallelic variants located in the ~10 Mb short-read mappable region of the Y-chromosome<sup>100</sup>. Recurrent mutations of lower phylogenetic confidence were removed by checking whether any identical SNP (genomic position and derived base) was found downstream of BT. The haplogroup placement was further confirmed by checking whether the individual carried the ancestral alleles for downstream haplogroups BT, CT, CF, DE, F, GHIJK, IJ, P1 and NO. All SNP calls are reported in Supplementary Table S9–S13. Code and data for the Y-chromosome analyses are available at [github.com/tpinotti/Pompeya](https://github.com/tpinotti/Pompeya).

**PCA and model-based clustering.** To assess the genetic relationship between the Pompeian individual A and other ancient/modern populations, we merged the whole-genome shotgun data generated in this study with published datasets of modern and ancient populations by using Plink ([www.cog-genomics.org/plink/1.9/](http://www.cog-genomics.org/plink/1.9/))<sup>101</sup>. A subset panel of the Human Origins panel published by Lazaridis and colleagues<sup>56</sup> based on modern Western Eurasian populations (Albanians, Ashkenazi, Basques, Belarusians, British, Bulgarians, Croatians, Cypriots, Czechs, Estonians, Finns, the French, Greeks, Hungarians, Icelanders, Italians, Lithuanians, Maltese, Mordovians, Norwegians, Orcadians, Russians, Sardinians, Scottish, Spanish, Turks, Ukrainians), and selected 1030 published ancient individuals<sup>13–15,29,41–66</sup> was used (Supplementary Table S2). We performed a Principal Component Analysis (PCA) using the EIGENSTRAT method<sup>102</sup> included in the EIGENSOFT package<sup>67</sup>, projecting the ancient individuals onto the components calculated for modern Western Eurasian populations using “lsqproject” and “shrinkmode” options of smartpca.

**D- and f-statistics and population differentiation.** To investigate patterns of shared ancestry in our dataset, D-statistics analyses were performed using the Admixtools 7.0 package<sup>70</sup>. Only values of D-statistics for which  $|Z| > 3$  were considered as statistically significant. The individuals were grouped in the same population according to the cultural and period information reported to the Supplementary Table S2 (column: Label for analyses).

**qpAdm.** We performed *qpAdm* on the pseudo-haploid genotype of the Pompeian individual using *qpAdm* v7.0<sup>70</sup> applying the option “allsnps: YES”. A set of 15 outgroups was used as “right populations”: Mbuti, UstIshim, Caucasian hunter-gatherers (CHG), Eastern hunter-gatherers (EHG), Villabruna, Russia\_MA1\_HG, Natufian, Jordan\_PPNB. We tested several models, and we only plotted the model with  $p > 0.05$ . Because we are modelling a single individual, we set a minimum threshold of 100,000 SNPs<sup>45</sup>.

Using the same parameters we also performed the *qpAdm* analysis on already published Roman Imperial Age individuals<sup>29</sup> (Supplementary Table S8) and we only plotted the models with the higher p-value ( $p > 0.05$ ) where the inferred admixture proportions were also inside the interval [0, 1] (Supplementary Fig. S5).

**Pathogen screening.** We used Kraken2<sup>77</sup>, a k-mer-based competitive metagenomic classifier, to find matches on a database of all archaea, bacterial and viral genomes, as well as the human reference, with custom parameters (-confidence 0.05 -minimum-map-quality 30). 3.90% of reads were classified as Bacteria, and 14,096 of those were ranked at the *Mycobacterium* genus (Supplementary Table S7). However, only 403 reads could be uniquely assigned to the *Mycobacterium tuberculosis* species complex (Supplementary Table S7). We extracted those reads and mapped those to *Mycobacterium tuberculosis* human strain H37Rv using bowtie2<sup>103</sup> using custom parameters (-D 20 -R 3 -N 1 -L 20 -i S,1,0.50 -end-to-end<sup>104</sup>) obtaining 161 and 45 unique reads, respectively (Supplementary Table S7). While these reads could stem from actual circulating *Mycobacterium tuberculosis* found in the individual, the low amount of data does not provide enough resolution to authenticate or to perform downstream analysis on it.

## Data availability

The alignment bam file generated in this study has been deposited in the Zenodo database under the permanent DOI: <https://doi.org/10.5281/zenodo.6468368>.

Received: 17 December 2021; Accepted: 8 April 2022

Published online: 26 May 2022

## References

1. Sigurdsson, H., Carey, S., Cornell, W. & Pescatore, T. The eruption of vesuvius in AD 79. *Natl. Geogr. Res.* **1**, 332–387 (1985).
2. Luongo, G. *et al.* Impact of the AD 79 explosive eruption on Pompeii, II. Causes of death of the inhabitants inferred by stratigraphic analysis and areal distribution of the human casualties. *J. Volcanol. Geotherm. Res.* **126**, 169–200 (2003).
3. Martyn, R. *et al.* A re-evaluation of manner of death at Roman Herculaneum following the AD 79 eruption of Vesuvius. *Antiquity* **94**, 76–91 (2020).
4. Scorrano, G. *et al.* Methodological strategies to assess the degree of bone preservation for ancient DNA studies. *Ann. Hum. Biol.* **42**, 10–19 (2015).
5. Matsuo, S., Toyokumi, S., Osaka, M., Hamazaki, S. & Sugiyama, T. Degradation of DNA in dried tissues by atmospheric oxygen. *Biochem. Biophys. Res. Commun.* **208**, 1021–1027 (1995).
6. Cipollaro, M. *et al.* Ancient DNA in human bone remains from Pompeii archaeological site. *Biochem. Biophys. Res. Commun.* **247**, 901–904 (1998).
7. Di Bernardo, G., Del Gaudio, S., Galderisi, U. & Cipollaro, M. 2000 Year-old ancient equids: An ancient-DNA lesson from pompeii remains. *J. Exp. Zool. B Mol. Dev. Evol.* **302**, 550–556 (2004).
8. Di Bernardo, G. *et al.* Genetic characterization of Pompeii and Herculaneum Equidae buried by Vesuvius in 79 AD. *J. Cell. Physiol.* **199**, 200–205 (2004).
9. Guarino, F. M. *et al.* Recovery and amplification of ancient DNA from Herculaneum victims killed by the 79 AD Vesuvius hot surges. *Turk. J. Biol.* **41**, 640–648 (2017).
10. Di Bernardo, G., Del Gaudio, S., Galderisi, U., Cascino, A. & Cipollaro, M. Ancient DNA and family relationships in a Pompeian house. *Ann. Hum. Gen.* **73**, 429–437 (2009).
11. Cipollaro, M. *et al.* Histological analysis and ancient DNA amplification of human bone remains found in Caius iulius polybius house in pompeii. *Croat. Med. J.* **40**, 392–397 (1999).
12. Di Bernardo, G. *et al.* Enzymatic repair of selected cross-linked homoduplex molecules enhances nuclear gene rescue from Pompeii and Herculaneum remains. *Nucleic Acids Res.* **30**, e16 (2002).
13. Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
14. Olalde, I. *et al.* The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).
15. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
16. Hansen, H. B. *et al.* Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLoS One* **12**, e0170940 (2017).
17. Dellino, P., Dioguardi, F., Isaia, R., Sulpizio, R. & Miele, D. The impact of pyroclastic density currents duration on humans: The case of the AD 79 eruption of Vesuvius. *Sci. Rep.* **11**, 4959 (2021).
18. Pearson, K. Mathematical contribution to the theory of evolution. On the reconstruction of the stature of prehistoric races. *Philos. Trans. R. Soc. A.* **192**, 169–244 (1899).
19. Trotter, M. & Gleser, G. C. Estimation of stature from long bones of American Whites and Negroes. *Am. J. Phys. Anthropol.* **10**, 463–514 (1952).
20. Trotter, M. & Gleser, G. C. A re-evaluation of estimation of stature based on measurements of stature taken during life and of long bones after death. *Am. J. Phys. Anthropol.* **16**, 79–123 (1958).
21. Trotter, M. & Gleser, G. C. Estimation of stature from long limb bones of American Whites and Negroes—Reply. *Am. J. Phys. Anthropol.* **47**, 355–356 (1977).
22. Giannecchini, M. & Moggi-Cecchi, J. Stature in archeological samples from central Italy: Methodological issues and diachronic changes. *Am. J. Phys. Anthropol.* **135**, 284–292 (2008).
23. Henneberg, M. & Henneberg, R. J. Reconstructing medical knowledge in ancient Pompeii from the hard evidence of bones and teeth. In *Homo Faber: Studies on Nature, Technology, and Science at the Time of Pompeii* (eds. Renn, J. & Castagnetti, G.) 169–187 (L'Erma di Bretschneider, 2002).
24. Becker, M. & Capasso, L. I fuggiaschi di Ercolano: paleobiologia delle vittime dell'eruzione vesuviana del 79 dC, Roma. *J. Rom. Stud.* **93**, 404–406 (2001).
25. Sawyer, S., Krause, J., Guschanski, K., Savolainen, V. & Paabo, S. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131 (2012).
26. Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).
27. Weissensteiner, H. *et al.* HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–63 (2016).
28. van Oven, M. PhyloTree Build 17: Growing the human mitochondrial DNA tree. *Forens. Sci. Int.-Gen. S.* **5**, E392–E394 (2015).
29. Antonio, M. L. *et al.* Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708–714 (2019).
30. Hervella, M. *et al.* Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the neolithic expansion into Europe. *PLoS One* **7**, e34417 (2012).
31. Mannino, M. A. *et al.* Origin and diet of the prehistoric hunter-gatherers on the mediterranean island of Favignana (Egadi Islands, Sicily). *PLoS One* **7**, e49802 (2012).
32. Richards, M. B., Macaulay, V. A., Bandelt, H. J. & Sykes, B. C. Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* **62**, 241–260 (1998).
33. Shamoony-Pour, M., Li, M. & Merriwether, D. A. Rare human mitochondrial HV lineages spread from the Near East and Caucasus during post-LGM and Neolithic expansions. *Sci. Rep.* **9**, 14751 (2019).
34. De Fanti, S. *et al.* Fine dissection of human mitochondrial DNA haplogroup HV lineages reveals paleolithic signatures from European glacial refugia. *PLoS One* **10**, e0144391 (2015).
35. Scorrano, G. *et al.* The genetic landscape of Serbian populations through mitochondrial DNA sequencing and non-recombining region of the Y chromosome microsatellites. *Coll. Antropol.* **41**, 275–296 (2017).
36. Fraumene, C. *et al.* High resolution analysis and phylogenetic network construction using complete mtDNA sequences in Sardinian genetic isolates. *Mol. Biol. Evol.* **23**, 2101–2111 (2006).
37. Cinnioglu, C. *et al.* Excavating Y-chromosome haplotype strata in Anatolia. *Hum. Genet.* **114**, 127–148 (2004).
38. Luis, J. R. *et al.* The Levant versus the Horn of Africa: Evidence for bidirectional corridors of human migrations. *Am. J. Hum. Genet.* **74**, 532–544 (2004).
39. Francalacci, P. *et al.* Low-pass DNA sequencing of 1200 Sardinians reconstructs European Y-chromosome phylogeny. *Science* **341**, 565–569 (2013).
40. D'Atanasio, E. *et al.* The peopling of the last Green Sahara revealed by high-coverage resequencing of trans-Saharan patrilineages. *Genome Biol.* **19**, 1–15 (2018).
41. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
42. Broushaki, F. *et al.* Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016).
43. Catalano, G. *et al.* Late Upper Palaeolithic hunter-gatherers in the Central Mediterranean: New T archaeological and genetic data from the Late Epigravettian burial Oriente C (Favignana, Sicily). *Quat. Int.* **537**, 24–32 (2020).
44. de Barros Damgaard, P. *et al.* The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, eaar7711 (2018).

45. Fernandes, D. M. *et al.* The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nat. Ecol. Evol.* **4**, 334–345 (2020).
46. Fregel, R. J. *et al.* Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc. Natl. Acad. Sci. USA* **115**, 6774–6779 (2018).
47. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
48. Furtwängler, A. *et al.* Ancient genomes reveal social and genetic structure of Late Neolithic Switzerland. *Nat. Commun.* **11**, 1–11 (2020).
49. Gallego-Llorente, M. *et al.* The genetics of an early Neolithic pastoralist from the Zagros, Iran. *Sci. Rep.* **6**, 31326 (2016).
50. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
51. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
52. Hofmanova, Z. *et al.* Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl. Acad. Sci. USA* **113**, 6886–6891 (2016).
53. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).
54. Keller, A. *et al.* New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat. Commun.* **3**, 698 (2012).
55. Kilinc, G. M. *et al.* The demographic development of the first farmers in anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).
56. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
57. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
58. Lazaridis, I. *et al.* Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218 (2017).
59. Lipson, M. *et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
60. Marcus, J. H. *et al.* Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nat. Commun.* **11**, 939 (2020).
61. Mittnik, A. *et al.* The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).
62. Mittnik, A. *et al.* Kinship-based social inequality in Bronze Age Europe. *Science* **366**, 731–734 (2019).
63. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
64. Olalde, I. *et al.* The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196 (2018).
65. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505**, 87–91 (2014).
66. Posth, C. *et al.* The origin and legacy of the Etruscans through a 2000-year archeogenomic time transect. *Sci. Adv.* **7**, eabi7673 (2021).
67. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
68. Scorrano, G., Eylem Yediay, F., Pinotti, T., Feizabadifarahan, M. & Kristiansen, K. The genetic and cultural impact of the Steppe migration into Europe. *Ann. Hum. Biol.* **48**, 223–233 (2021).
69. Saube, T. *et al.* Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula. *Curr. Biol.* **31**, 2576–2591 (2021).
70. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
71. Ortner, D. J. *Identification of Pathological Conditions in Human Skeletal Remains* (Academic Press, 2003).
72. Rothschild, B., Naples, V. & Barbian, L. Bone manifestations of actinomycosis. *Ann. Diagn. Pathol.* **10**, 24–27 (2006).
73. Robbins, S. L. & Cotran, R. S. *Pathologic Basis of Disease* (Elsevier, 2002).
74. Marson, P. & Ripa Bonatti, M. Today's rheumatological paleopathology. *Reumatismo* **55**, 1–4 (2003).
75. Roberts, C. & Buikstra, J. E. *The Bioarchaeology of Tuberculosis. A Global View on a Reemerging Disease.* (University Press of Florida, 2003).
76. Margaryan, A. *et al.* Ancient pathogen DNA in human teeth and petrous bones. *Ecol. Evol.* **8**, 3534–3542 (2018).
77. Wood, D. *et al.* Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257–269 (2019).
78. Shenai, S. *et al.* Exploring alternative biomaterials for diagnosis of pulmonary tuberculosis in HIV-negative patients by use of the GeneXpert MTB/RIF assay. *J. Clin. Microb.* **51**, 4161–4166 (2013).
79. Wood, R. *et al.* Detection of *Mycobacterium tuberculosis* DNA on the oral mucosa of tuberculosis patients. *Sci. Rep.* **5**, 8668 (2015).
80. Mesman, A. *et al.* Molecular detection of *Mycobacterium tuberculosis* from buccal swabs among adults in Peru. *Sci. Rep.* **10**, 22231 (2020).
81. Borowka, P. *et al.* Screening methods for detection of ancient *Mycobacterium tuberculosis* complex fingerprints in next-generation sequencing data derived from skeletal samples. *GigaScience* **8**, giz065 (2019).
82. De Angelis, F. *et al.* Mitochondrial variability in the Mediterranean area: A complex stage for human migrations. *Ann. Hum. Biol.* **45**, 5–19 (2018).
83. O'Connell, T. C. *et al.* Living and dying at the Portus Romae. *Antiquity* **93**, 719–734 (2019).
84. Emery, M. V. *et al.* Ancient Roman mitochondrial genomes and isotopes reveal relationships and geographic origins at the local and pan-Mediterranean scales. *J. Archaeol. Sci. Rep.* **20**, 200–209 (2018).
85. Moatti, C. Translation, migration, and communication in the Roman Empire: Three aspects of movement in history. *Class. Antiq.* **25**, 109–140 (2006).
86. Canci, A. *et al.* A case of healing spinal infection from Classical Rome. *Int. J. Osteoarchaeol.* **15**, 77–83 (2005).
87. Eddy, J. J. The ancient city of Rome, its empire, and the spread of tuberculosis in Europe. *Tuberculosis* **95**, S23–28 (2015).
88. Minozzi, S., Catalano, P., Caldarini, C. & Fornaciari, G. Palaeopathology of human remains from the Roman Imperial Age. *Pathobiology* **79**, 268–283 (2012).
89. Murail, P., Bruzeki, J., Houët, F. & Cunha, E. DSP: A tool for probabilistic sex diagnosis using worldwide variability in hip bone measurement. *Bulletins et Mémoires de la Société d'Anthropologie de Paris* **17**, 167–176 (2005).
90. Todd, T. W. Age changes in the Pubic Bone VI. The interpretation of variations in the symphyseal area. *Am. J. Phys. Anthropol.* **4**, 407–424 (1921).
91. Buckberry, J. L. & Chamberlain, A. T. Age estimation from the auricular surface of the ilium: A revised method. *Am. J. Phys. Anthropol.* **119**, 231–239 (2002).
92. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold. Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).
93. Lindgreen, S. AdapterRemoval: Easy cleaning of next-generation sequencing reads. *BMC. Res. Notes.* **5**, 337 (2012).
94. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
95. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
96. Jonsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. & Orlando, L. mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).

97. Korneliusen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of next generation sequencing data. *BMC Bioinform.* **15**, 356 (2014).
98. Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
99. Martiniano, R., De Sanctis, B., Hallast, P. & Durbin, R. Placing ancient DNA in reference phylogenies. *bioRxiv* <https://doi.org/10.1101/2020.12.19.423614> (2020).
100. Poznik, G. D. *et al.* Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. *Science* **341**, 562–565 (2013).
101. Chang, C. C. *et al.* Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7–22 (2015).
102. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
103. Langmead, B. & Salzberg, S. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
104. Mühlemann, B. *et al.* Diverse variola virus (smallpox) strains were widespread in northern Europe in the Viking Age. *Science* **369**, eaaw8977 (2021).

## Acknowledgements

We thank the Parco Archeologico di Pompei for the authorization to publish this paper (prot. 2917 of 29.03.2021) in the journal "Scientific Reports" (prot. 6081 of 24.06.2021).

Digital radiograph images were made by X-Ray Sas radiology center (Aradeo, LE). We would like to thank James Fallon for his assistance with the English revision of the manuscript. Support for this project was also provided by PRIN MIUR (Italian Ministry for the Universities) 2009–11 (3 years) prot.2010EL8TXP National Scientific Coordinator and Principal Investigator OR: Biological and cultural heritage of the central-southern Italian population through 30 thousand Years EPIC and by an in-kind contribution of the Laboratory of molecular Psychiatry at the University of California, Irvine (UCI).

## Author contributions

O.R., P.F.F., F.M. supervised the study; G.S. performed laboratory work and analysed the genetic data; F.M. sequenced the libraries built; P.F.F. and S.V. provided the Pompeian individuals, all the relevant information about the archaeological context and performed the anthropological analysis; G.S. performed the genetic data interpretation; T.P. performed the Y-chromosome data analyses and interpretation; G.S., S.V. and F.M. wrote the manuscript with input from all co-authors.

## 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/s41598-022-10899-1>.

**Correspondence** and requests for materials should be addressed to G.S., P.F.F. or F.M.

**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) 2022