No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy

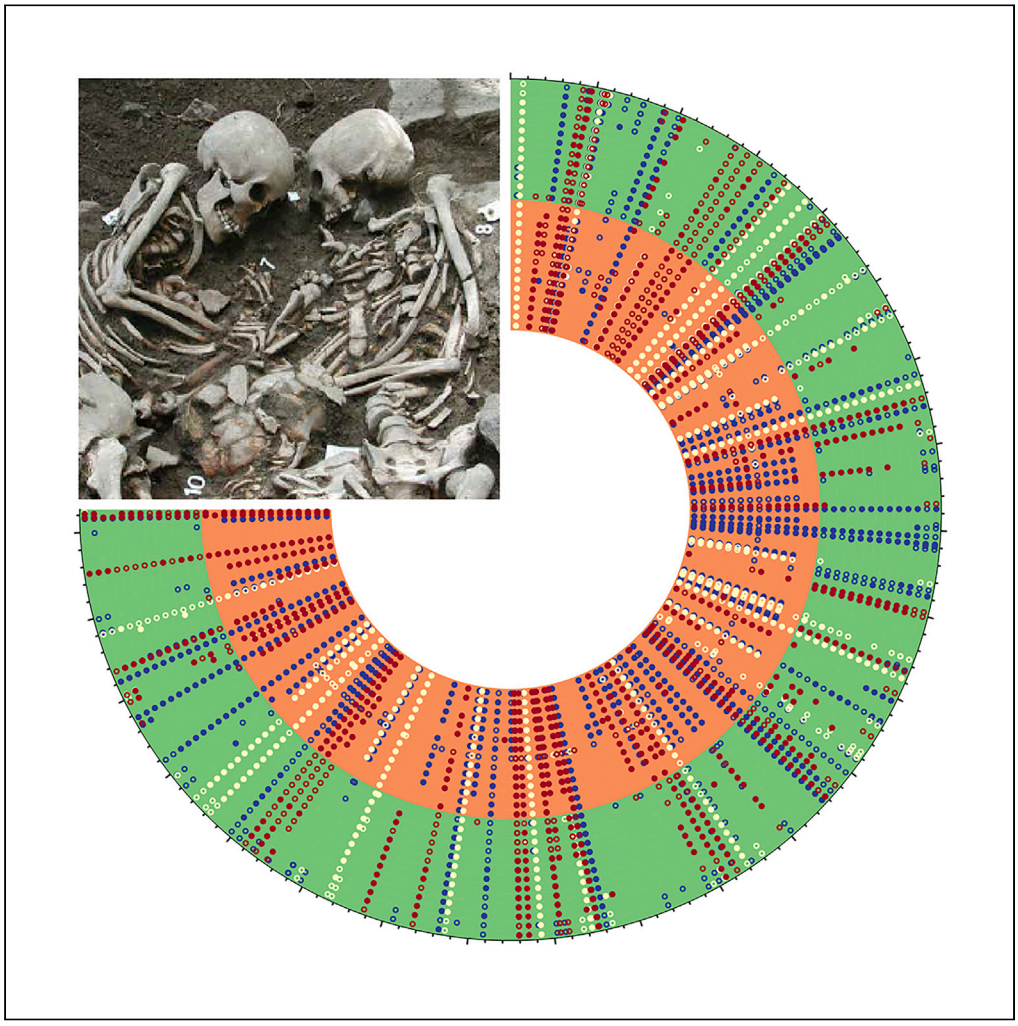
Andaine Seguin-Orlando, Caroline Costedoat, Clio Der Sarkissian, Stéfan Tzortzis, Célia Kamel, Norbert Telmon, Love Dalén, Catherine Thèves, Michel Signoli, Ludovic Orlando

iScience

CellPress
OPEN ACCESS

Article

No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy



Andaine Seguin-Orlando, Caroline Costedoat, Clio Der Sarkissian, ..., Catherine Thèves, Michel Signoli, Ludovic Orlando

ludovic.orlando@univ-tlse3.fr

Highlights

DNA screening of 26 individuals from a 1629-1630 French Alps cemetery for plague

Sequencing two *Yersinia pestis* plague genomes at 2.3- and 13.7-fold coverage

These genomes are phylogenetically closest to those from the 1636 Italian outbreak

No genetic changes explain the particularly increased spread of the disease in Italy

Seguin-Orlando et al., iScience *24*, 102383 April 23, 2021
https://doi.org/10.1016/j.isci.2021.102383

iScience

CellPress
OPEN ACCESS

Article

No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy

Andaine Seguin-Orlando,[1,2,7] Caroline Costedoat,[3,7] Clio Der Sarkissian,[1] Stéfan Tzortzis,[4] Célia Kamel,[3] Norbert Telmon,[1] Love Dalén,[5,6] Catherine Thèves,[1] Michel Signoli,[3] and Ludovic Orlando[1,8,9,*]

SUMMARY

The 17th century plague epidemic had a particularly strong demographic toll in Southern Europe, especially Italy, where it caused long-lasting economical damage. Whether this resulted from ineffective sanitation measures or more pathogenic *Yersinia pestis* strains remains unknown. DNA screening of 26 skeletons from the 1629-1630 plague cemetery of Lariey (French Alps) identified two teeth rich in plague genetic material. Further sequencing revealed two *Y. pestis* genomes phylogenetically closest to those from the 1636 outbreak of San Procolo a Naturno, Italy. They both belonged to a cluster extending from the Alps to Northern Germany that probably propagated during the Thirty Years war. Sequence variation did not support faster evolutionary rates in the Italian genomes and revealed only rare private non-synonymous mutations not affecting virulence genes. This, and the more heterogeneous spatial diffusion of the epidemic outside Italy, suggests environmental or social rather than biological causes for the severe Italian epidemic trajectory.

INTRODUCTION

With the advent of next-generation DNA sequencing, ancient DNA research has moved from single locus studies to the characterization of the complete genomes of ancient individuals, including from extinct hominids such as Neanderthals and Denisovans (see Orlando et al., 2021 for a review). The variation present in the genome of ancient individuals now provides a novel type of historical source that can help rewrite the history of population movements across the Old World and into the New World (see Nielsen et al., 2017 for a review). Time-stamped genome data have also provided unprecedented resolution to the study of the process by which plants and animals have been domesticated, selected and propagated around the world (see Frantz et al., 2020 and Kistler et al., 2020 for reviews). As the DNA fragments of ancient pathogens can survive together with those from their hosts, ancient genomic data have largely contributed to better understand past epidemics. Such data have not only contributed to solve the mysterious origins of past epidemics, such as that, that have decimated Mexican populations following their first contact with Europeans in the 16th century CE (Common Era) (Vågene et al., 2018) but also to track their evolutionary dynamics and the genetic changes that could facilitate transmission and may have affected virulence (see Spyrou et al., 2019a for a review).

Together with *Mycobacterium tuberculosis* (the agent of tuberculosis), *Yersinia pestis* (the agent of plague) represents the bacterial pathogen that thus far has received most attention in ancient DNA research. Although historians have identified the Justinian plague of the sixth century CE as marking the beginning of the first major plague pandemics, ancient DNA data have revealed that plague pathogens had in fact already started to infect human populations thousand years earlier, between the third and sixth millennium BCE (Before Common Era) (Rasmussen et al., 2015; Andrades Valtueña et al., 2017; Spyrou et al., 2018; Rascovan et al., 2019). Additionally, genome sequencing has revealed a pathogenic genetic toolkit much different then than at the time of the Justinian plague (Wagner et al., 2014; Feldman et al., 2016; Namouchi et al., 2018; Keller et al., 2019) and the infamous Black Death (Bos et al., 2011, 2016; Spyrou et al., 2016, 2019b; Guellil et al., 2020; Morozova et al., 2020; Susat et al., 2020), which marked the beginning of the so-called second pandemic by decimating 30-60% of the European population in the 14th century CE (Eckert, 1978; Benedictow, 2004). For example, the absence of the *ymt* locus, which is normally present on the

[1]Centre for Anthropobiology and Genomics of Toulouse CAGT, UMR 5288, CNRS, Université Toulouse III Paul Sabatier, Faculté de Médecine Purpan, Bâtiment A, 37 allées Jules Guesde, 31000 Toulouse, France

[2]Institute for Advanced Study in Toulouse IAST, Université Toulouse I Capitole, Esplanade de l'Université, 31080 Toulouse cedex 06, France

[3]Anthropologie bio-culturelle, droit, éthique et santé ADES, UMR 7268 CNRS EFS, Aix-Marseille Université, Faculté de Médecine, Secteur Nord Bâtiment A CS80011, Boulevard Pierre Dramard, 13344 Marseille Cedex 15, France

[4]Ministère de la Culture et de la Communication, Direction Régionale des Affaires Culturelles de PACA, Service Régional de l'Archéologie, 23 bd du Roi René, 13617 Aix-en-Provence cedex, France

[5]Centre for Palaeogenetics, Svante Arrhenius väg 20C, 10691 Stockholm, Sweden

[6]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, 10405 Stockholm, Sweden

[7]These authors contributed equally

[8]Twitter: @LudovicLorlando

[9]Lead contact

*Correspondence: ludovic.orlando@univ-tlse3.fr

https://doi.org/10.1016/j.isci.2021.102383

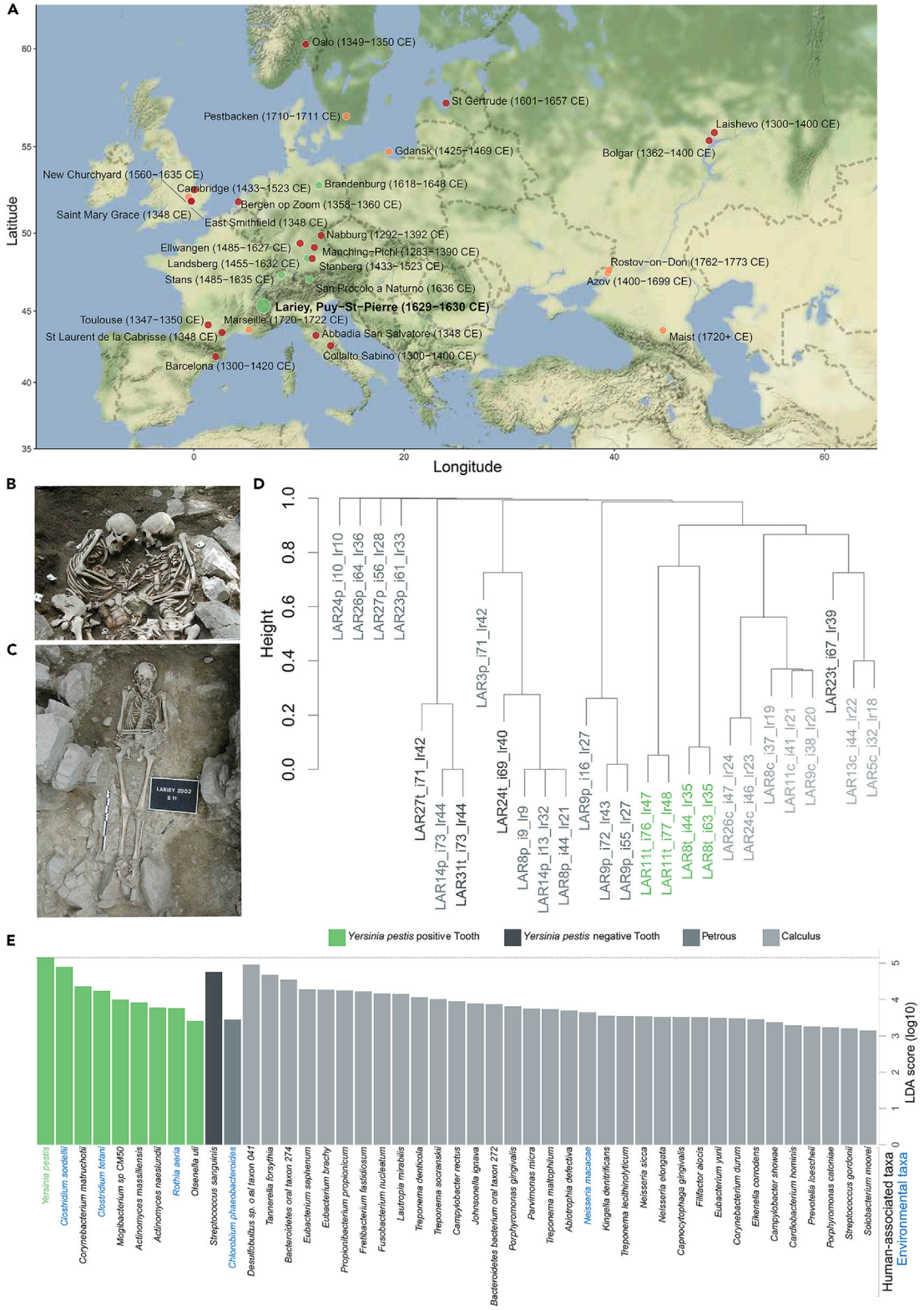


Figure 1. Sample information and metagenomic analyses

(A) Map showing the location and temporal range of the Lariey-Puy-Saint-Pierre cemetery together with previously published second pandemic plague genomes. Colors are indicated with respect to Spyrou et al. (2019b) and according to the main phylogenetic clusters shown on Figure 3.

(B) Multiple burial showing the LAR8 individual (right).

Figure 1. *Continued*

(C) Simple burial showing the LAR11 individual.

(D) Hierarchical clustering dendrogram of Bray-Curtis distances between MetaPhlAn2 (Truong et al., 2015) bacterial abundance profiles (10,000 bootstrap pseudo-replicates) and disregarding abundances below 1%. All clusters are supported with a pvclust (Suzuki and Shimodaira, 2006) approximately unbiased p value of 100.

(E) LEfSe (Segata et al., 2011) Linear Discriminant Analysis indicating those microbial species with most contrasted abundance patterns (LDA scores >3). Higher log10-LDA scores identify those bacterial species contributing the most to the differences in the metagenomic content of teeth positive or negative for *Yersinia pestis*, petrosal and dental calculus remains. The source of each species was predicted on the basis of the literature and conservatively categorized as '*environmental*' whenever both sources were likely.

See also Figures S1 and S2.

pMT1 plasmid and is key for the transmission of the disease by fleas (Sun et al., 2014), suggested different transmission modes for plague epidemics during the Bronze Age and the Iron Age (Rasmussen et al., 2015; Andrades Valtueña et al., 2017). Comparative genomic work has also provided examples of convergent evolution between strains from the first and second pandemics, including the parallel loss of magnesium transporters essential for survival into the macrophage phagosome (Keller et al., 2019; Spyrou et al., 2019b). Furthermore, strains from the 14th century CE have been found remarkably homogeneous genetically across Europe and the Caucasus, suggesting a history of extremely rapid spread from a unique source (Bos et al., 2016). Genome characterization of additional strains accompanying the epidemic waves from the following centuries, and until the 18th century CE (Biraben, 1975), have been found to form at least two main phylogenetic groups, potentially indicative of different origins from strains descending from the Black Death and either surviving in Europe (Bos et al., 2016) or in nearby foci (Guellil et al., 2020, see also Barbieri et al., 2020 for a review).

The second pandemic represents the plague pandemic that is currently best documented at the genetic level, with over 70 complete ancient plague genomes hitherto sequenced. Despite this, many areas require further research. One such area relates to the particular historical trajectory that plague epidemics from the 17th century CE have had in Italy. There, the pathogen has been reported to have been more pervasive geographically than in most other European regions, especially in 1629-1631 CE where it caused a massive demographic impact in both the main cities (e.g. 62% mortality in Verona (Donazzolo and Saibante, 1926) and around 50% in Bologna, Mantua, Pavia (Del Panta and Livi Bacci, 1977) or Parma (Lucchetti et al., 1998)) and the countryside (Manfredini et al., 2002; Alfani and Murphy, 2017). This unleashed a long-lasting economic crisis, as tax incomes became considerably reduced and cities could not be rapidly repopulated from nearby villages so as to provide the manpower yet necessary to sustain their activities (Alfani, 2013). This particular context has been proposed to have marked the beginning of the Great Divergence between Italy and the other European countries, in which economies could restart much quicker and could, for some, benefit from established and growing colonial empires. As the Italian economy and institutions were amongst the best in the continent at the time when the epidemic struck (e.g. cities were equipped with permanent health boards from the 15th century CE (Cipolla, 1976; 1981) and anti-plague tracts and other measures available at the time have been largely inspired from Italian publications (Cohn, 2009)), historians have proposed that the emergence of a new, more virulent strain may have contributed to the particularly dramatic impact measured in Italy (Alfani, 2013). In this study, we sequenced two complete plague genomes from individuals who died in the French Alps in 1629-1630 CE. These genomes were phylogenetically closest to those previously characterized in Italy in the following decade (Guellil et al., 2020). This provided us with a unique opportunity to investigate whether or not the pathogen developed a particularly harmful genetic set that could explain the dramatic epidemic striking Italy at the time.

RESULTS AND DISCUSSION

Genomic and metagenomic screening

A total of 12 ancient human petrosal bones, 10 teeth and 7 dental calculus from 15 individuals excavated at the Lariey-Puy-Saint-Pierre cemetery (France) were prepared in ancient DNA facilities for shallow shotgun sequencing on the MiniSeq Illumina instrument (transparent methods; Table S1 and Figure S1). This cemetery represents the only French site from the 17th century CE that can unambiguously be linked to the 1629-1630 CE plague pandemic (Signoli et al., 2003a) (Figures 1A–1C). Sequence mapping against the hs37d5 human reference genome revealed substantial variation in human DNA content, with petrosal bones showing maximal proportions (median = 72.43%, range = 1.57%–97.96%), followed by teeth (median = 2.67%, range = 0.02%–18.82%) and dental calculus (median = 0.08%, range = 0.05%–0.34%; Table

S1A). This is in line with the generally better postmortem DNA preservation reported for petrosal bones (Pinhasi et al., 2015) and dental calculus deriving mainly from oral bacterial biofilms (Warinner et al., 2015). The high variation in the human endogenous DNA content measured across samples of similar types supports that DNA preservation is driven by micro-environmental factors instead of global physico-chemical parameters characteristics of the site.

Bacterial taxonomic profiling with metaBIT (Louvel et al., 2016) against the MetaPhLAn2 database (Truong et al., 2015) indicated marked differences in 29 samples successfully characterized, with only one single tooth from individual LAR23 and one dental calculus tissue from individual LAR13 clustering together with modern dental plaque samples (Figure S2). Most of the other remains occupied a central position in the Principal Coordinate Analysis, indicating low diversity probably due to the presence of environmental microbes contaminating archaeological remains after death. However, abundance profiles possibly suggested non-negligible DNA proportions of *Yersinia pestis* in the teeth of two individuals LAR8 (10.8–10.9%) and LAR11 (40.7–42.1%). Polymerase chain reaction (PCR) amplifications of a 133-bp fragment located in the pPCP1 *pla* gene returned positive results on the tooth DNA extracts of these two individuals. They remained, however, negative across all the other tooth extracts tested, except one (LAR27), including 13 additional extracts that were not previously screened by sequencing (Table S1). The fact that only a small fraction of the individuals was detected positive for plague reflects the extensive postmortem DNA fragmentation and the relatively limited power of shotgun sequencing in identifying plague from DNA extracts dominated by human and/or environmental bacterial templates.

The microbial profiles from the LAR8 and LAR11 teeth were closer to each other than to any other material sequenced, including other teeth (Figure 1D). *Yersinia pestis* was furthermore confirmed through Linear Discriminant Analyses in LEfSe (Segata et al., 2011) as the top bacterial species driving the abundance profiles of those two teeth versus those that were negative for *Yersinia pestis* and the other remains. The presence of a number of oral microbes, such as *Tannerella forsythia* and *Treponema denticola*, was also characteristic of microbial profiles obtained from ancient dental calculus tissues (Figure 1E).

Read mapping against the CO92 plague reference genome with the stringent parameters described in previous studies (Spyrou et al., 2019b; Keller et al., 2019) revealed clear signatures of postmortem DNA damage, confirming the likely presence of the pathogen in both the LAR8 and LAR11 individuals (Figure S3). Such signatures appeared mainly in the form of an excess of cytosines at those genomic positions preceding mapped reads and increased (although faint) C-to-T (G-to-A) nucleotide mis-incorporation rates at read starts (ends) (Briggs et al., 2007). This is so because DNA extracts were treated with the USER enzymatic mix that breaks the DNA backbone downstream of those cytosine residues that have been deaminated after death (Rohland et al., 2015). Interestingly, mapDamage (Jónsson et al., 2013) inferred higher nucleotide mis-incorporation rates in human sequence alignments than in plague data. This was not indicative of different postmortem DNA degradation in the host and the plague genome. Instead, this reflected faster postmortem cytosine deamination at methylated CpG dinucleotides (Smith et al., 2014; Seguin-Orlando et al., 2015), which are found in the human but not the plague genome (Figure S4).

Altogether, our results suggested the likely presence of *Yersinia pestis* DNA in the tooth extracts of the LAR8 and LAR11 individuals. The virtual absence of plague DNA in the petrosal bones of individuals otherwise positive for the infection confirms previous reports on plague (Margaryan et al., 2018). The absence of plague DNA in the dental calculus of individuals otherwise positive for the infection contrasts, however, with the recent successful characterization of *Mycobacterium leprae* from ancient dental calculus remains (Fotakis et al., 2020). This likely reflects the different etiology of the diseases, with leprosy, but not plague, causing lesions in the mucous membranes of the upper respiratory tract (de Abreu et al., 2006).

Plague genome sequencing

The fraction of plague DNA sequences identified during our preliminary screening indicated that complete bacterial genomes could be obtained with reasonable sequencing efforts. Further stringent mapping of an additional 139.0 and 188.6 million reads generated on the NovaSeq instrument from the DNA content of LAR8 and LAR11 tooth libraries resulted in the characterization of two plague genomes at an average 2.3-fold and 13.7-fold coverage, respectively (Figures 2A–2D). The pCD1, pMT1 and pPCP1 plasmids were also sequenced at 4.7–219.3-fold coverage in both individuals (Table S2). Edit distance distributions confirmed the genetic proximity to *Yersinia pestis* relative to close outgroups, such as *Yersinia pseudotuberculosis*

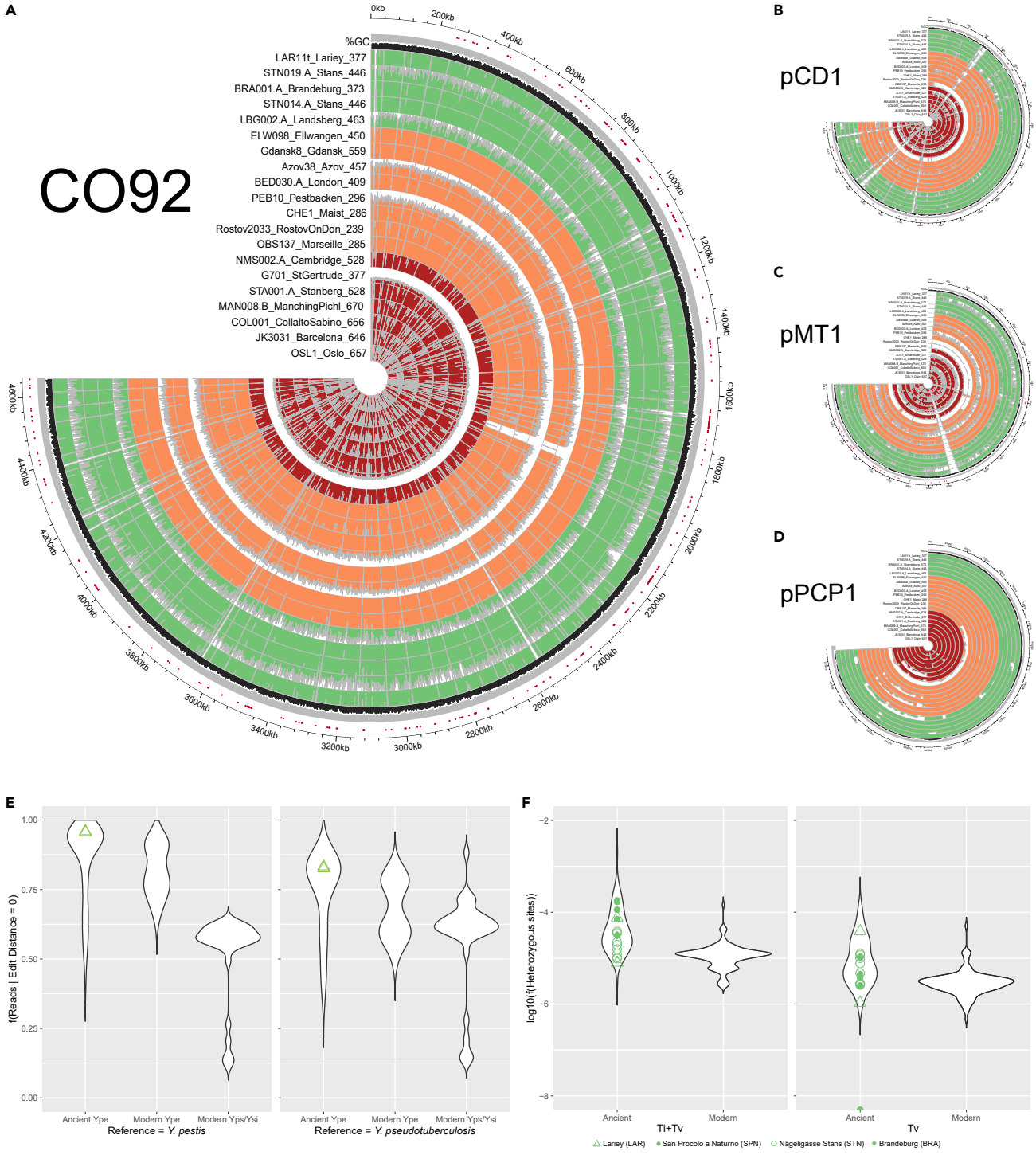


Figure 2. Plague chromosome and plasmid sequence coverage

(A) Coverage and %GC variation along 1,000 bp windows along the CO92 plague reference genome (GenBank: NC_003143.1) (Parkhill et al., 2001). Average depth-of-coverage was calculated using Paleomix (Schubert et al., 2014). %GC composition was calculated using seqtk (https://github.com/lh3/seqtk). Circular plots were traced using the circlize library in R (Gu et al., 2014) (max = 10-fold). In case multiple genomes are available for one site, the genome showing maximal coverage is shown.

(B) Coverage and %GC variation along 100 bp windows along the pCD1 plague plasmid (GenBank: NC_003131.1).

(C) Same as panel B, except that the pMT1 plasmid is shown (GenBank: NC_003134.1).

(D) Same as panel B, except that the pPCP1 plasmid is shown (GenBank: AL109969.1).

Figure 2. *Continued*

(E) Edit distance profiles obtained when mapping the sequence data underlying the LAR8 and LAR11 genomes and a comparative panel of 78 ancient and 155 modern *Yersinia pestis* (Ype), *Yersinia pseudotuberculosis* (Yps) and *Yersinia similis* (Ysi) sequence datasets aligned against the Ype and Yps reference genomes, respectively (Table S2). Edit distance distributions were considered as long as based on a minimum of 500 reads.

(F) Heterozygosity profiles of the LAR8 and LAR11 genomes and a comparative panel of 78 ancient and 129 modern *Yersinia pestis* sequence datasets aligned against the CO92 *Yersinia pestis* reference chromosome (Table S2). Transition SNPs were disregarded to account for the differential rates of postmortem DNA damage amongst the comparative panel.

and *Yersinia similis* (Figure 2E). Additionally, the fraction of heterozygous calls was limited in both individuals and comparable to that observed in modern individuals infected by a single bacterial strain and most ancient genomes previously characterized (Figure 2F). This indicated that a single infection rather than multiple co-infections from highly divergent strains likely caused the death of the two individuals.

Human genome analyses

Further mapping of the tooth sequence data provided limited coverage of the human genome for the LAR8 and LAR11 individuals (0.377-fold and 0.085-fold, respectively). This was, however, sufficient to confirm previous sex determination of LAR8 as a female individual based on the shape of the coxal bone (Signoli et al., 2003b) and to identify LAR11, an immature individual who could not be sexed anatomically (Figure S1B), as a male individual. Both the presence of typical postmortem DNA damage profiles (Figure S3B) and the calculation of negligible contamination rates ($\leq$1%) supported data authenticity. The latter were obtained on the basis of mitochondrial sequence variation for both individuals, or the X chromosome for the male individual (Tables S3 and S4). We found that the two individuals were not first- or second-degree relatives (Table S5) and projected onto the genomic variation of modern western European (Figure S5A), close to present-day French, Spanish, and Italian individuals. f3-Outgroup statistics also supported genetic affinities to present-day western Europeans (Figures S5B and S5C), and the two individuals carried mitochondrial (T1a1 and H2a1 for LAR8 and LAR11, respectively) and Y-chromosomal (R1b1a1b1 for LAR11) haplogroups that are relatively common in this region today (Tables S3 and S4).

Phylogenetic analyses

We next placed the LAR8 and LAR11 plague genomes into the phylogeny of modern and ancient plagues using Maximum Likelihood reconstruction with IQTree (Minh et al., 2020) (Figure 3A). This confirmed previous findings showing that Neolithic-Bronze Age strains (Rasmussen et al., 2015; Andrades Valtueña et al., 2017; Spyrou et al., 2019a; Rascovan et al., 2019), and those strains underlying the first (Wagner et al., 2014; Namouchi et al., 2018; Keller et al., 2019) and second pandemics (Bos et al., 2011, 2016; Morozova et al., 2020; Spyrou et al., 2016, 2019b; Susat et al., 2020) had various evolutionary origins. Additionally, no phylogenetic structure was found amongst 14th century CE second-pandemic strains. This is in line with their rapid, almost clonal spread across Europe at the time of the Black Death (Bos et al., 2016).

Post-Black-Death strains appeared, however, differentiated into two main phylogenetic groups. The LAR8 and LAR11 plague genomes were nested within a first phylogenetic cluster grouping together strains retrieved from individuals buried during the late 15th to the mid-17th century CE in Italy (San Procolo a Naturno, SPN; 1636 CE), Switzerland (Stans, 1485-1635 CE), and Germany (Landsberg, 1455-1632 CE and Brandenburg, 1618-1648 CE) (Figures 1A and 3A). The second phylogenetic cluster included strains stretched throughout the Caucasus and Europe and spanning the mid-15th to the late 18th century CE. Bayesian phylogenetic analyses in BEAST v2 (Bouckaert et al., 2019) indicated that both clusters split between 1379 and 1434 CE (median = 1407 CE) while the most common recent ancestor of second pandemic strains most likely lived between 1228 and 1321 CE (median = 1283 CE) (Figure 3B; Table S2). We noticed that each of the second-pandemic and post-second pandemic clusters showed a clear temporal structure in which older genomes generally branched first. Interestingly, these data are in line with historical evidence indicating that plague would have circulated from Germany to France and Italy following Thirty Years War troops movements. This, and the coexistence of two differentiated phylogenetic clusters within Europe, suggest a history of outbreaks deriving from two different bacterial strains, both descending from the Black Death. The geographic restriction of the first phylogenetic cluster along the Alps and Germany (Figure 1A) may indicate the persistence of local foci in the region, possibly adapted to new rodent secondary hosts, as previously suggested (Carmichael 2014; Bos et al., 2016; Susat et al., 2020). It may, however, also reflect the insufficient sampling currently available throughout Europe, the Caucasus and Russia. Therefore, further work is required to test the possibility of alternative sources.

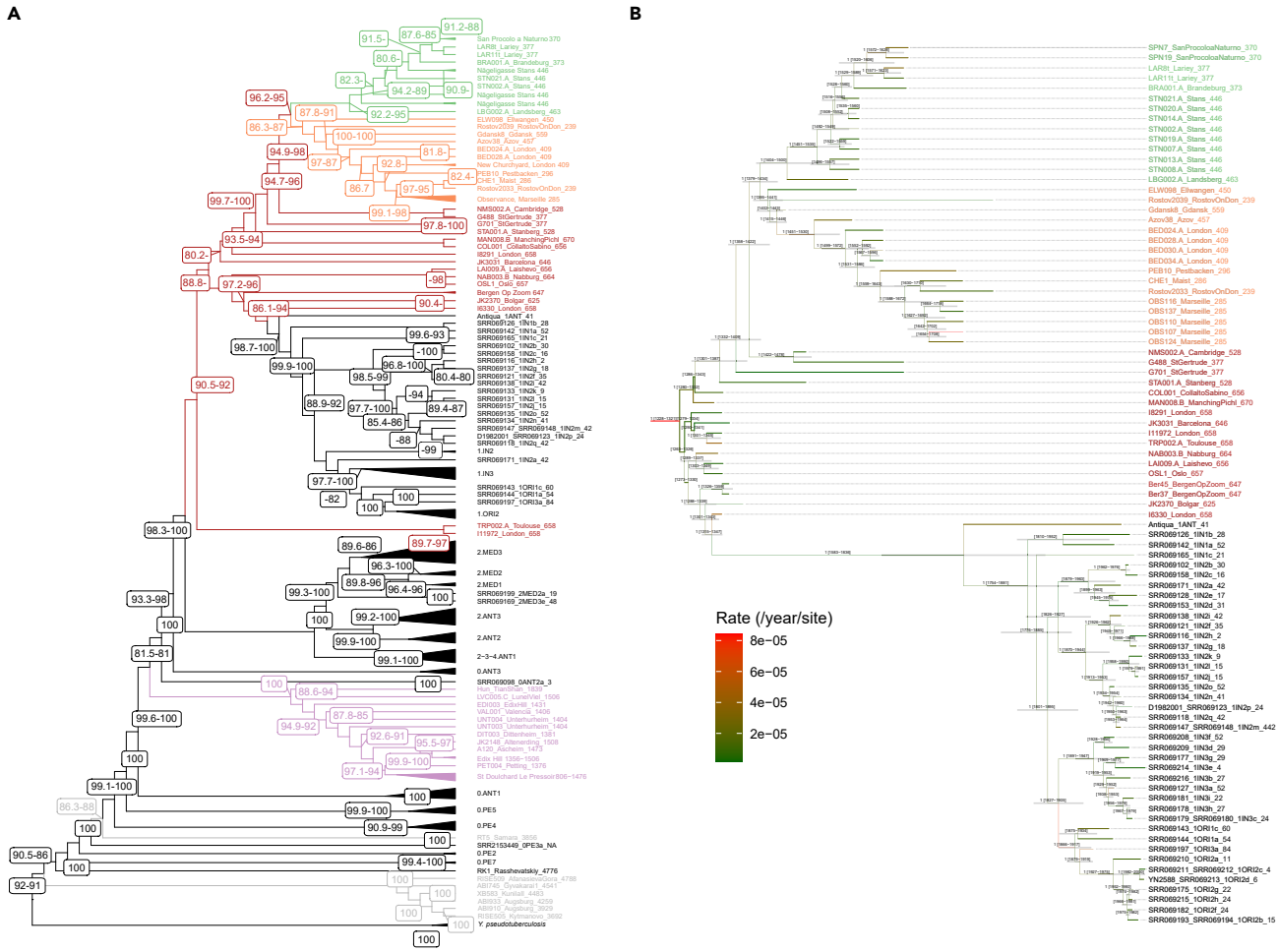


Figure 3. Phylogenetic reconstructions

(A) Maximum Likelihood topology returned by IQTree (Minh et al., 2020) using a TVM + F + R6 substitution model and a total of 21,279 polymorphic sites. Node supports are indicated using SH-aLRT support (Guindon et al., 2010) (left), as well as the ultrafast bootstrap approximation (Hoang et al., 2018) (right) when superior or equal to 80%. Only one value is shown when both supports are equal to 100%.

(B) Consensus maximal clade credibility phylogeny obtained with BEAST v2 (Bouckaert et al., 2019) restricting the sequence alignment to ancient and modern plague genomes phylogenetic clustering with or descending from second pandemic strains.

Interestingly, the two genomes from Lariey were monophyletic and appeared phylogenetically extremely close, suggesting that both individuals died during a single outbreak (Figure 3B). This contrasts to the situation reported in another site from the 17th century CE of Latvia, where two local genetically divergent strains could be documented (G488 and G701) (Susat et al., 2020). Importantly, the LAR8 and LAR11 genomes were closest to those from SPN, characterized from individuals who died at this Italian location in 1636 CE. Both Lariey and SPN genomes seem directly related to one genome from Brandenburg (BRA001), Germany, that was sequenced from the remains of one foreign Swedish soldier who occupied the city in 1631 CE during the Thirty Years War (Spyrou et al., 2019b). The direct genetic connection found between these different genomes adds to multiple historical sources highlighting the role that this war played in spreading the disease (Wilson, 2009).

Genome evolution in Italy

The sister phylogenetic relationship found between the SPN and Lariey genomes provided a unique opportunity to test the possibility of local biological adaptation for the plague circulating at the time in Italy, where the epidemic had a more profound demographic impact than in most other European countries (Alfani, 2013). There, the epidemic spread and killed in cities, hamlets and villages alike, which considerably limited the repopulation potential and the available workforce of the country. This has been proposed to

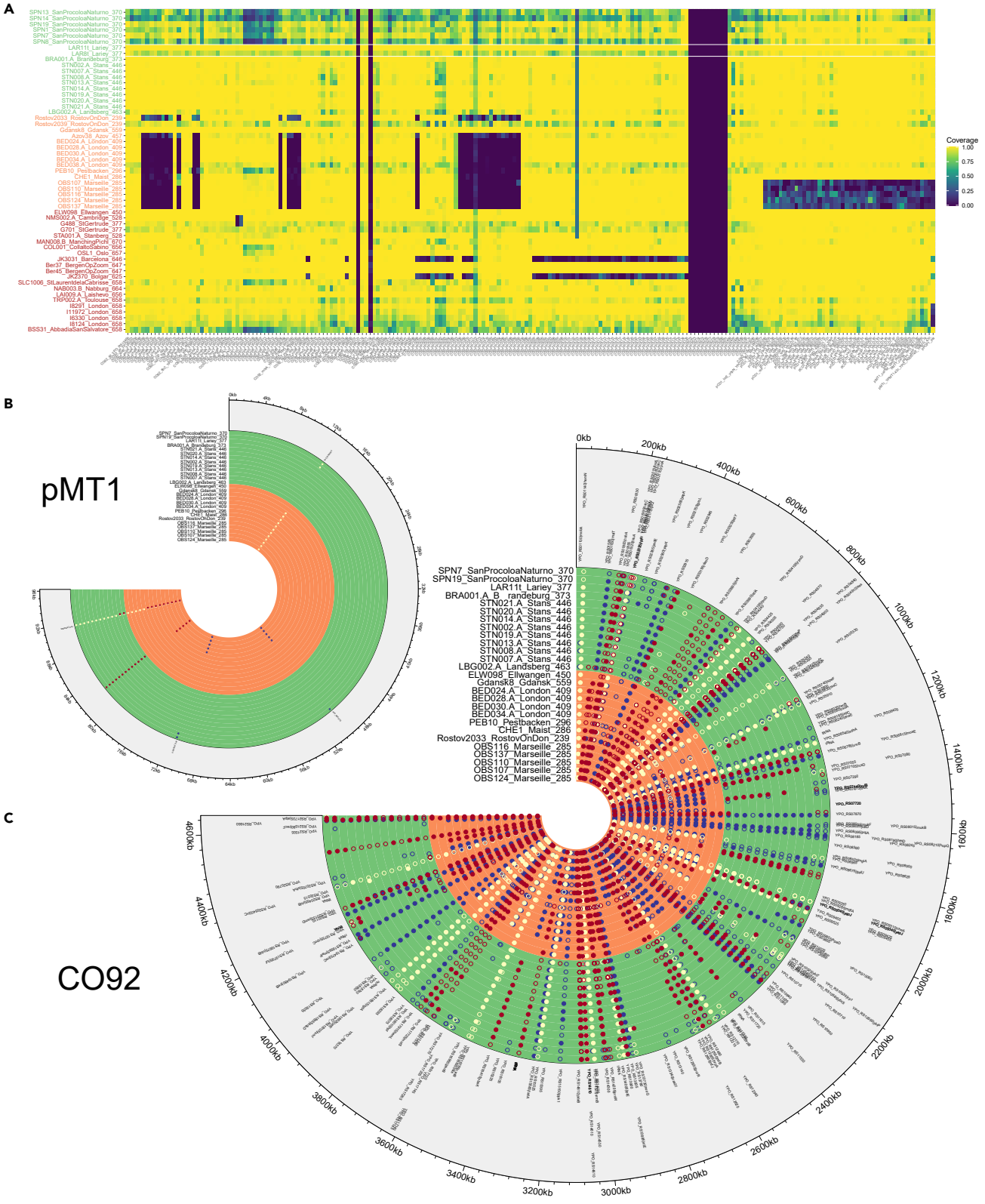

A
Coverage
1.00
0.75
0.50
0.25
0.00
B
pMT1
C
CO92

Figure 4. Genome composition of SPN, Lariey and other second pandemic plague genomes

(A) Sequence coverage at 207 gene loci involved in the pathogen virulence and transmission. Coverage represents the fraction of the gene CDS at least covered once (0 = not covered, 1 = fully covered).

(B) Non-synonymous (blue), synonymous (yellow) and intergenic (red) mutations present in a subset of 27 second pandemic ancient plague pMT1 plasmids. Open circles indicate sites not covered by at least 2 independent sequencing reads in an individual plasmid. The sequence data was trimmed and individual base quality scores rescaled in order to limit the potential impact of postmortem DNA damage.

(C) Same as Panel B, except that those mutations affecting the CO92 reference chromosome are shown.

See also Figure S6.

have significantly delayed the economic recovery of major city centers and to have played an important role in the economic divergence that followed between Italy and neighbor countries, especially those with increasing colonial power (Alfani, 2013; Parker, 2017).

In order to assess potential differences in the gene composition of the SPN and Lariey plague genomes, we looked at patterns of coverage variation at 207 virulence loci (Figure 4A) (Cui et al., 2013). This approach confirmed the previously described deletion of the *mgt* and *mgtC* genes in several second pandemic strains (Spyrou et al., 2019b; Guellil et al., 2020). These deletions were, however, not present in the phylogenetic cluster that included the SPN and Lariey genomes. Some virulence factors showed limited coverage across several SPN genomes (irp1-irp8; Figure 4A). Since this was only the case for those genomes characterized at minimal average coverage (SNP1, SNP8, SNP13, and SPN14; 1.2–2.6-fold average coverage), it was indicative of local coverage drop due to limited sequencing efforts, rather than deletions. Therefore, the SPN strain likely did not benefit from the increased intracellular survival potential within macrophages associated with *mgtB* and *mgtC* loss (Ford et al., 2014). Likewise, the SPN genomes were not different from other genomes in their phylogenetic clusters, as those did not show any particular deletion of the *inv* gene either (Figure 4A). The product of this gene is involved in epithelial colonization in *Yersinia pseudotuberculosis* but not in *Yersinia pestis* (Simonet et al., 1996), and deleted strains were previously reported in post-Black Death strains from Cambridge, UK (Spyrou et al., 2019b) and St Gertrude, Latvia (Susat et al., 2020). Furthermore, it has been hypothesized that the coexistence of *pla*-deleted and regular pPCP1 plasmids in some post-Black Death plague strains may have reduced the rapid spread of the disease through flea vectors and favored the emergence of septicemic rather than bubonic symptoms (Susat et al., 2020). We recovered those coverage drops previously reported for some second pandemic strains and confirmed their presence in both the SPN and Lariey genomes, which all showed sequencing read spanning the whole locus (Figure S6). This suggests no *pla*-dependent differential transmission rates between those strains.

We finally established the list of private mutations found amongst SPN genomes compared to other second pandemic plague genomes, including from Lariey (Figures 4B and 4C, Table S6). This was restricted to the SPN7 and SPN19 genomes, as the only one sequenced at sufficient coverage to confidently identify SNPs (i.e. approximately ~6.1-fold following trimming and base rescaling for handling possible nucleotide mis-incorporation arising from postmortem DNA damage; Table S2). Similarly, LAR8 was dismissed due to limited coverage. Only two non-synonymous mutations were common to the SPN7 and SPN19 genomes and not found in LAR11 or any other second pandemic genomes present in our comparative panel (Table S6). The first such mutation was located at the *treC* locus (CO92, NC_003143.1:4,130,262 G > A) (Figure 4C), a gene encoding the trehalose-6-phosphate hydrolase (also known as the alpha,alpha-phosphotrehalase). This enzyme is not known to affect virulence but is involved in starch and sucrose metabolism and acts as osmoprotectant in *Escherichia coli* (Rimmele and Boos, 1994). The second non-synonymous variant found specifically in SPN genomes only affected the YPO_RS00910 gene (pMT1, NC_003134.1:71,016A > G) (Figure 4B), which is involved in the type II toxin-antitoxin system RelE/ParE family toxin, a system ensuring stable plasmid inheritance for the bacteria (Guglielmini and Van Melderen, 2011). The SPN and Lariey genomes differed for a third mutation affecting one tRNA gene (CO92, NC_003143.1:3,336,035). This site was, however, found polymorphic across the other second pandemic strains and was, thus, an unlikely candidate for driving SPN-specific virulence phenotypes.

The relatively limited number of variants distinguishing the SPN genomes from those of Lariey and other second pandemic strains suggested no excessive accumulation of beneficial mutations along the SPN lineage. As hypermutability can, however, lead to the quicker emergence of beneficial mutations providing a fitness advantage to the pathogen in the co-evolutionary arms race against their host (Elena and Lenski,

2003), we further explicitly tested whether the SPN lineage displayed particularly accelerated evolutionary rates. Root-to-tip regression in TempEst (Rambaut et al., 2016) indicated temporality in the sequence data available for second pandemic strains (Supplemental information). However, BEAST analyses did not support any specific substantial shift in the mutational clock along the branch leading to the SPN cluster (Figure 3B). Our approach did not show limited sensitivity since a major acceleration in other phylogenetic branches could be detected. This, and the absence of private mutations affecting genes involved in DNA repair, rules out potentially beneficial hyper-mutator phenotypes amongst SPN strains. Hypermutability can thus be dismissed from the list of possible drivers of the increased damage to the Italian human population observed at the time.

Conclusion

Overall, the Italian plague genomes from the 17th century CE analyzed in this study showed only minute genetic differences with their closest evolutionary relatives. While the existence of other diverging strains taking over SPN cannot formally be ruled out without extensive genome sampling in Italy at the time, the relative genetic proximity amongst other second pandemic plague genomes suggests this as an unlikely alternative. The reason why more virulent strains would then remain endemic to Italy at the time of trans-European wars is also unclear. Overall, this suggests that the strains circulating in Italy during the Thirty Years War were likely not more virulent than their close phylogenetic relatives found in France, Switzerland and Germany. It follows that the underlying cause of the specific epidemic trajectory in Italy may not lie in the pathogen biology itself but rather in factors such as environmental, social and political that ultimately facilitated territorial pervasiveness and spread of the disease to villages, hamlets and cities altogether (Alfani, 2013). Previous work suggested that this extreme situation prevented the country from a quick demographic recovery, which limited the total production and fiscal income and resulted in important, long-lasting economical damage. Further work will be required to investigate the role of pathogens other than plague (e.g. typhus) and climate change in this crisis as the two other common scourges at the time.

Limitations of the study

In this study, the sequence variation present among the ancient plague strains was only investigated following read alignment against one single reference genome and not through *de novo* genome assembly. Structural variants and their possible functional consequences, thus, remain overlooked. Additionally, the genetic diversity present in Italy during the 17th century CE may not have been fully characterized from the sequence variation present in the single archaeological site of San Procolo a Naturno.

Resource availability

Lead contact

Further information and requests for resources, material and reagents should be addressed and will be fulfilled by the lead contact, Ludovic Orlando (ludovic.orlando@univ-tlse3.fr).

Materials availability

This study did not yield new unique reagents.

Data and code availability

Raw sequence data and alignments are publicly available at the European Nucleotide Archive (ENA) under accession number ENA: PRJEB43291. Previously published genomic data used in this study are available at the sources referenced in the transparent methods. The software and computational procedures are detailed in the transparent methods, including the individual versions used as well as the command parameters.

METHODS

All methods can be found in the accompanying Transparent methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.102383.

ACKNOWLEDGMENTS

We thank Loreleï Chauvey, Stéphanie Schiavinato, and Laure Tonasso-Calvière for managing the ancient DNA lab facilities in Toulouse and running the MiniSeq sequencing instrument. We thank Claudia Gillet, Pierre Clavel and all other members of the AGES research group at CAGT for discussions. The authors acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure (NGI) in Sweden, the Knut and Alice Wallenberg Foundation, and UPPMAX for providing assistance in massively parallel DNA sequencing and computational infrastructure. This project has received funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010; the ANR LifeChange; the CNRS MITI '*Défi Ecologie de la Santé 2020*' program; the Simone and Cino Del Duca Foundation (Subventions scientifiques 2020, HealthTimeTravel), and; the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 681605).

AUTHOR CONTRIBUTIONS

A.S.O., C.C., N.T., M.S. and L.O. conceived the project. L.O. designed research. C.C. and M.S. provided samples. C.C., S.T., C.K. and M.S. provided information about historical and archaeological context. A.S.O., C.C., M.S. and L.O. performed sampling. A.S.O., C.D.S. and C.T. carried out ancient DNA laboratory work. L.O. carried out computational analyses, with input from A.S.O. and C.D.S. L.D. and L.O. provided reagents and material. A.S.O. and L.O. wrote the supplemental Information, with input from C.D.S. L.O. wrote the paper, with input from A.S.O. and C.D.S., and all coauthors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 21, 2020
Revised: February 25, 2021
Accepted: March 29, 2021
Published: April 23, 2021

REFERENCES

Cui, Y., Yu, C., Yan, Y., Li, D., Li, Y., Jombart, T., Weinert, L.A., Wang, Z., Guo, Z., Xu, L., et al. (2013). Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. Proc. Natl. Acad. Sci. U S A *110*, 577–582.

de Abreu, M.A.M.M., Michalany, N.S., Weckx, L.L.M., Neto Pimentel, D.R., Hirata, C.H.W., and de Avelar Alchorne, M.M. (2006). The oral mucosa in leprosy: a clinical and histopathological study. Braz. J. Otorhinolaryngol. *72*, 312–316.

Alfani, G. (2013). Plague in seventeenth century Europe and the decline of Italy: an epidemiological hypothesis. Eur. Rev. Eco. Hist. *17*, 408–430.

Alfani, G., and Murphy, T.E. (2017). Plague and lethal epidemics in the pre-industrial world. J. Eco. Hist. *77*, 314–343.

Andrades Valtueña, A., Mittnik, A., Key, F.M., Haak, W., Allmäe, R., Belinskij, A., Daubaras, M., Feldman, M., Jankauskas, R., Janković, I., et al. (2017). The stone age plague and its persistence in Eurasia. Curr. Biol. *27*, 3683–3691.e8.

Barbieri, R., Signoli, M., Chevé, D., Costedoat, C., Tzortzis, S., Aboudharam, G., Raoult, D., and Drancourt, M. (2020). Yersinia pestis: the natural history of Plague. Clin. Microbiol. Rev. *34*, e00044–19.

Benedictow, O.J. (2004). The Black Death 1346-1353. The Complete History (Boydell Press).

Biraben, J.N. (1975). La peste dans l'histoire. Les hommes et la peste en France et dans les pays européens et méditerranéens, *volume I* (de Gruyter).

Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., DeWitte, S.N., Meyer, M., Schmedes, S., et al. (2011). A draft genome of *Yersinia pestis* from victims of the Black Death. Nature *478*, 506–510.

Bos, K.I., Herbig, A., Sahl, J., Waglechner, N., Fourment, M., Forrest, S.A., Klunk, J., Schuenemann, V.J., Poinar, D., Kuch, M., et al. (2016). Eighteenth century *Yersinia pestis* genomes reveal the long-term persistence of an historical plague focus. Elife *5*, e12994.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., et al. (2019). Beast 2.5: an advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. *15*, e1006650.

Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. U S A *104*, 14616–14621.

Carmichael, A.G. (2014). Plague persistence in western Europe: a hypothesis. In Pandemic Disease in the Medieval World: Rethinking the Black Death, M. Green, ed. (Arc Humanities Press), pp. 157–192.

Cipolla, C.M. (1976). Publich Health and the Medical Profession in the Renaissance (Cambridge University Press).

Cipolla, C.M. (1981). Fighting the Plague in Seventeenth-Century Italy (University of Wisconsin Press).

Cohn, S.K. (2009). Cultures of plague. Medical Thought at the End of the Renaissance (Oxford University Press).

Donazzolo, P., and Saibante, M. (1926). Lo sviluppo demografico di Verona e délia sua provincia dalla fine del sec. XV ai nostri giorni. Metron *VI*, 3–4.

Eckert, E.A. (1978). Boundary formation and diffusion of plague: swiss epidemics from 1562-1669. Ann. de Démographie Historique *1*, 49–80.

Elena, S.F., and Lenski, R.E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat. Rev. Genet. *4*, 457–469.

Feldman, M., Harbeck, M., Keller, M., Spyrou, M.A., Rott, A., Trautmann, B., Scholz, H.C., Päffgen, B., Peters, J., McCormick, M., et al. (2016). A high-coverage *Yersinia pestis* genome

from a sixth-century justinianic plague victim. Mol. Biol. Evol. *33*, 2911–2923.

Ford, D.C., Joshua, G.W.P., Wren, B.W., and Oyston, P.C.F. (2014). The importance of the magnesium transporter MgtB for virulence of *Yersinia pseudotuberculosis* and *Yersinia pestis*. Microbiology (Reading) *160*, 2710–2717.

Fotakis, A.K., Denham, S.D., Mackie, M., Orbegozo, M.I., Mylopotamitaki, D., Gopalakrishnan, S., Sicheritz-Pontén, T., Olsen, J.V., Cappellini, E., Zhang, G., et al. (2020). Multi-omic detection of *Mycobacterium leprae* in archaeological human dental calculus. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190584.

Frantz, L.A.F., Bradley, D.G., Larson, G., and Orlando, L. (2020). Animal domestication in the era of ancient genomics. Nat. Rev. Genet. *21*, 449–460.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812.

Guellil, M., Kersten, O., Namouchi, A., Luciani, S., Marota, I., Arcini, C.A., Iregren, E., Lindemann, R.A., Warfvinge, G., Bakanidze, L., et al. (2020). A genomic and historical synthesis of plague in 18th century Eurasia. Proc. Natl. Acad. Sci. U S A *117*, 28328–28335.

Guglielmini, J., and Van Melderen, L. (2011). Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. Mob. Genet. Elem. *1*, 283–290.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. *35*, 518–522.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics *29*, 1682–1684.

Keller, M., Spyrou, M.A., Scheib, C.L., Neumann, G.U., Kröpelin, A., Haas-Gebhard, B., Päffgen, B., Haberstroh, J., Ribera I Lacomba, A., Raynaud, C., et al. (2019). Ancient *Yersinia pestis* genomes from across western Europe reveal early diversification during the first pandemic (541-750). Proc. Natl. Acad. Sci. U S A *116*, 12363–12372.

Kistler, L., Bieker, V.C., Martin, M.D., Pedersen, M.W., Ramos Madrigal, J., and Wales, N. (2020). Ancient plant genomics in archaeology, herbaria, and the environment. Annu. Rev. Plant Biol. *71*, 605–629.

Louvel, G., Der Sarkissian, C., Hanghøj, K., and Orlando, L. (2016). metaBIT, an integrative and automated metagenomic pipeline for analysing microbial profiles from high-throughput sequencing shotgun data. Mol. Ecol. Resour. *16*, 1415–1427.

Lucchetti, E., Manfredini, M., and De Iasio, S. (1998). La peste de 1630 dans la ville et dans le territoire de Parme (Italie). BMSAP *10*, 411–424.

Manfredini, M., De Iasio, S., and Lucchetti, E. (2002). The plague of 1630 in the territory of Parma : outbreak and effects of a crisis. Int. J. Anthropol. *17*, 41–57.

Margaryan, A., Hansen, H.B., Rasmussen, S., Sikora, M., Moiseyev, V., Khoklov, A., Epimakhov, A., Yepiskoposyan, L., Kriiska, A., Varul, L., et al. (2018). Ancient pathogen DNA in human teeth and petrous bones. Ecol. Evol. *8*, 3534–3542.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534.

Morozova, I., Kasianov, A., Bruskin, S., Neukamm, J., Molak, M., Batieva, E., Pudło, A., Rühli, F.J., and Schuenemann, V.J. (2020). New ancient Eastern European *Yersinia pestis* genomes illuminate the dispersal of plague in Europe. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190569.

Namouchi, A., Guellil, M., Kersten, O., Hänsch, S., Ottoni, C., Schmid, B.V., Pacciani, E., Quaglia, L., Vermunt, M., Bauer, E.L., et al. (2018). Integrative approach using *Yersinia pestis* genomes to revisit the historical landscape of plague during the Medieval Period. Proc. Natl. Acad. Sci. U S A *115*, E11790–E11797.

Nielsen, R., Akey, J.M., Jakobsson, M., Pritchard, J.K., Tishkoff, S., and Willerslev, E. (2017). Tracing the peopling of the world through genomics. Nature *541*, 302–310.

Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Avila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., and Warinner, C. (2021). Ancient DNA analysis. Nat. Rev. Methods Primers *1*, 15, https://doi.org/10.1038/s43586-021-00016-3.

Del Panta, L., and Livi Bacci, M. (1977). Chronologie, intensité et diffusion des crises de mortalité en Italie: 1600-1850. Population *32*, 401–446.

Parker, G. (2017). Global Crisis: War, Climate Change and Catastrophe in the Seventeenth Century (Yale University Press).

Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T., Prentice, M.B., Sebaihia, M., James, K.D., Churcher, C., Mungall, K.L., et al. (2001). Genome sequence of *Yersinia pestis*, the causative agent of plague. Nature *413*, 523–527.

Pinhasi, R., Fernandes, D., Sirak, K., Novak, M., Connell, S., Alpaslan-Roodenberg, S., Gerritsen, F., Moiseyev, V., Gromov, A., Raczky, P., et al. (2015). Optimal Ancient DNA yields from the inner ear part of the human petrous bone. PLoS One *10*, e0129102.

Rambaut, A., Lam, T.T., de Carvalho, L.M., and Pybus, O.G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst. Virus Evol. *2*, vew007.

Rascovan, N., Sjögren, K.-G., Kristiansen, K., Nielsen, R., Willerslev, E., Desnues, C., and Rasmussen, S. (2019). Emergence and spread of basal lineages of *Yersinia pestis* during the neolithic decline. Cell *176*, 295–305.e10.

Rasmussen, S., Allentoft, M.E., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K.G., Pederseen, A.G., Schubert, M., Van Dam, A., Outzen Kapel, C.M., et al. (2015). Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. Cell *163*, 571–582.

Rimmele, M., and Boos, W. (1994). Trehalose-6-phosphate hydrolase of *Escherichia coli*. J. Bacteriol. *176*, 5654–5664.

Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., and Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *370*, 20130624.

Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernandez, R., Kircher, M., McCue, M., et al. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nat. Protoc. *9*, 1056–1082.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. Genome Biol. *12*, R60.

Seguin-Orlando, A., Gamba, C., Der Sarkissian, C., Ermini, L., Louvel, G., Boulygina, E., Sokolov, A., Nedoluzhko, A., Lorenzen, E.D., Lopez, P., et al. (2015). Pros and cons of methylation-based enrichment methods for ancient DNA. Sci Rep *5*, 11826.

Signoli, M., Tzortzis, S., Bizot, B., Ardagna, Y., Rigeade, C., Acotto, J., Chève, D., and Seguy, I. (2003a). Puy-Saint-Pierre, Lariey: Cimetière en lien avec l'épidémie de peste de 1629-1630 (D.R.A.C. - P.A.C.A. - S.R.A.), pp. 41–43.

Signoli, M., Tzortzis, S., Bizot, B., Ardagna, Y., Rigeade, C., Acotto, J., Chève, D., and Seguy, I. (2003b). Le cimetière de Lariey (Puy-Saint-Pierre, Hautes-Alpes): Un ensemble funéraire de l'épidémie de peste de 1629-1630. Population, sociétés et patrimoines dans les milieux alpins (S.C.E.R.E.N., C.C.D.P. des Hautes-Alpes, C.R.D.P. de l'Académie d'Aix-Marseille), pp. 29–33.

Simonet, M., Riot, B., Fortineau, N., and Berche, P. (1996). Invasin production by *Yersinia pestis* is abolished by insertion of an IS200-like element within the inv gene. Infect. Immun. *64*, 375–379.

Smith, O., Clapham, A.J., Rose, P., Liu, Y., Wang, J., and Allaby, R.G. (2014). Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process. Sci Rep *4*, 5559.

Spyrou, M.A., Tukhbatova, R.I., Feldman, M., Drath, J., Kacki, S., Beltrán de Heredia, J., Arnold, S., Sitdikov, A.G., Castex, D., Wahl, J., et al. (2016). Historical *Y. pestis* genomes reveal the European Black death as the source of ancient and modern plague pandemics. Cell Host Microbe *19*, 874–881.

Spyrou, M.A., Tukhbatova, R.I., Wang, C.-C., Valtueña, A.A., Lankapalli, A.K., Kondrashin, V.V., Tsybin, V.A., Khokhlov, A., Kühnert, D., Herbig, A., et al. (2018). Analysis of 3800-year-old *Yersinia*

pestis genomes suggests Bronze Age origin for bubonic plague. Nat. Commun. *9*, 2234.

Spyrou, M.A., Bos, K.I., Herbig, A., and Krause, J. (2019a). Ancient pathogen genomics as an emerging tool for infectious disease research. Nat. Rev. Genet. *20*, 323–340.

Spyrou, M., Keller, M., Tukhbatova, R.I., Scheib, C.L., Nelson, E.A., Andrades Valtueña, A., Neumann, G.U., Walker, D., Alterauge, A., Carty, N., et al. (2019b). Phylogeography of the second plague pandemic revealed through analysis of historical *Yersinia pestis* genomes. Nat. Commun. *10*, 4470.

Sun, Y.-C., Jarrett, C.O., Bosio, C.F., and Hinnebusch, B.J. (2014). Retracing the evolutionary path that led to flea-borne Transmission of *Yersinia pestis*. Cell Host Microbe *15*, 578–586.

Susat, J., Bonczarowska, J.H., Pētersone-Gordina, E., Immel, A., Nebel, A., Gerhards, G., and Krause-Kyora, B. (2020). *Yersinia pestis* strains from Latvia show depletion of the pla virulence gene at the end of the second plague pandemic. Sci. Rep. *10*, 14628.

Suzuki, R., and Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics *22*, 1540–1542.

Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., and Segata, N. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods *12*, 902–903.

Vågene, Å.J., Herbig, A., Campana, M.G., Robles García, N.M., Warinner, C., Sabin, S., Spyrou, M.A., Andrades Valtueña, A., Huson, D., Tuross, N., et al. (2018). *Salmonella enterica* genomes from victims of a major sixteenth-century epidemic in Mexico. Nat. Ecol. Evol. *2*, 520–528.

Wagner, D.M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J.W., Enk, J., Birdsell, D.N., Kuch, M., Lumibao, C., et al. (2014). *Yersinia pestis* and the plague of Justinian 541-543 AD: a genomic analysis. Lancet Infect. Dis. *14*, 319–326.

Warinner, C., Speller, C., and Collins, M.J. (2015). A new era in palaeomicrobiology: prospects for ancient dental calculus as a long-term record of the human oral microbiome. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20130376.

Wilson, P.H. (2009). The Thirty Years War: Europe's Tragedy (Belknap Press).

iScience, Volume *24*

Supplemental information

No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy

Andaine Seguin-Orlando, Caroline Costedoat, Clio Der Sarkissian, Stéfan Tzortzis, Célia Kamel, Norbert Telmon, Love Dalén, Catherine Thèves, Michel Signoli, and Ludovic Orlando

Supplemental Figures

Figure S1. Preservation states of the cranial and post-cranial skeletons of the LAR8 and LAR11 individuals. (A) Sample LAR8. (B) Sample LAR11. Related to Figure 1.

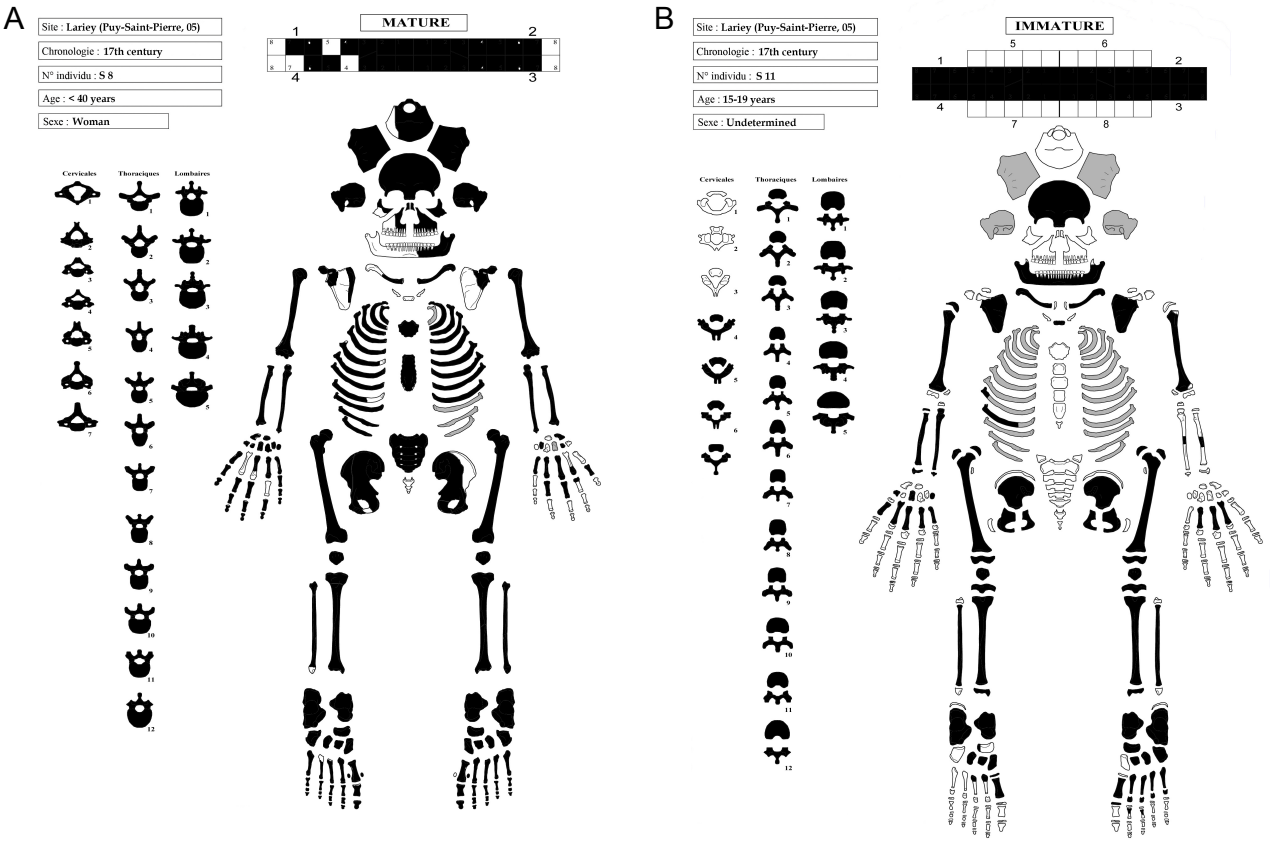


Figure S2. Genus-level Principal Coordinates Analysis (PCoA) of MetaPhlAn2 (Truong et al., 2015) bacterial abundance profiles. The 29 profiles obtained in this study are compared to a panel of 15 soil and 689 human-associated modern microbiota obtained from (Fierer et al., 2012; The Human Microbiome Project Consortium, 2012). Those remains analyzed here with shotgun sequencing DNA data are indicated with respect to the tissue originally sampled (c: dental calculus; p: petrosal bone, and; t: tooth) and the external (i) and internal indices (lr) used for constructing triple-indexed double-stranded DNA libraries. Related to Figure 1.

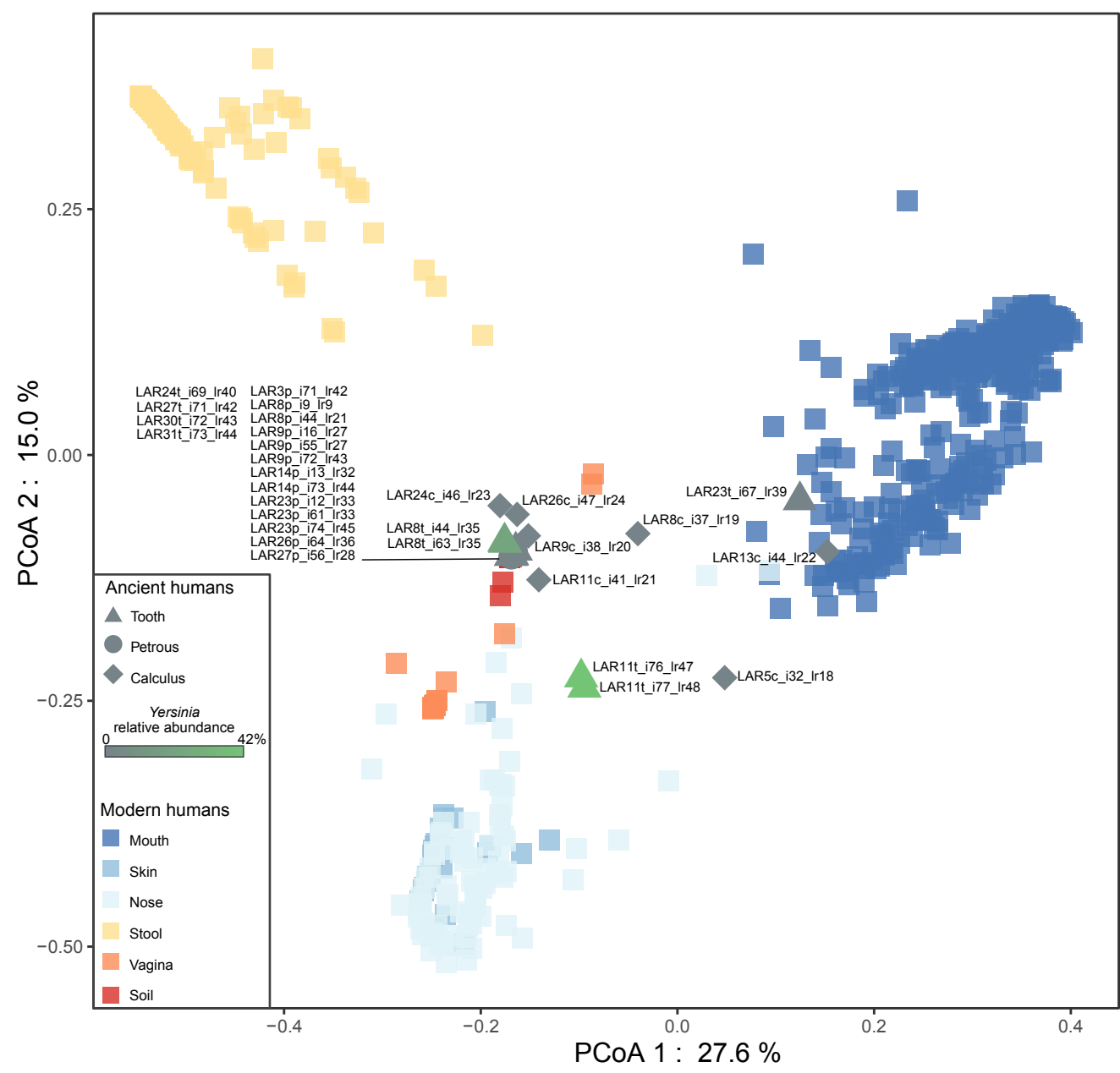

PCoA 2 : 15.0 %
PCoA 1 : 27.6 %
0.25
0.00
−0.25
−0.50
−0.4
−0.2
0.0
0.2
0.4
LAR24t_i69_lr40
LAR27t_i71_lr42
LAR30t_i72_lr43
LAR31t_i73_lr44
LAR3p_i71_lr42
LAR8p_i9_lr9
LAR8p_i44_lr21
LAR9p_i16_lr27
LAR9p_i55_lr27
LAR9p_i72_lr43
LAR14p_i13_lr32
LAR14p_i73_lr44
LAR23p_i12_lr33
LAR23p_i61_lr33
LAR23p_i74_lr45
LAR26p_i64_lr36
LAR27p_i56_lr28
LAR24c_i46_lr23
LAR26c_i47_lr24
LAR8t_i44_lr35
LAR8t_i63_lr35
LAR9c_i38_lr20
LAR11c_i41_lr21
LAR23t_i67_lr39
LAR8c_i37_lr19
LAR13c_i44_lr22
LAR11t_i76_lr47
LAR11t_i77_lr48
LAR5c_i32_lr18
Ancient humans
Tooth
Petrous
Calculus
Yersinia
relative abundance
0
42%
Modern humans
Mouth
Skin
Nose
Stool
Vagina
Soil

Figure S3. DNA fragmentation profiles. (A) DNA fragmentation profiles underlying the two ancient plague genomes characterized in this study (top: DNA library LAR8t_i63_lr35; bottom: DNA library LAR11t_i76_lr47). All aligned reads were processed using mapDamage2 (Jónsson et al., 2013) considering base with quality scores superior or equal to 30. Position-wise base compositional profiles are provided with the 10 first (left, 1 to 10) and 10 last (right, N-9 to N) read positions within sequencing reads, and the 5 genomic positions preceding read starts (left, -1 to - 5) or following read ends (right, N+1 to N+5). (B) DNA fragmentation profiles underlying the two ancient human genomes characterized in this study (top: DNA library LAR8t_i63_lr35; bottom: DNA library LAR11t_i76_lr47). Related to *Figure 2.*

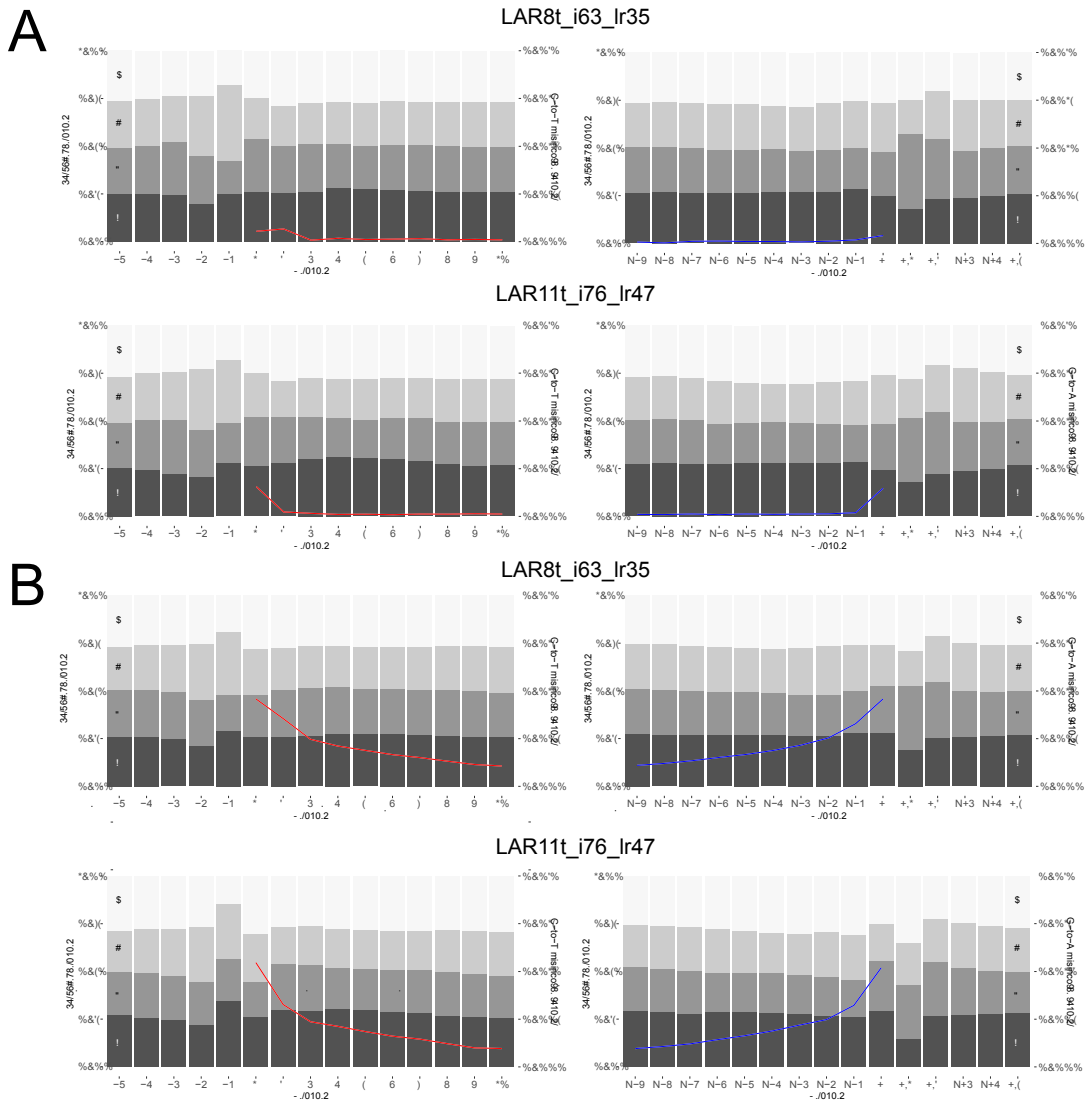


Figure S4. Nucleotide mis-incorporation profiles within CpG and CpH contexts in human and plague aligned sequencing reads. Nucleotide mis-incorporation rates are calculated within CpG and outside of CpG contexts (CpH refers to CpA, CpC and CpT contexts) in order to account for the difference in DNA methylation patterns in the human and the bacterial genome. Ancient human DNA data show evidence for higher Cytosine deamination rates than ancient bacterial DNA data in CpG contexts, in line with the presence of methylated CpGs in the human genome, that are known to undergo fast post-mortem decay (Smith et al., 2014, Seguin-Orlando et al., 2015). No differences in Cytosine mis-incorporation rates are found outside of CpG contexts across the human and bacterial data, confirming similar post-mortem deamination rates. The difference in overall Cytosine deamination rates observed by mapDamage (Figure S3) is thus fully driven by the different DNA methylation properties of plague and human genomes. (A) specimen LAR8, plague data. (B) specimen LAR11, plague data. (C) specimen LAR8, human data. (D) specimen LAR11, human data. C⟶T: Cytosine mis-incorporation rates in CpH contexts. All others: Average nucleotide mis-incorporation rates affecting adenine, guanine and thymine residues, but not cytosines. Related to Figure 2.

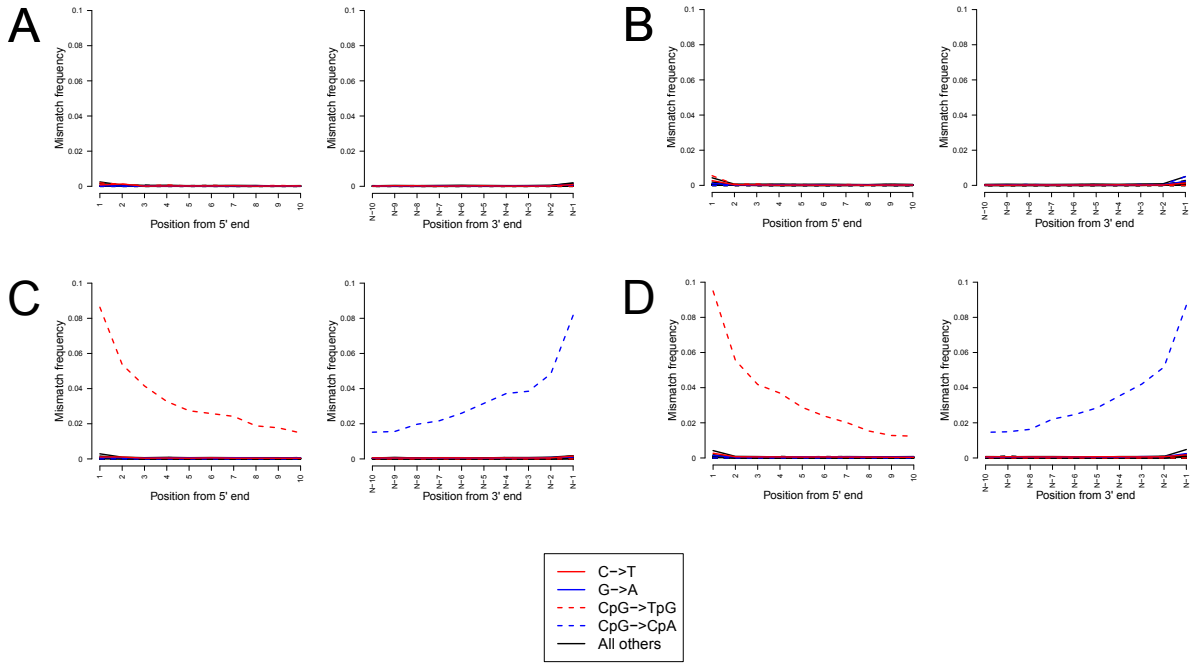
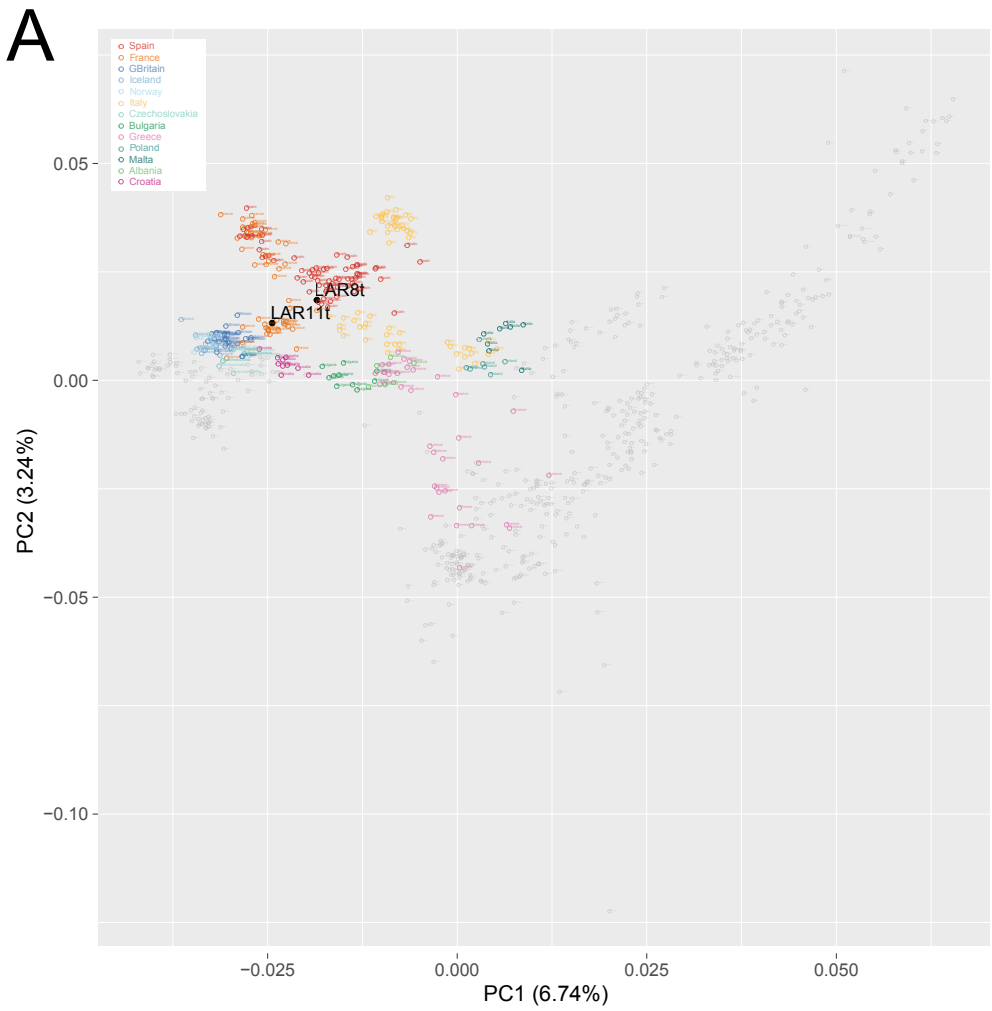


Figure S5. Human Population Genetic Affinities. (A) Principal Component Analysis (PCA) showing the two human genomes from Lariey projected onto the genetic variation present across 792 modern humans from Western Eurasia (Lazaridis et al., 2014). (B) Top-25 (Mbuti; LAR8t, X) f3-Outgroups, where X is a modern human individual from the 1240K panel. (C) Same as Panel B, for the top-25 (Mbuti; LAR11t, X) f3-Outgroups. Related to Figure 1.


A
Spain
France
GBritain
Iceland
Norway
Italy
Czechoslovakia
Bulgaria
Greece
Poland
Malta
Albania
Croatia
LAR8t
LAR11t
0.05
0.00
−0.05
−0.10
PC2 (3.24%)
−0.025
0.000
0.025
0.050
PC1 (6.74%)

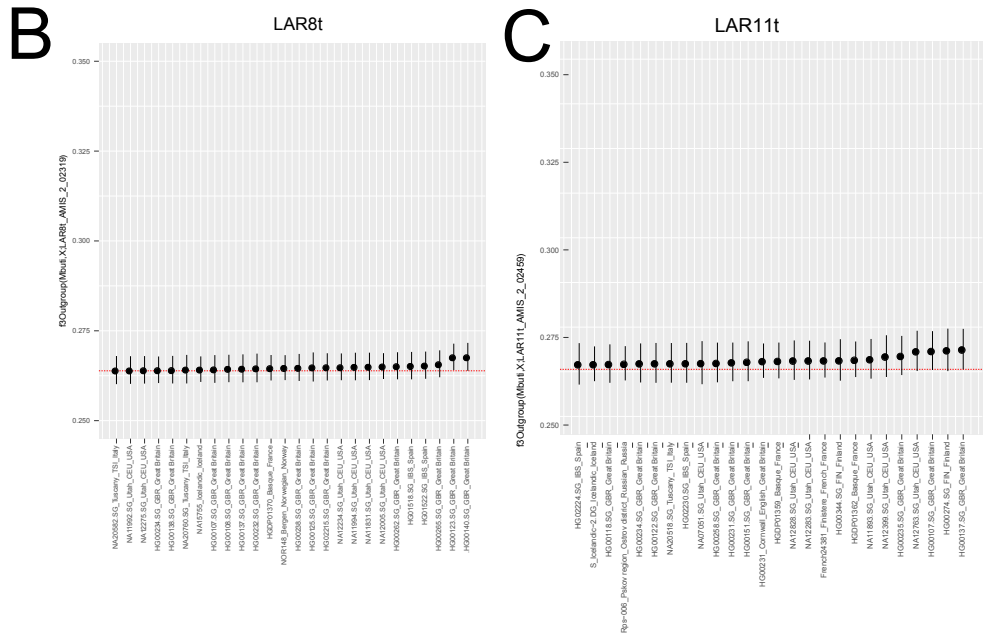
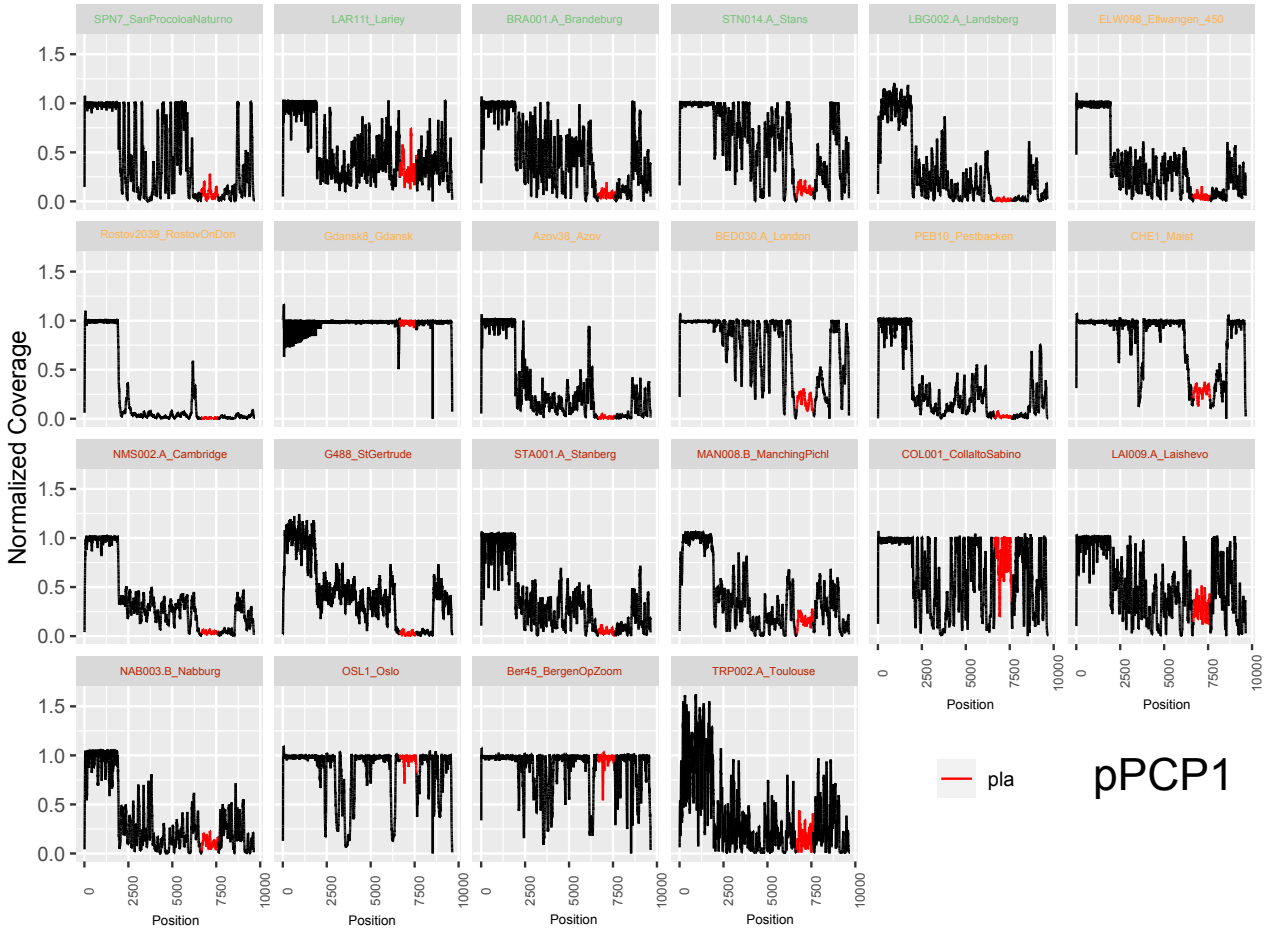

B
LAR8t
f3Outgroup(Mbuti.X,LAR8t_AMIS_2_02319)
0.350
0.325
0.300
0.275
0.250
C
LAR11t
f3Outgroup(Mbuti.X,LAR11t_AMIS_2_02459)
0.350
0.325
0.300
0.275
0.250

Figure S6. Sequence coverage profiles along the pPCP1 plasmid and the *pla* gene (red). The depth of coverage is calculated per-position and normalized to the average coverage value represented by the 90% quantile. Related to Figure 4.



Transparent Methods

Archaeological sites, archaeological and anthropological information

The cemetery from Lariey-Puy Saint Pierre was excavated in 2002 CE (Signoli et al., 2003a) on the basis of prospective archaeological digs carried out in 2001 CE (Signoli et al., 2001). This site is located in the French Alps, approximately two kilometers west of Briançon (France), and 90 kilometers west of Turin (Italy) (Figure 1A). It represents the only French site unambiguously assigned to the 1628-1632 CE plague epidemic, based on direct radiocarbon dating of four individuals (Signoli et al., 2003a) and on the presence of rare ceramic artefacts of local origin with Piedmontese influence. In addition, historical archive confirms the presence of plague in 1630 CE in Le Pinet, one of Puy St Pierre hamlets (Briançon municipal archives). The cemetery from Lariey-Puy Saint Pierre is therefore traditionally associated with years 1629-1630 CE (Bergé, 1989; Bigny, 1982). Interestingly, the plague cemetery is still part of the local popular memory today, as testified by a wooden cross engraved "*cimetière de la peste*" (literally, standing for "*plague cemetery*" in French) erected on the site and regularly replaced, and by the annual pilgrimage organized by Puy Saint Pierre villagers for Saint-Roch's day, commemorating the Saint specially invoked against the plague. The funerary ensemble covers an area of approximately ~1,800 square meters and is delimited by a stone wall, while the cemetery *per se* is limited to ~53 square meters, in which 17 adult individuals and 17 immatures could be identified. Both individual and multiple sepultures are present on site and are displayed along a north-to-south organization (Signoli et al., 2007). While single burials are concentrated in the northern area of the funerary ensemble, the presence of double and multiple burials increases towards the south. This has suggested that the funerary ensemble was first opened from the north at the time where infection and mortality rates were still limited. As infection and mortality rates increased, bodies had to be buried together within double and multiple burials, possibly as an attempt to react promptly and prevent further spread. The LAR8 individual was identified as an adult female on the basis of the coxal morpho-anatomy (Murail et al., 2005; Brooks and Suchey, 1990; Schmitt, 2005). The LAR11 individual showed characteristic of an immature individual, probably belonging to the 15-19 age class (Schaffer et al, 2009; Scheuer and Black, 2004). Preservation states of the cranial and post-cranial skeletons of the LAR8 and LAR11 individuals are illustrated in Figure S1. Archaeological evidence indicates that LAR11 was one of the first victims of the epidemic, and was buried in a single sepulture in the most remote part of the cemetery. On the other hand, LAR8, buried in the southern part of the cemetery, was part of a group of four individuals who were buried simultaneously (LAR7, a young female adult; as well as LAR9 and LAR10: children aged between 5 and 9 years old) (Figure 1B-C; Signoli et al., 2003a). As both the administrative and scientific director of the Lariey excavation, co-author M. Signoli has obtained full official permission from the French Ministry of Culture and Communication, as well as from the regional archaeology authorities (*Service Régional de l'Archéologie* SRA PACA) to carry out sampling and scientific analyses on the material excavated (site reference 05 109 7 AH). Accordingly, the analyses presented in this study have directly involved the regional archaeology authorities SRA PACA, which are represented by co-author S. Tzortzis who took part of data collection and interpretation. The Lariey remains are currently located at the laboratory *Anthropologie bio-culturelle, droit, éthique et santé* ADES, UMR 7268, Aix-Marseille University, France. Sampling took place in February and March 2019 under the authority of co-author M. Signoli, who was then the director of this laboratory.

DNA extraction

Remains belonging to 26 different individuals were processed for this study, of which 15 were subjected to DNA sequencing and 11 were tested for the presence of *Yersinia pestis* through PCR amplification targeting a short fragment within the *pla* locus. Sample preparation was performed in the CAGT laboratory (UMR5288, Toulouse, France), following strict procedures to avoid and detect potential contamination. Samples were processed (down to the PCR setup step) in state-of-the-art ancient DNA facilities, fully dedicated to the analyses of archaeological remains and physically separated from post-PCR and modern DNA laboratories. Additionally, all experimental procedures were implemented following a systematic decontamination of material and surfaces, using disposable personal protective equipment and co-processing negative blank controls together with individual samples. For 12 individuals, DNA was extracted

from 50-200mg of petrosal bone that were obtained through the pulverization on a Retsch MM 200 instrument of a short bone piece cut with a Dremel instrument. For 23 individuals, 15-50mg of dental pulp and dentin powder was drilled from the inner surface of the tooth roots and crowns, following the procedure described by Neumann and colleagues (2020). A total of 7 teeth also showed the presence of dental calculus that we collected for DNA extraction (the weight of each sample was below the scale range, *i.e.* inferior to 10mg). For each sample (including petrosal bone, tooth and dental calculus), DNA content was extracted following the protocol described in (Seguin-Orlando et al., 2021), using 963µL of lysis buffer (0.45M EDTA, 0.25mg/mL proteinase K and 0.5% N-lauryl Sarcosyl) for a first pre-digestion step of 1 hour at 37°C, before carrying out a full digestion of the pellet overnight at 42°C in 963µL of fresh lysis buffer. Dental calculus was processed using the same extraction procedure, except that the pre-digestion step was skipped. The overnight digestion buffer was centrifuged at 12,000 rpm for 2 minutes in order to remove potentially remaining pellets, before an aliquot of 200µL of the supernatant was purified on a MinElute column (QIAGEN©) and eluted in 26µL of elution buffer (EB + 0.05% Tween 20).

PCR screening

The presence of *Yersinia pestis* DNA in the tooth extracts was assessed through a PCR assay targeting a sequence of the pPCP1 *pla* gene, as previously described (Seifert et al., 2013). A total PCR reaction volume of 25µL containing 2µL of DNA extract, 1.5 units of AmpliTaq Gold™ polymerase (ThermoFisher Scientific), 1X Gold Buffer and 200nM of each primer, was subjected to a touch-down amplification program (2 cycles of 30 seconds at 94°C, 30 seconds at 64°C and 60 seconds at 72°C, followed by 2 cycles of 30 seconds at 94°C, 30 seconds at 62°C and 60 seconds at 72°C, and finally; 46 cycles of 30 seconds at 94°C, 30 seconds at 60°C and 60 seconds at 72°C). The presence of a specific 133bp amplified product was visualized by electrophoresis on a 2% agarose gel.

Uracil excision, library construction and amplification

For a total of 29 remains, originating from 15 distinct individuals (Table S1), a 22.8µL aliquot of DNA extract was subjected to Uracil-Specific Excision Reagent treatment by incubation at 37°C for 3 hours with 7µL of USER enzyme (NEB®). Illumina sequencing libraries were constructed following a well-established protocol (originally described in (Rohland et al., 2015) and modified as described in (Fages et al., 2019)) to introduce a unique 7-nucleotide barcode within both adapters P5 and P7 (those indices are referred to hereafter as '*internal*' indices as they form the first 7 nucleotide positions within each sequencing read; see lrXX, where XX is identified as a number on Table S1). Libraries were amplified and indexed by performing 8-12 PCR cycles in 25µL reaction volumes using 1 unit of AccuPrime™ Pfx DNA polymerase, 4-6µL of DNA library and with an overall concentration of 200nM of both the InPE1.0 primer and one custom PCR primer including a unique 6-nucleotide index (this index is hereafter referred to as '*external*' index as its sequence is obtained following the priming of a sequencing reaction independent from that leading to each read pair; see iXX, where XX is identified as a number on Table S1). Amplified products were purified using Agencourt Ampure XP beads (1.4:1 or 1.6:1 as beads:DNA ratio) and eluted in 20µL EB+0.05% tween. Library molarity, size and concentration were checked on a Tapestation 4200 instrument (Agilent Technologies) and on a QuBit HS dsDNA assay (Invitrogen). Up to 3 amplified libraries were obtained per sample, for a total of 45 DNA libraries (Table S1).

DNA Sequencing

Amplified libraries were pooled with other indexed libraries and sequenced using the Paired-End mode on an Illumina MiniSeq instrument (2x80 bp reads) at the CAGT laboratory (Toulouse, France) or on a NovaSeq S4 (2x150 bp reads) at SciLifeLab (Stockholm, Sweden) in order to obtain deeper genome coverage.

Read processing, alignment, trimming and rescaling

Illumina paired-end reads were demultiplexed according to the 7-bp 'internal' indexes present at both read starts (Rohland et al., 2015), using a maximum edit distance of 1 mismatch per individual index. Demultiplexed reads were then trimmed for adapter sequences (--mm 5) and poor-quality ends (Phred quality scores ≤ 2), and collapsed into single reads using AdapterRemoval2 (Schubert et al., 2016). Collapsed, collapsed truncated and non-collapsed pairs were mapped against a number of reference genome sequences, including the human reference genome (GRCh37, hg19), the human revised Cambridge reference sequence (rCRS, Genbank Accession Number NC_120920.1), the plague reference genome (strain CO92, Accession Number = NC_003143.1 ; (Parkhill et al., 2001)), and each individual plasmid, including pCD1 (Accession Number = NC_003131.1), pMT1 (Accession Number = NC_003134.1), and pPCP1 (Accession Number = AL109969.1). As the plague chromosome is circular, the first and last 30 bp of the CO92 chromosome sequence were duplicated at the end and the start of the CO92 chromosome sequence in order to identify those reads extending across both regions. The same was done to handle the circularity of each plasmid, except that only the first 30 bp were copied at the end of each individual plasmid reference sequence. Read alignment against the human nuclear and mitochondrial genomes was carried out using Bowtie2 (Langmead et al., 2012) according to the recommendations of Poullet and Orlando (2020) and the procedures described by Seguin-Orlando and colleagues (2021). Read alignment against the plague chromosome and plasmids was carried out using BWA backtrack v0.7.17-r1194-dirty and followed the stringent parameters described by Spyrou and colleagues (2019b). All aligned reads shorter than 25 bp and alignments showing mapping quality strictly inferior to 30 were disregarded. PCR duplicates were removed using MarkDuplicates from Picard Tools (version 2.18.0, http://broadinstitute.github.io/picard/) and read were locally re-aligned around indels using GATK (version 3.8.1, (McKenna et al., 2010)). Read collapsing, trimming and mapping, as well as PCR duplicate removal and local realignment were carried out using the Paleomix automated computational pipeline (version 1.2.13.2 (Schubert et al., 2014)). Sequencing statistics, including numbers of sequencing reads, endogenous DNA content, clonality, and coverage are provided in Table S1. We downloaded previously published sequence data from 30 *Yersinia pseudotuberculosis* strains, 2 *Yersinia similis*, as well as 127 modern and 133 ancient *Yersinia pestis* strains (Table S2), representing a total of 24.8 billion reads. Raw fastq reads (or read pairs) were subjected to the same processing steps as described above so as to obtain an extensive comparative genome panel. We noticed that the ancient genomes previously characterized combined a full array of archaeological sites and experimental methods (including DNA extracts treated or not treated using the USER-enzyme (Rohland et al., 2015), shotgun sequencing and capture data). They were, thus, likely affected by various levels of post-mortem DNA damage, resulting in the presence of different proportions of nucleotide mis-incorporations in our comparative panel. In order to mitigate such effects, all plague BAM alignment files were subjected to the following procedure. First, we used PMDtools (Skoglund et al., 2014) to identify those aligned reads carrying post-mortem DNA damage signatures, using the conservative score of 1. The base quality of all read positions affected by signatures of post-mortem Cytosine deamination (i.e. C-to-T and G-to-A transitions, relative to the reference) was further downscaled using mapDamage v2.7 (Jónsson et al., 2013). This effectively resulted in the elimination of all mutations potentially introduced by post-mortem DNA damage. Once rescaled, read alignments were further trimmed across the first and last 10 positions to eliminate additional sources of spurious variation potentially present. Those aligned reads that did not show any evidence of post-mortem DNA damage (PMD score threshold < 1) were trimmed the first and last 5 read positions only. Aligned reads were then merged together using samtools merge (Li et al., 2009) in order to obtain final BAM files for downstream analyses.

Human genome analyses

Individual sex was inferred on the basis of X-to-autosomal sequence coverage. Mitochondrial haplotypes were called using haplogrep (version 2.2 (Kloss-Brandstätter et al., 2011)), minimal mapping and base quality thresholds of 30 and a minimal depth filter of 5 (or 3 for those individuals showing minimal coverage; Table S1), following the procedure from Seguin-Orlando and colleagues (2021). Contamination rates based on mitochondrial data were estimated using Schmutzi (Renaud et al., 2014). The Y-chromosome haplotype carried by individual LAR11 was called using the Yleaf statistical package (Ralf et al., 2018).

Contamination rates were also estimated for males using heterozygosity measurements at polymorphic sites present on the X chromosome, following the methodology from (Rasmussen et al., 2011) and implemented in ANGSD (Korneliussen et al., 2013), excluding transition substitutions and sites covered only once or more than 200-times. Relatedness between the LAR8 and LAR11 individuals was assessed using lcMLkin (Lipatov et al., 2015), considering all autosomal positions overlapping the 1240K dataset (v42.4 available at https://reichdata.hms.harvard.edu/pub/datasets/amh_repo/curated_releases/V42/V42.4/SHARE/public.dir/v42.4.1240K_HO.tar) and random sampling one read at those sites covered multiple times. The sequence data was converted to vcf prior to running lcMLkin using VCFtools (version 0.1.17) (Danecek et al., 2011), and pruned using the SNPbam2vcf.py script for sites located at least 100 kb apart and present at minimal 5% allelic frequencies (--thin 100000 --maf 0.05). Principal Component Analysis (PCA) was carried out using the Human Origin reference panel for 592,998 autosomal genotypes in 796 modern west Eurasian individuals, as reported by Patterson and colleagues (2012) and Lazaridis and colleagues (2014). The analysis was based on pseudo-diploid genotype calls due to the limited genome coverage achieved for both the LAR8 and LAR11 individuals. PCA was carried out using smartPCA from EIGENSOFT version 7.2.1 (Patterson et al., 2006), and turning the lsqproject and shrink options on for projecting the two ancient individuals onto the two first principal components (PCs) obtained from modern reference individuals. Finally, we carried out f3-Outgroup statistics using q3Pop from Admixtools (version 5.0 (Patterson et al., 2012)) to identify those modern individuals showing highest genetic affinities with the two ancient individuals sequenced in this study. The comparative panel included a total number of 6,152 individuals from across the world, including 4 Mbuti individuals who were used as outgroup (S_Mbuti-3.DG, B_Mbuti-4.DG, S_Mbuti-2.DG, S_Mbuti-1.DG).

Metagenomic profiling

Microbial taxonomic profiles of each individual DNA library were determined using the metaBIT automated computational package (Louvel et al., 2016), disregarding collapsed truncated and uncollapsed pairs as well as collapsed reads showing high-quality (mapping quality ≥ 30) unique alignments against the human genome, and restricting assignments to non-viral and non-eukaryotic taxa present in the MetaPhlAn2 diversity database (Truong et al., 2015). Bacterial taxa supported by abundances lower than 1% at either the species level (Figure 1D) or the genus level (Figure S2) were disregarded. Unsupervised clustering of abundance profiles was carried out at the species level using the pvclust R package (https://cran.r-project.org/web/packages/pvclust/index.html) (Suzuki and Shimodaira, 2006), assessing node support through approximately-unbiased (au) tests and 10,000 bootstrap pseudo-replicates. Principal Coordinate Analyses were carried out based on Bray-Curtis distances to compare the taxonomic profiles of all individual DNA libraries sequenced in this study. The analyses were repeated to a broader panel, including a range of modern microbial profiles, including from 15 soil samples as well as 689 human-associated modern microbiota obtained from (Fierer et al., 2012; The Human Microbiome Project Consortium, 2012). Linear Discriminant Analyses were carried out in LEfSe with default parameters (Segata et al., 2011), and grouping samples in four categories corresponding to *Yersinia pestis*-positive teeth, *Yersinia pestis*-negative teeth, petrosal bones and dental calculus.

Phylogenetic analyses

Maximum Likelihood phylogenetic trees were constructed using IQ-Tree v1.6.12 (Minh et al., 2020). The best mutational model (TVM+F+R6) was automatically selected using AICc from a panel of 283 individual models corresponding to various combinations of 22 common substitution models, site frequency models (F), and rate heterogeneity across sites (G). Base frequencies were optimized by Maximum Likelihood (ML) during phylogenetic reconstruction. Sequence alignments corresponded to those CO92 chromosome positions that were polymorphic (disregarding indels) and covered in at least half of the plague chromosomes. Any strain showing less than 75% of the resulting sites was further removed from the alignment, resulting in a total of 21,279 sites, including 17,688 parsimony-informative and 3,591 singletons. Individual base positions were called using BCFtools (version 1.8-31-g9ba4024, using htslib 1.7-41-g816a220; (https://github.com/samtools/bcftools)), especially considering minimum base and mapping

Phred scores of 30 in the mpileup module, turning the --ploidy flag to 1 in the call module, and filtering those variants located within a 10 bp range of indels and those showing a genotype Phred score strictly inferior to 30. Positions showing coverage superior to the 99.5% quantile of the position-wise coverage distribution were disregarded as potentially indicative of undetected chromosomal rearrangements such as duplications. The individual threshold per strain was obtained from the Paleomix depths command with default parameter (Schubert et al., 2014). Node support was estimated using a total of 1,000 ultrafast bootstrap (UFBoot) approximation (Hoang et al., 2018) and the SH-like approximate likelihood ratio (aLRT) test with 1,000 replicates (Guindon et al., 2010).

Bayesian phylogenetic reconstructions were carried out in BEAST v2.6.3 (Bouckaert et al., 2019), following the methodology from Spyrou and colleagues (2019b). In short, the sequence alignment used in IQ-Tree was restricted to those strains that formed the second plague pandemic as well as those belonging to the same main phylogenetic cluster, which included modern third pandemic strains (see Table S2 for the final list of samples considered). The Beast2 xml file was prepared in BEAUti 2 (Drummond et al., 2012), importing the fasta alignment and providing absolute dates for tips (before present). The GTR+G8 substitution model was considered and mutation rate variation (uniform prior range between 1.e-2 to 1.e-10 substitution per site per year) along the tree branch lengths was modeled using the Relaxed Clock Log Normal clock model (Drummond et al., 2006). Root-to-top regression in TempEst v1.5.3 (http://tree.bio.ed.ac.uk/software/tempest/) using the IQtree ML-tree obtained on the data underlying the Beast2 analysis, and constructed without molecular clock assumption, supported temporal structure in the data (R squared function, correlation coefficient = 0.769). Effective population size could vary according to the stepwise coalescent Skyline model (Drummond et al., 2005), considering a total of 20 individual time bins. The tree topology, branch length and the other model parameters were estimated following 750,000,000 MCMC iterations, sampling 1 very 100 state. A final set of 50,000 trees was obtained using LogCombiner disregarding the first 20% as burn-in and sampling a final subset of 50,000 trees with LogCombiner (https://www.beast2.org/programs/). The maximum clade credibility tree was obtained in TreeAnnotator (https://www.beast2.org/treeannotator/). MCMC convergence was assessed through ESS values superior to 200 for all relevant model parameters. All trees were plotted using the R ggtree library (Yu et al., 2018).

Genome characteristics

Patterns of depth-of-coverage variation along the CO92 chromosome were calculated for each individual strain using 1,000 bp non-overlapping sliding windows using the Paleomix coverage command (Schubert et al., 2014) (Figure 2). Depth-of-overage calculations were carried out within 100 bp non-overlapping sliding windows for the three individual plasmids. Additionally, patterns of %GC variation along the CO92 chromosome and the three plasmids were calculated within the same windows and using seqtk (https://github.com/lh3/seqtk). Circular plots were plotted using the R circlize library (Gu, 2014). Gene coverage was calculated using the Paleomix depths module (Schubert et al., 2014) for a total of 163 chromosome and 44 plasmid regions previously described to be associated with virulence characteristics (Zhou et al., 2004; Zhou and Yang, 2009; Keller et al., 2019; Spyrou et al., 2019b). Here, gene coverage was calculated as the fraction of the positions that was covered at least once. Gene coverage was plotted using the heatmap function from the R ggplot2 library (https://cran.r-project.org/web/packages/ggplot2/index.html) (Figure 4A). Synonymous and non-synonymous sequence polymorphisms were annotated using snpToolkit v2.0.6 (Namouchi et al., 2018) and the individual vcf files generated as described in the previous section. The minimal depth threshold was set to 3 (DP4≥3) for bases showing Phred scores above or equal to 30. The allele present in a given strain was then identified as long as it was present in at least 90% of the reads. We also calculated the overall heterozygosity observed along the CO92 chromosome using VCFtools and the same parameters as those described in the previous section, except that the –ploidy option was set to 2. Edit distance distributions were generated from the NM:i field obtained while running the samtools view command. Coverage variation along the pPCP1 plasmid shown as Figure S6 is calculated per-position and normalized to the average coverage value represented by the 90% quantile.

Data availability

Raw sequence data and alignments are available at the European Nucleotide Archive (ENA) under accession number PRJEB43291.

Supplemental References

Bergé, F. (1989). Puy-Saint-Pierre : aperçu historique, mémoire d'un village (Escartons Briançon).

Bligny, B. (1982). Histoire du Dauphiné (Privat).

Brooks, S. and Suchey, J. (1990). Skeletal age determination base on the Os Pubis: A Comparison of the Acsádi-Nemeskéri and Suchey-Brooks Methods. Hum Evol*, 5,* 227-238.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Drummond, A.J., Rambaut, A., Shapiro, B., and Pybus, O.G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol Evol. *22*, 1185–1192.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., and Rambaut, A. (2006). Relaxed Phylogenetics and Dating with Confidence. PLoS Biol. *4*, e88.

Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. *29*, 1969–1973.

Fages, A., Hanghøj, K., Khan, N., Gaunitz, C., Seguin-Orlando, A., Leonardi, M., McCrory Constantz, C., Gamba, C., Al-Rasheid, K.A.S., Albizuri, S., et al. (2019). Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series. Cell *177*, 1419-1435.e31.

Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H., and Caporaso, J.G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proc. Natl. Acad. Sci. U S A *109*, 21390–21395.

The Human Microbiome Project Consortium (2012). A framework for human microbiome research. Nature *486*, 215–221.

Korneliussen, T.S., Moltke, I., Albrechtsen, A., and Nielsen, R. (2013). Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics *14*, 289.

Kloss-Brandstätter, A., Pacher, D., Schönherr, S., Weissensteiner, H., Binna, R., Specht, G., Kronenberg, F. (2011). HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Hum. Mutat. *32*, 25–32.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., et al. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature *513*, 409–413.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lipatov, M., Sanjeev, K., Patro, R. and Veeramah, K. (2015). Maximum Likelihood Estimation of Biological Relatedness from Low Coverage Sequencing Data. doi:10.1101/023374.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

Murail, P., Bruzek, J., Houët, F. and Cunha, E. (2005). DSP : a tool for probabilistic sex diagnosis using worlwide variability in hip bone measurements. BMSAP *17*, 167-176.

Neumann, G.U., Andrades Valtueña, A., Fellows Yates, J.A., Stahl, R., and Brandt, G. (2020). Tooth Sampling from the inner pulp chamber for ancient DNA Extraction v1 (protocols.io.bakqicvw). https://doi.org/10.17504/protocols.io.bakqicvw

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

Poullet, M., and Orlando, L. (2020). Assessing DNA Sequence Alignment Methods for Characterizing Ancient Genomes and Methylomes. Frontiers Ecol. Evol. *8*, 105.

Ralf, A., Montiel González, D., Zhong, K., and Kayser, M. (2018). Yleaf: Software for Human Y-Chromosomal Haplogroup Inference from Next-Generation Sequencing Data. Mol. Biol. Evol. *35*, 1291–1294.

Rasmussen, M., Guo, X., Wang, Y., Lohmueller, K.E., Rasmussen, S., Albrechtsen, A., Skotte, L., Lindgreen, S., Metspalu, M., Jombart, T., et al. (2011). An Aboriginal Australian genome reveals separate human dispersals into Asia. Science *334*, 94–98.

Renaud, G., Slon, V., Duggan, A.T. and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. *16*, 224.

Schaefer, M., Black, S. and Scheuer, L. (2009) Juvenile Osteology: A laboratory and Field Manual (Academic Press London).

Scheuer, L. and Black, S. (2004).The juvenile skeleton. (Academic Press London).

Schmitt, A. (2005). Une nouvelle méthode pour estimer l'âge au décès des adultes à partir de la surface sacro-pelvienne iliaque. *BMSAP 17, 1-13.*

Schubert, M., Lindgreen, S. and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes *9*, 88.

Seguin-Orlando, A., Donat, R., Der Sarkissian, C., Southon, J., Thèves, C., Manen, C., Tchérémissinoff, Y., Crubézy, E., Shapiro, B., Deleuze, J.F., et al. (2021). Heterogeneous hunter-gatherer and steppe-related ancestries in Late Neolithic and Bell Beaker genomes from present-day France. Curr. Biol. https://doi.org/10.1016/j.cub.2020.12.015.

Seifert, L., Harbeck, M., Thomas, A., Hoke, N., Zöller, L., Wiechmann, I., Grupe, G., Scholz, H.C. and Riehm, J.M. (2013). Strategy for Sensitive and Specific Detection of *Yersinia pestis* in Skeletons of the Black Death Pandemic. PLoS ONE *8*, e75742.

Signoli, M. (2001). *Le cimetière des pestiférés de Lariey, (Puy-Saint-Pierre, Hautes-Alpes), rapport d'opération de sondages* (SRA PACA).

Signoli, M., Tzortzis, S., Bizot, B., Ardagna, Y., Rigeade, C., and Seguy, I. (2007). Découverte d'un cimetière de pestiférés du XVII^{ème} siècle (Puy-Saint-Pierre, Hautes-Alpes, France). *In* LA PESTE : entre épidémies et sociétés, *Signoli, M., Cheve, D., Adalian, P., Boëtsch, G. and Dutour, O., ed. (*Firenze University Press).

Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J., and Jakobsson, M. (2014). Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proc. Natl. Acad. Sci. U S A *111*, 2229–2234.

Yu, G., Lam, T.T.-Y., Zhu, H., and Guan, Y. (2018). Two methods for mapping and visualizing associated data on phylogeny using ggtree. Mol. Biol. Evol *35*, 3041–3043.

Zhou, D., Tong, Z., Song, Y., Han, Y., Pei, D., Pang, X., Zhai, J., Li, M., Cui, B., Qi, Z., et al. (2004). Genetics of Metabolic Variations between *Yersinia pestis* Biovars and the Proposal of a New Biovar, *microtus*. J. Bacteriol. *186*, 5147–5152.

Zhou, D., and Yang, R. (2009). Molecular Darwinian Evolution of Virulence in *Yersinia pestis*. Infect. Immun. *77*, 2242–2250.