Improving the extraction of ancient Yersinia pestis genomes from the dental pulp

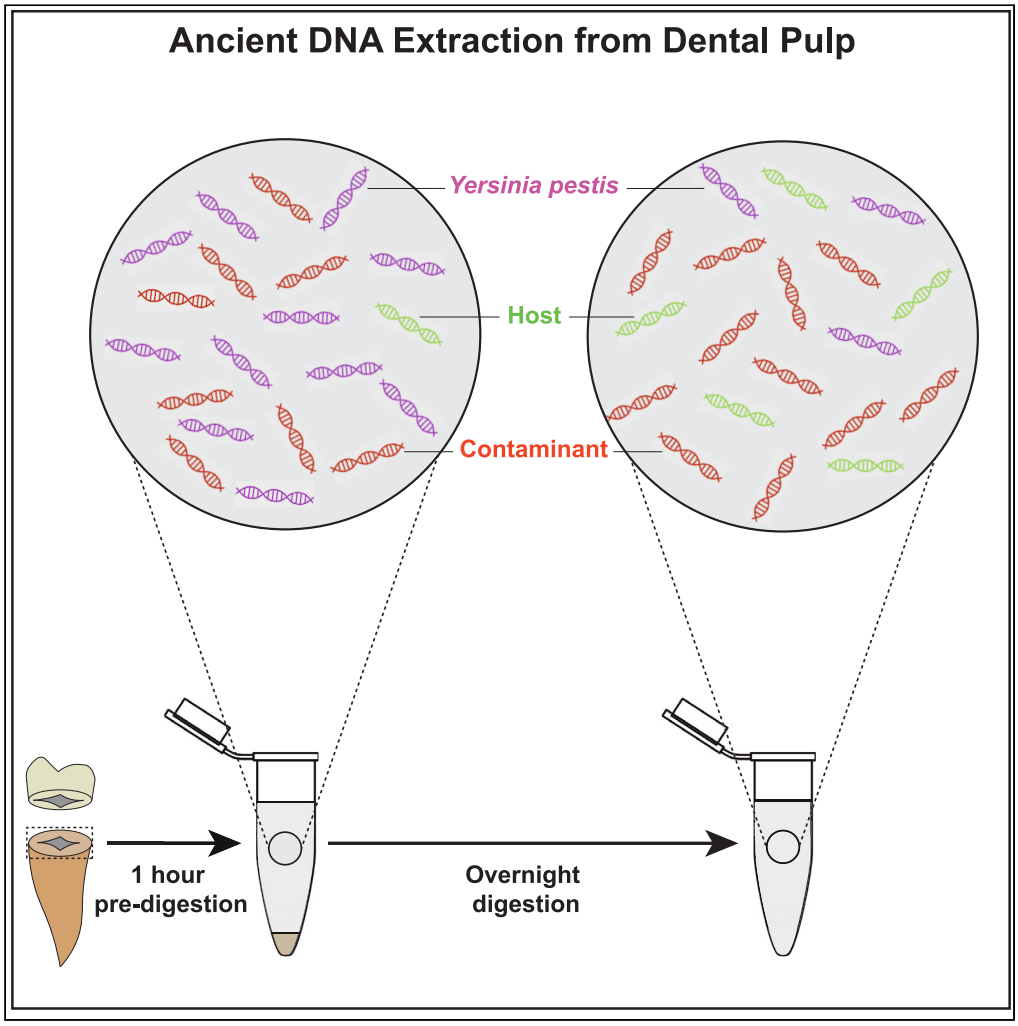
Pierre Clavel, Lexane Louis, Clio Der Sarkissian, Catherine Thèves, Claudia Gillet, Lorelei Chauvey, Gaétan Tressières, Stéphanie Schiavinato, Laure Calvière-Tonasso, Norbert Telmon, et al.

iScience

CellPress
OPEN ACCESS

Article

Improving the extraction of ancient *Yersinia pestis* genomes from the dental pulp



Pierre Clavel, Lexane Louis, Clio Der Sarkissian, ..., Maria A. Spyrou, Andaine Seguin-Orlando, Ludovic Orlando

ludovic.orlando@univ-tlse3.fr

Highlights

Short tooth powder digestion times improve recovery of *Yersinia pestis* ancient DNA

Variable *Y. pestis* detection success across teeth from the same individual

Sequencing of 12 second plague pandemic genomes from France

Different strains during the Thirty Years' War and the Great Plague of Marseille

Clavel et al., iScience *26*, 106787
May 19, 2023
https://doi.org/10.1016/j.isci.2023.106787

iScience

CellPress
OPEN ACCESS

Article

Improving the extraction of ancient *Yersinia pestis* genomes from the dental pulp

Pierre Clavel,[1] Lexane Louis,[1] Clio Der Sarkissian,[1] Catherine Thèves,[1] Claudia Gillet,[1] Lorelei Chauvey,[1] Gaétan Tressières,[1] Stéphanie Schiavinato,[1] Laure Calvière-Tonasso,[1] Norbert Telmon,[1] Benoît Clavel,[2] Richard Jonvel,[3] Stéfan Tzortzis,[4] Laetitia Bouniol,[5] Jean-Marc Fémolant,[5] Jennifer Klunk,[6] Hendrik Poinar,[7,8,9] Michel Signoli,[10] Caroline Costedoat,[10] Maria A. Spyrou,[11] Andaine Seguin-Orlando,[1] and Ludovic Orlando[1,12,*]

SUMMARY

Ancient DNA preserved in the dental pulp offers the opportunity to characterize the genome of some of the deadliest pathogens in human history. However, while DNA capture technologies help, focus sequencing efforts, and therefore, reduce experimental costs, the recovery of ancient pathogen DNA remains challenging. Here, we tracked the kinetics of ancient *Yersinia pestis* DNA release in solution during a pre-digestion of the dental pulp. We found that most of the ancient *Y. pestis* DNA is released within 60 min at 37°C in our experimental conditions. We recommend a simple pre-digestion as an economical procedure to obtain extracts enriched in ancient pathogen DNA, as longer digestion times release other types of templates, including host DNA. Combining this procedure with DNA capture, we characterized the genome sequences of 12 ancient *Y. pestis* bacteria from France dating to the second pandemic outbreaks of the 17th and 18th centuries Common Era.

INTRODUCTION

Ancient DNA preserved within archaeological and paleontological remains provides genetic information about our evolutionary past.[1] The increasing throughput of next-generation sequencing instruments has considerably improved our capacity to characterize ancient genomes, despite environmental microbes often producing the dominant fraction of the DNA extracts.[2] Specific types of remains, including petrosal bones,[3] ossicles,[4] and the tooth cementum,[5] are known to be generally less prone to environmental DNA contamination. Together with improved decontamination and DNA extraction procedures, involving various pre-digestion treatments of sample powder through bleaching[6,7] and/or washing buffers,[8–10] these biological remains have facilitated the characterization of ancient genomes. The development of DNA capture methodologies targeting million-scale SNP (Single Nucleotide Polymorphism) panels,[11,12] whole chromosomes[13] or even the entire genome[14,15] have provided a further opportunity to focus sequencing efforts on the DNA fraction of interest, therefore, increasing sensitivity and reducing sequencing costs. As a result, there has been an acceleration since the first ancient human genomes have been sequenced in 2010[16–18] to the thousands of ancient human genomes characterized, with recent human paleogenomic studies typically including hundreds of individuals.[19,20]

Besides host and environmental microbial material, ancient DNA extracts can also preserve molecules originating from pathogens that circulated in the blood system when the host died.[21] In cases where the pathogenic load is high and environmental contamination is low, shotgun sequencing has allowed for the successful characterization of the entire genome sequence from ancient pathogens, including viruses (e.g., smallpox,[22,23] hepatitis B[24,25]), bacteria (e.g., *Mycobacterium leprae*[26]), and protozoan parasites (e.g., *Plasmodium vivax*[27]). Reconstructing ancient pathogen genomes, however, most commonly requires DNA capture technologies, both due to the molecular complexity of ancient DNA extracts and the generally limited genome size of pathogens, relative to their hosts.[21,28] Additionally, the type of material considered for ancient pathogen genome characterization is critical, as tissues favoring host DNA preservation, such as petrosal bones, show virtually no bacterial pathogen DNA.[29,30] Despite these limitations, more than five hundred ancient pathogen genomes have been sequenced,[31] revealing important insights into the epidemiological outbreaks with substantive morbidity and mortality.

Continued

[1]Centre d'Anthropobiologie et de Génomique de Toulouse (CAGT), CNRS UMR5288, Université Paul Sabatier, 37 allées Jules Guesde, 31000 Toulouse, France

[2]Archéozoologie, Archéobotanique: Sociétés, Pratiques et Environnements (AASPE), CNRS-UMR7209, Muséum national d'histoire naturelle, 55 Rue Buffon, 75005 Paris, France

[3]Amiens Métropole Service Archéologie Préventive, 2 rue Colbert, 80000 Amiens, France

[4]Service Régional de l'Archéologie, 21 allée Claude Forbin, 13100 Aix-en-Provence, France

[5]Service archéologique de la ville de Beauvais, 1 rue Desgroux, 60021 Beauvais, France

[6]Daicel Arbor Biosciences, Ann Arbor, MI 48103 USA

[7]McMaster Ancient DNA Centre, Departments of Anthropology, Biology and Biochemistry, McMaster University, Hamilton, ON L8S 4L9, Canada

[8]Michael G. DeGroote Institute of Infectious Disease Research, McMaster University, Hamilton, ON L8S, 4L9, Canada

[9]Humans and the Microbiome Program, Canadian Institute for Advanced Research, Toronto, ON, Canada

[10]Aix-Marseille Université, CNRS, EFS, ADES, 13005 Marseille, France

[11]Institute for Archaeological Sciences, Eberhard Karls University of Tübingen, Tübingen, Germany

[12]Lead contact

Plague represents the best-studied infectious disease from an ancient DNA perspective.[32] DNA analyses of human individuals buried during the Justinian plague and the Black Death have confirmed *Y. pestis* as the etiological agent responsible for massive human loss in the sixth-eighth century Common Era (c. CE)[33,34] and the 14th c. CE,[35] respectively. Such analyses have, however, revealed that plague already circulated during the late Neolithic and early Bronze Age,[36–41] and have started reconstructing the evolutionary trajectory of key genetic changes underlying virulence and transmission.[38,42] Genetic evidence has also tracked the expansion of the second plague pandemic from a Central Asian homeland[43] and the following diversification of several lineages through space and time.[44]

Despite such advances, the characterization of ancient pathogen genomes remains difficult. Relatively low detection rates and minimal target ancient DNA contents are commonly observed in mass graves, even when the association with known pandemic, such as the second plague pandemic, is documented historically.[45] While the performance of various DNA capture systems is well established,[46–48] no single study has attempted to optimize the molecular toolkit for pathogen DNA extraction and capture. In this study, we investigated the temporal dynamics of plague and host DNA release from dental pulp, using a total of 120 human teeth dated to the 17th and 18th c. CE. These remains originated from five French archaeological sites associated with the second plague pandemic. We found that the vast majority of the pathogen DNA can be recovered following a short pre-digestion wash of the dental pulp, while the recovery of the host DNA benefits from further digestion. Therefore, focusing on the pre-digestion wash improves the ratio of *Y. pestis*-to-host DNA. Whether such an approach could also enhance the recovery of other ancient pathogens in different preservation contexts and/or tissue types remains to be tested. Combined with targeted capture, pre-digestion washes helped us characterize 12 ancient *Y. pestis* genomes, including from material showing limited DNA preservation. Our work contributes to ongoing efforts aimed at mapping the genetic diversity of second pandemic plague strains.

RESULTS AND DISCUSSION

Sample characteristics

We collected 120 teeth from 89 human individuals that were excavated from French mass graves and cemeteries related to the second plague pandemic (Figure 1, Data S1). More specifically, a total of 56 teeth originated from the sites of La Major (1720 CE, Marseille) and Les Rayettes (1721 CE, Martigues), both associated with the last major plague outbreak in France. This epidemic started after the ship *Grand Saint Antoine* departed from the Levant and landed on the Marseille harbor in 1720 CE. It rapidly decimated half of the population of the city, before it spread to the Provence region, hitting Marseille a second time (causing considerably lower mortality), and finally vanished in 1722 CE.[49] Overall, it is estimated that this plague outbreak killed more than a quarter of the population of the Provence region. The remaining 64 teeth were collected from 58 individuals buried during the 17th c. CE, including from the Lariey Puy-Saint-Pierre cemetery (1629–1630 CE), as well as the Maladrerie Saint-Lazare (1630–1670 CE, Beauvais) and the Hôtel-Dieu Saint-Jean-Baptiste (1630–1670 CE, Amiens). While the Lariey cemetery was previously analyzed for *Y. pestis* and human ancient DNA,[45] the latter two sites were genetically investigated here for the first time. These sites correspond to plague mass graves formed at the time of the Thirty Years' War.

A total of 33 teeth were first subjected to the standard procedure for ancient DNA extraction routinely applied at the Centre for Anthropobiology and Genomics of Toulouse (CAGT) ancient DNA facilities. Briefly, the dental pulp powder is collected by cutting the tooth at the cementum-enamel junction in order to drill the pulp chamber from inside with a diamond ball bit. The powder is first pre-digested in a DNA extraction buffer for 1 h at 37°C. The undigested tooth pellets are then collected and digested until completion in a fresh DNA extraction buffer, following overnight incubation with agitation at 42°C. This procedure was previously optimized for recovering ancient DNA from the host,[50] but nonetheless led to the successful characterization of two ancient human and *Y. pestis* genomes from the Lariey cemetery, following shotgun sequencing on Illumina platforms.[45] Stringent alignment of shallow sequencing data against the *Y. pestis* reference genome[34,51] revealed limited, if any, *Y. pestis* DNA preservation across all the samples investigated (< 0.93%), regardless of the archaeological site considered (Table S1). DNA capture procedures were, thus, implemented to attempt to enrich for *Y. pestis* and human mtDNA templates, using the synthetic RNA 80-mers probes designed by Wagner and colleagues.[33] In this experiment, a total of 5 and 55 amplified DNA libraries were selected from individuals excavated at Lariey and La Major, respectively, as representatives of the entire range of DNA preservation conditions. Bioinformatic demultiplexing and collapsing of capture DNA reads provided between 14,386 and 2,250,565 sequences per

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

https://doi.org/10.1016/j.isci.2023.106787

Figure 1. Sample corpus

(A) Sample location and dates. The archaeological sites analyzed in this study are all from France and are shown in purple, with respect to those previously published, following the coloring scheme from Spyrou and colleagues,[43] for consistency. Pictures of the mass graves studied with the exception of the archaeological site of Lariey (12 teeth), which was investigated by Seguin-Orlando and colleagues[45] (the number of teeth investigated, N, is indicated on each picture).

(B) La Major (Marseille, 1720 CE).

(C) Les Rayettes (Martigues, 1721 CE).

(D) Hôtel-Dieu Saint-Jean-Baptiste (Amiens, 1630–1670 CE).

(E) Maladrerie Saint-Lazare (Beauvais, 1630–1670 CE). See also Tables S1 and S2.

individual, of which only those from two individuals from Lariey (LAR8 and LAR11) provided sufficient coverage of the *Y. pestis* genome (Table S1). Further metagenomics analyses in MetaPhlAn 4[52] confirmed a greater *Y. pestis* content in capture versus shotgun collapsed reads for 12 libraries prepared from six individual teeth (Figures 2A and 2B). This supported numbers of classified microbial species dropping with capture, as expected for a procedure aimed at focusing sequencing efforts on *Y. pestis* only. Accordingly, the fraction of environmental and oral bacteria identified following capture was reduced relative to shotgun experiments (Figure 2B).

While targeted DNA capture improved the *Y. pestis* DNA recovery rate, the vast majority of the sequence data for other individuals consisted of PCR duplicates (37.72–98.03%), which limited the cost-effectiveness as an approach for characterizing the plague genome. The high proportion of PCR duplicates indicated that the total amount of plague DNA templates present in the original extracts was extremely limited. We, thus, decided to identify DNA extraction procedures that could enhance the recovery of ancient plague DNA molecules.

Pre-digestion DNA release

Previous work reported increased proportions of host DNA in libraries constructed following complete digestion, relative to those obtained pre-digestion.[8,9] This suggested that pre-digestion could wash away a fraction of the environmental microbial DNA in contrast to the host DNA, which would be preserved deeper into the calcified matrix, and thus would require longer digestion time to be released.[53] We reasoned that the *Y. pestis* DNA would be mostly restricted to the blood circulation and would not

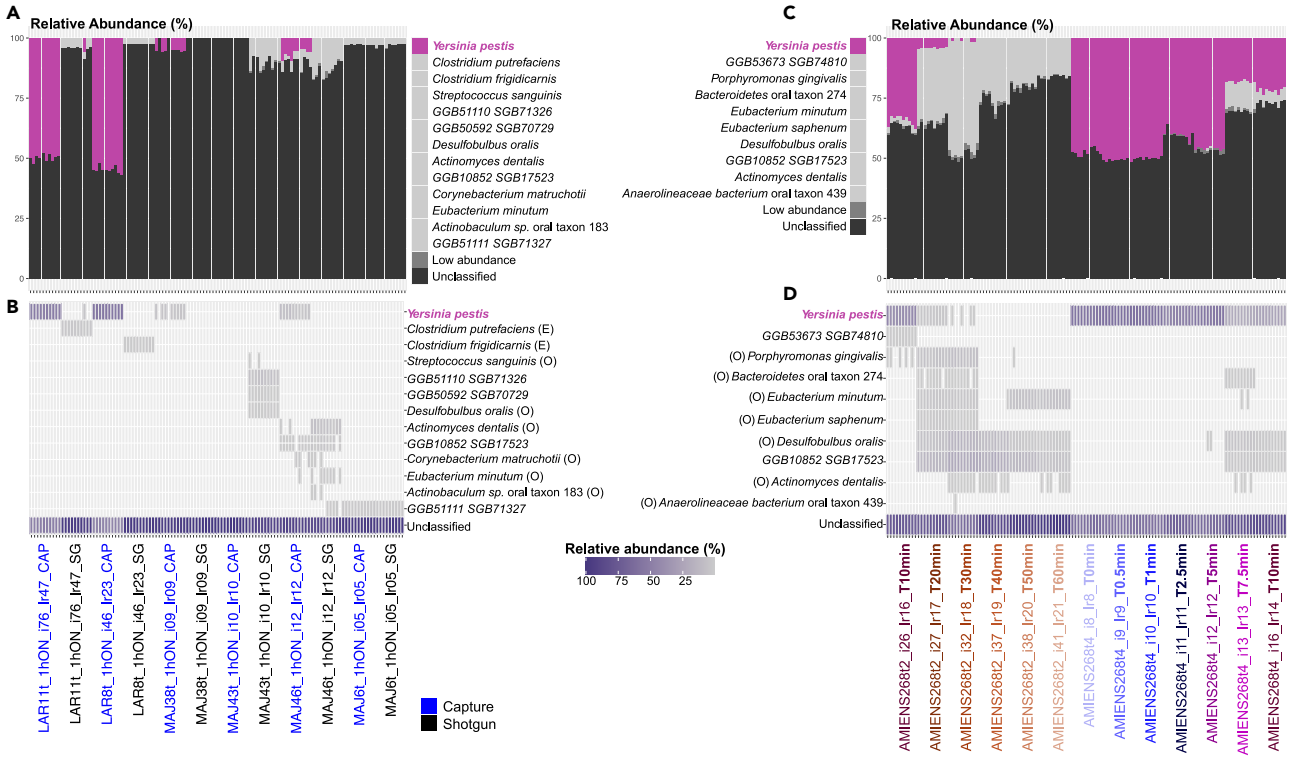


Figure 2. MetaPhlAn 4 microbial taxonomic profiles of the tooth material analyzed

(A and B) Comparison of microbial diversity in shotgun versus capture sequence datasets. These experiments were carried out on the extracts resulting from a 1-h pre-digestion, followed by an overnight digestion of the remaining pellets (1h+ON). (A, B) Relative abundances of microbial species.

(C and D) Same as A, B except that different pre-digestion times were contrasted from 10 to 60 min and from 0 to 10 min. The analyses were restricted to two teeth from the same sample (AMIENS268t4 and AMIENS268t2).

penetrate the deeper extracellular matrix. Under this assumption the pathogen DNA would follow release kinetics similar to the environmental microbial DNA. To test this theory, we purified the DNA from 17 1-h pre-digested fractions (1h) to prepare libraries for the capture of *Y. pestis* and human mtDNA genomes. The DNA contents of the pre-digested samples were compared to those obtained following capture of the DNA libraries prepared after complete pellet digestion (1h+O/N; Figures 3A and 3B). Normalizing for sequencing efforts after *Y. pestis* whole-genome capture, we found that most of the libraries constructed following 1 h pre-digestion (1h, without overnight digestion) contained larger proportions of *Y. pestis* sequences. We also observed inverse trends with clonality, which is reduced in 1-h pre-digested samples (Figure S1). Combined, these results support the presence of more *Y. pestis* DNA molecules in the pre-digested fraction.

Human mtDNA content followed no particular trend, representing lower, similar or greater proportions of the sequence data depending on the specimen considered. This indicated that only some of the remains analyzed still contained a substantial fraction of human DNA in the undigested pellets and that DNA preservation was considerably more limited in others.

Encouraged by the increased proportions of *Y. pestis* DNA present in the pre-digested fraction, we further investigated the kinetics of DNA release during the 60 min (min) pre-digestion wash. To achieve this, we generated fresh dental pulp powder for four additional teeth from individuals that showed positive plague identification in our previous experiments. The powder was then subjected to six consecutive partial digestion steps at 37°C of 10 min each (10 min, 20 min, 30 min, 40 min, 50 min, and 60 min; Figures 3C and 3D). The remaining undigested pellets were collected after each step, and incubated further in the same volume of a fresh extraction buffer. The resulting six pre-digested fractions were purified and used for library preparation, amplification, DNA capture, and sequencing (Figures 3C and 3D). Normalizing for sequencing efforts after *Y. pestis* whole-genome capture, we found that most of the *Y. pestis* DNA was released in the

extraction buffer during the first 10 min pre-digestion. The proportion of *Y. pestis* DNA released during further pre-digestion steps was consistently lower than that released during the previous incubations. This indicated rapid *Y. pestis* DNA release in the extraction buffer, and diminishing returns with increasing pre-digestion times. Human mtDNA release followed a similar trend for one tooth (MAJ46t9), but exhibited no consistent trend across the four teeth investigated. Interestingly, MetaPhlAn 4 taxonomic profiling of microbial diversity for the single of the four teeth for which sufficient sequencing data were generated ($\geq$ 100,000 collapsed reads; AMIENS268t2), revealed increasingly abundant off-target microbial hits with pre-digestion times (Figures 2C and 2D). This translated into an increasing number of classified microbial species, consistent with capture probes not being saturated by *Y. pestis* DNA templates only allowing aspecific annealing, including to oral bacterial DNA (e.g., *Desulfobulbus oralis*; Figure 2D).

We repeated the same experiment for four additional teeth belonging to the same four individuals, except that shorter pre-digestion times were considered (0.5 min, 1 min, 2.5 min, 5 min, 7.5 min, and 10 min; Figures 3E and 3F). The proportions of *Y. pestis* and human mtDNA present in each resulting pre-digested fraction were compared to those found in the absence of pre-digestion (0 min), which were estimated after suspending the dental pulp powder in the pre-heated extraction buffer, and immediately centrifugating the pellets to collect the supernatant.

Two of the teeth examined were found to contain almost no *Y. pestis* DNA considering pre-digestion times up to 10 min (MAJ46t11 and SQ8723t4; Figures 3E and 3F). Other teeth from these individuals, however, were tested positive for *Y. pestis* in previous experimental conditions. This highlights that the success of *Y. pestis* detection strongly depends on the tooth material considered, possibly due to differences in DNA preservation across tooth microenvironments and/or tooth physical integrity conditioning colonization success from the oral and environmental communities.

The other two of the four teeth examined showed inconsistent kinetics of DNA release, with one sample (AMIENS268t4) releasing roughly twice as much *Y. pestis* DNA between 2.5 and 5 min than during the following 2.5 min (between 5 and 7.5 min), but the same proportion again between 7.5 and 10 min. For the last tooth (LAR27t4), the proportion of *Y. pestis* DNA collected in the last experimental condition was 6.43-fold–10.10-fold larger than during the two previous steps, despite lasting for 2.5 min each. It was also equivalent to the proportion of *Y. pestis* DNA obtained following a 1 min pre-digestion (Figure 3E). The AMIENS268t4 tooth provided sufficient numbers of collapsed reads for metagenomic profiling with MetaPhlAn 4, which revealed the presence of other microbes than *Y. pestis* in some but not all pre-digestion conditions (Figures 2C and 2D). Combined, these results indicated stochastic effects dominating DNA release of both *Y. pestis* and other microbes during the first 10 min of pre-digestion. No specific trend could again be observed for the release of human mtDNA.

As the pre-digestion conditions between 10 and 60 min provided consistent results across the four teeth investigated, we decided to further characterize the underlying kinetics of DNA release in the pre-digestion buffer. We found that the proportion of unique stringent alignments characterized decreased exponentially through time (Figures 4A and 4B). This was true for both *Y. pestis* and human mtDNA, which followed trends that were not significantly different (ANCOVA test, p values $\geq$ 0.19). As a result, the cumulative proportion of *Y. pestis* and human mtDNA templates sequenced rapidly reached saturation, leaving only 5% (1%) of unreleased templates after 28.0–38.5 min (43.1–59.2 min) of pre-digestion (Figures 4C, 4D, and S2). This indicated that pre-digesting the dental pulp powder for 30 to 60 min in the experimental conditions tested would be sufficient to release most of the *Y. pestis* DNA present. We, thus, performed the DNA analysis of the remaining teeth following a conservative pre-digestion for 60 min at 37°C to extract the DNA fraction of the pathogen almost entirely. The supernatant was purified and concentrated for preparing a total of 176 DNA libraries for *Y. pestis* and human mtDNA enrichment.

Phylogenetic analyses

The improved DNA extraction procedures developed in this study allowed us to characterize the genome sequence of *Y. pestis* strains present in 12 ancient individuals. Patterns of nucleotide mis-incorporations and base compositional profiles at the genomic positions preceding and following read termini were characteristic of authentic ancient DNA data generated following USER-treatment of extracts (Figure S3). Specifically, USER enzymes cleave those cytosine residues that are deaminated postmortem, leaving only slightly inflated C→T (G→A) mis-incorporation rates at read starts (ends), as well as an excess of ancient

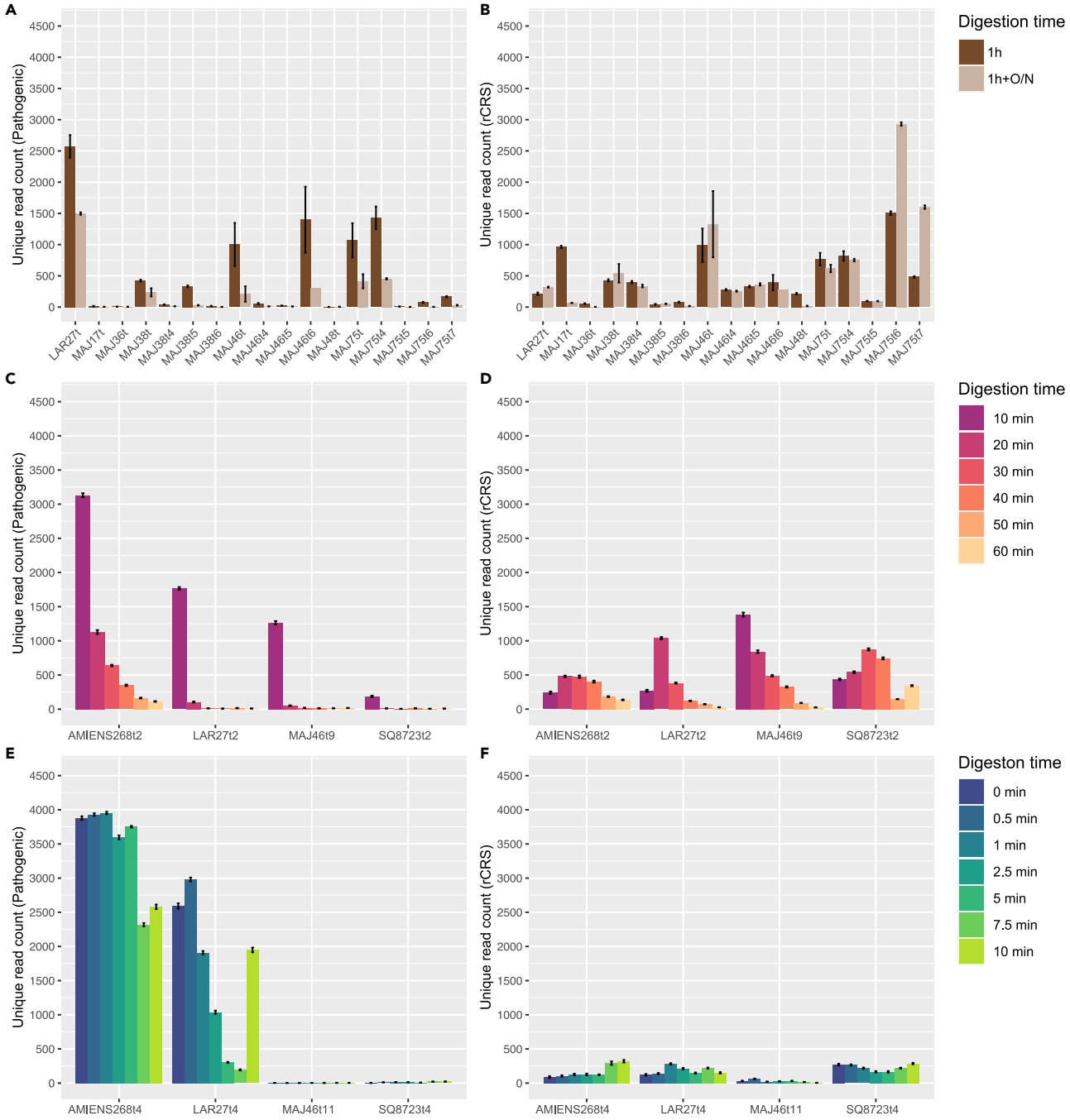


Figure 3. Number of unique high-quality alignments identified following different extraction conditions

(A and B) One hour of pre-digestion wash (1h, dark brown) versus complete pellet digestion overnight (1h+O/N, light brown).

(C and D) 10, 20, 30, 40, 50, and 60 min of pre-digestion wash (from purple to yellow).

(E and F) 0, 0.5, 1, 2.5, 5, 7.5, and 10 min of pre-digestion wash (from blue to green). Analyses are repeated 10 times, following random sampling of 4,557 collapsed reads in each experimental condition for normalizing sequencing efforts. Data are represented as mean ± SEM. Alignments are shown against the *Y. pestis* CO92 genome (left; A, C, E, pathogenic) and the human mtDNA genome (right; B, D, F, rCRS). See also Figure S1.

DNA templates starting immediately after cytosine residues.[2] Read-to-reference edit distance distributions confirmed a closer genetic proximity to *Y. pestis* relative to the closest outgroup, *Yersinia pseudotuberculosis* (Figure S4).

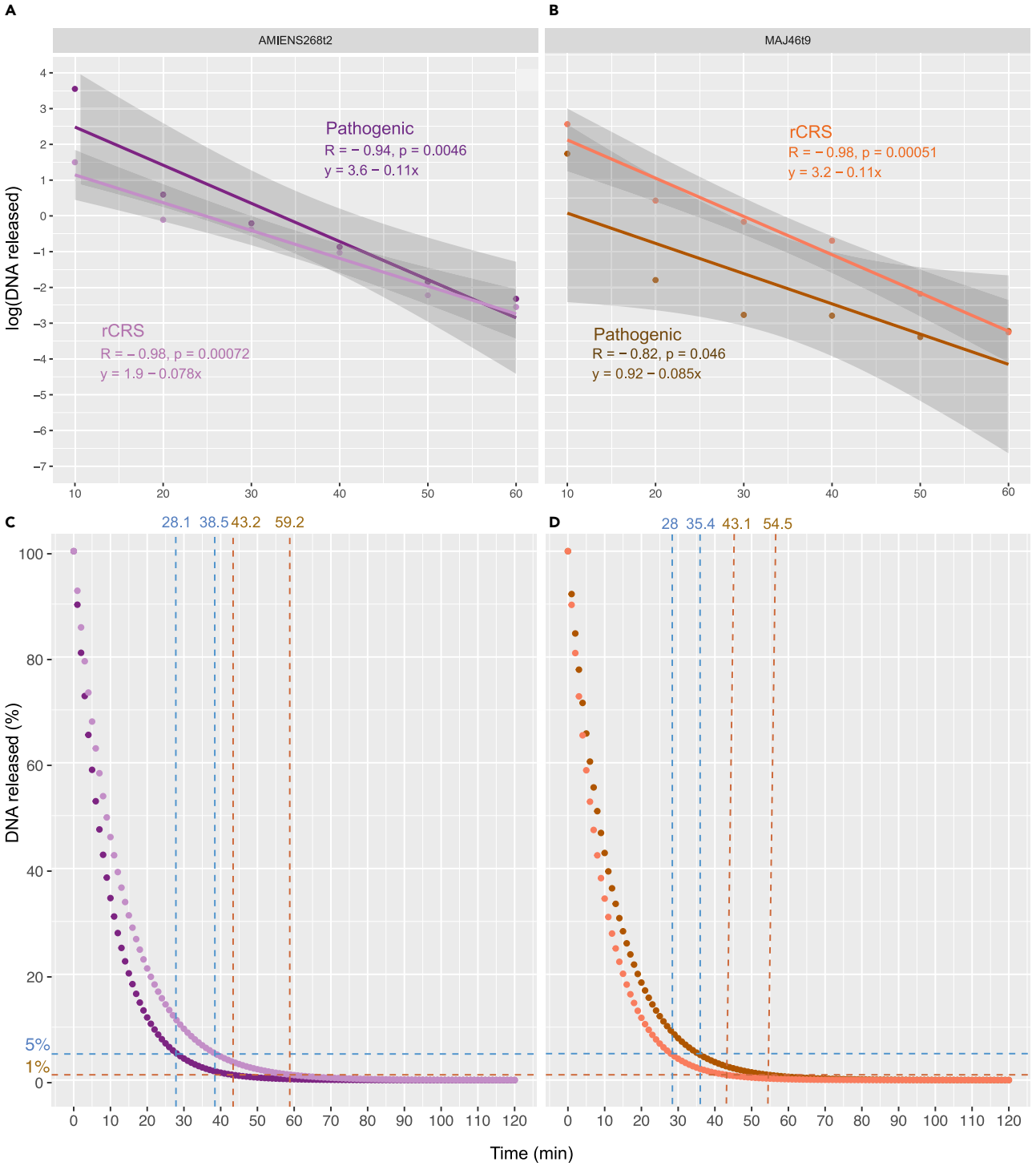


Figure 4. Kinetics of DNA release during pre-digestion wash

(A and B) Percentage of DNA released following pre-digestion wash of incremental time (10 min–1 h). The results are shown for the AMIENS268t2 (A) and MAJ46t9 (B) teeth (see Figure S2 for LAR27t2 and SQ8723t2 results). High-quality sequence alignments are aligned against the pathogen genome (CO92 and plasmids, dark purple, and dark orange) or the human mtDNA genome (Revised Cambridge Reference Sequence rCRS, light purple, and light orange). (C and D) Prediction for the DNA fraction remaining to be released following incremental pre-digestion wash. Predictions are based on the linear models shown in panels A and B for *Y. pestis* (dark purple and dark orange) and human mtDNA (light purple and light orange). The blue and dark orange dashed lines indicate the times when respectively 95% and 99% of the DNA has been released. See also Figure S2.

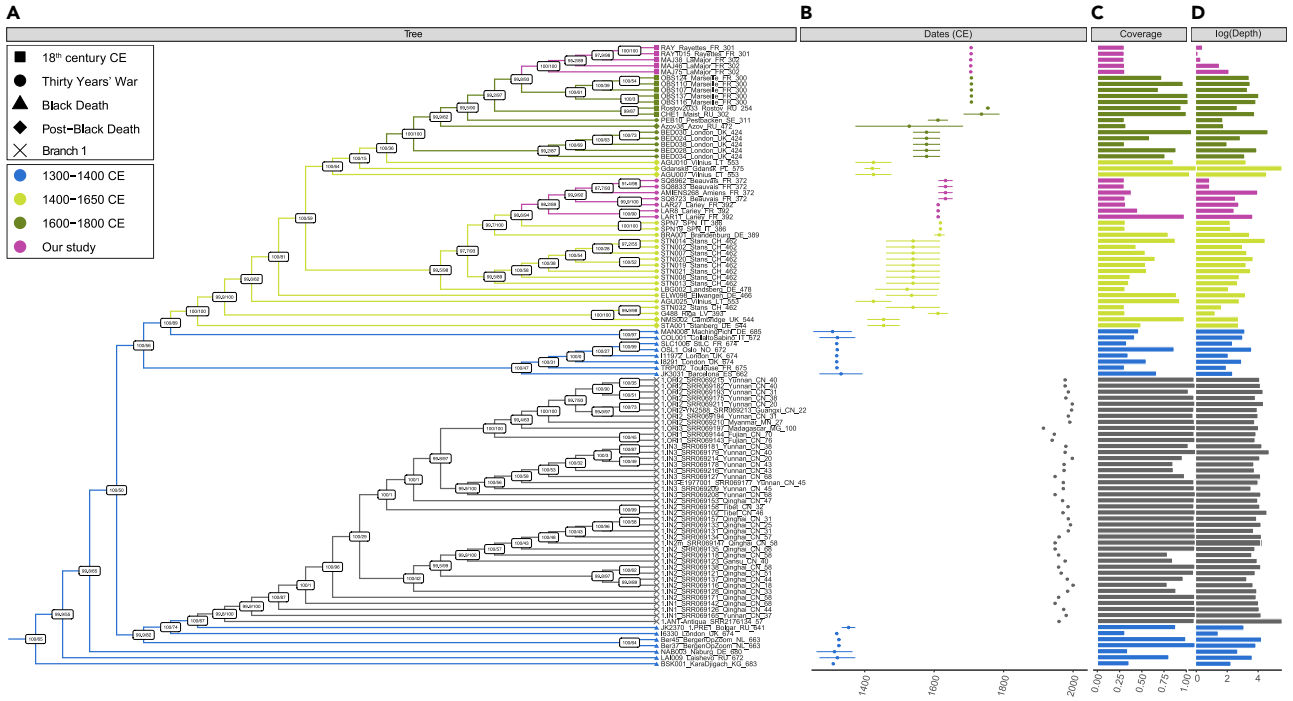


Figure 5. Phylogenetic reconstruction

(A) Maximum Likelihood tree (IQ-TREE), considering 29,609 genome-wide variant positions known to provide reliable phylogenetic signal. This analysis included 185 previously published genomes, but the tree shown is restricted to second and third pandemic genomes, disregarding branch lengths, for clarity (the tree with branch length is shown in Figure S5). Node supports are estimated from SH-aLRT calculations (left) and ultrafast bootstrap approximation (right).

(B) Sample dates CE, with error margins representing the lower and upper boundaries provided by archaeological contexts.

(C and D) represent the coverage and the logarithm of the average sequencing depth-of-coverage achieved for the different CO92 genomes considered. The colors illustrated follow the coloring scheme from Spyrou et al., 2022,[43] and the tip shapes indicate the historical context of archaeological samples investigated with respect to Figure 1. See also Figures S5 and S6.

Phylogenetic reconstructions revealed that most of the second pandemic *Y. pestis* genomes clustered together within a monophyletic group, including strains from the Black Death to the 18th c. CE (Figures 5 and S5, Table S2). Within this cluster, the 12 ancient *Y. pestis* genomes characterized in this study grouped together with all the other post-Black Death ancient *Y. pestis* genomes hitherto sequenced. This was true genome-wide (Figure S6), or when considering only a subset of 29,609 parsimony informative sites located outside non-core, repetitive or highly conserved regions, which are reported to be less prone to environmental contaminant mis-mapping (Figures 5 and S5).[54] The latter analysis confirmed the 1338–1339 CE genome from KaraDjigach, Kyrgystan, as basal to all subsequent strains sequenced, belonging to the second and third pandemic. Phylogenetic reconstructions also supported all modern branch 1 genomes responsible for the third pandemic descending from a subset of 14th c. CE strains identified in the Netherlands (BergenOpZoom), the United Kingdom (London), and Russia (Bolgar), as previously reported.[35,55,56] Additionally, the LAR27 genome from Lariey characterized in this study (1629–1630 CE) appeared closely related to the two genomes previously reported from the same site (LAR8 and LAR11), confirming their genetic proximity with contemporary *Y. pestis* genomes from San Procolo a Naturno (Italy, 1636 CE).[45] The *Y. pestis* genomes from Beauvais and Amiens (1630–1670 CE) also clustered with this group, suggesting common epidemiological origins. This lineage further spread during or in the aftermath of the Thirty Years' War (Figures 5 and S5), which is recognized as one of the most destructive conflicts in European history, it affected Central Europe between 1618 and 1648 CE, a time contemporary with the ancient *Y. pestis* genome from Brandenburg, which is located basally within this group.

The *Y. pestis* genome sequences retrieved from the samples of La Major (Marseille, 1720 CE) and Les Rayettes (Martigues, 1721 CE) formed a monophyletic group together with the previously reported genomes from Observance (Marseille, 1722 CE).[57] This supports historical sources that describe an initial

introduction of the plague in Marseille by the ship *Grand Saint Antoine* on May 1720 CE,[49] and a subsequent rapid spread to the entire Provence region. Interestingly, our phylogenetic reconstructions indicated that the ancient *Y. pestis* genomes from Observance do not directly descend from those entering Marseille two years earlier and present at La Major. Taken at face value, this may have provided support for the introduction of two closely related plague strains. However, we caution that due to extreme DNA degradation, the genomes from La Major and Les Rayettes could only be characterized with limited coverage (between 1.19-fold and 5.25-fold), despite the improved methodology described in this study. Furthermore, the majority of ancient strains presented in our phylogenetic analysis appear to have excessively long terminal branches, even when compared to modern genetic diversity (Figures S5 and S6). The combination of low-genome coverage, remnant sequencing errors due to DNA degradation, and residual contamination from environmental microbes, can result in excessive terminal branch lengths in ancient genomes. The presence of such long branches can introduce significant biases in phylogenetic placement,[58,59] possibly impacting fine-scale evolutionary predictions. Additionally, we note that the Observance genomes were characterized using a different DNA capture system,[57] introducing subtle technical batch effects, also potentially impacting phylogenetic reconstruction. This can be illustrated by the single Lariey genome characterized in this study, following DNA capture, which appears phylogenetic close, but does not form a monophyletic clade with those previously characterized by shotgun sequencing in Seguin-Orlando and colleagues.[45] In light of these caveats, we consider our phylogenetic reconstructions consistent with the abundance of historical sources supporting a unique introduction of the plague in Marseille in 1720 CE.[49] Determining whether the underlying strain first emerged in the Levant, or derived from a European source re-entering Europe through the *Grand Saint-Antoine*, requires further work. For example, mapping the genetic structure of *Y. pestis* strains at a finer-grained resolution during the 17th and the 18th c. CE may clarify the geographic sources and spread of related plague outbreaks. Integrating historical evidence will however be essential, given that the mutation rate of the *Y. pestis* genome provides genetic resolution at the 10-year timescale,[60] which is considerably slower than the pace of human circulation and trade.

Genome analyses

We further tested whether the ancient *Y. pestis* strains sequenced in this study showed specific genetic characteristics compared to their closest phylogenetic relatives. The average depth-of-coverage of the 12 individuals sequenced here ranged between 0.16 and 56.16-fold for the circular chromosome CO92 (median = 3.97-fold, average = 11.67-fold), while it ranged between 0.79 and to221.45-fold, 0.20– to73.46-fold and 5.71– to1092.67-fold for the pCD1, pMT1, and pPCP1 plasmids, respectively. Interestingly, the average proportion of high-quality alignments against the CO92 chromosome was lower-than-expected on the basis of the fraction of CO92 probes used for capture. The average proportions of high-quality alignment against the three different plasmids were, conversely, larger-than-expected on the basis of the probe distribution (Figures 6A–6D). This is in line with bacterial cells containing multiple copies of plasmids. The excess of plasmid DNA relative to probe proportions was, however, significantly more limited in the specimens dated from the 17th c. CE than in those from the 18th c. CE (Figures 6B–6D). This suggests that the latter contained a greater number of plasmid copies per cell than the former. Interestingly, the plasmid copy number can have an impact on virulence,[61] in particular the pCD1 plasmid, which carries the T3SS genes responsible for delivering *Yersinia* outer membrane proteins into the host cells.[62] The number of plasmids carried by the 17th and 18th c. CE strains may, thus, have driven differences in their virulence. Whether this genetic feature played a role in the excessive mortality observed during the Great Plague of Marseille, which decimated half the city's population in only a few months,[49] remains to be tested. Whether it also had equally dramatic consequences on the populations of rodent natural hosts, thus, precipitating disease disappearance is also unknown.

Furthermore, a 49-kb region from the CO92 chromosome was found to be largely absent in the strains from La Major and Les Rayettes (Figures 7A–7E). The *cspE* gene (cspE_1883773–1883985), which is involved in the survival and the multiplication at low temperature,[63] returned spurious non-null coverage, due to sporadic alignments resulting from environmental *Yersinia* bacteria. The 49-kb region spans the *mgtA* (mgtA_1888533–1891232, previously named *mgtB*) and *mgtC* (YPORS09280_1887434–1888132) genes involved in the survival of the bacteria under condition with low concentrations of Mg^{2+}.[64] It was previously reported to also be missing from the genomes of their closest phylogenetic relatives, including 18th c. CE strains from France (Observance, Marseille), Russia (Maist), and Sweden (Pestbacken), as well as mid-16th/mid-17th c. CE strains from the United Kingdom (London).[51] Moreover, a 47-kb region overlapping 11

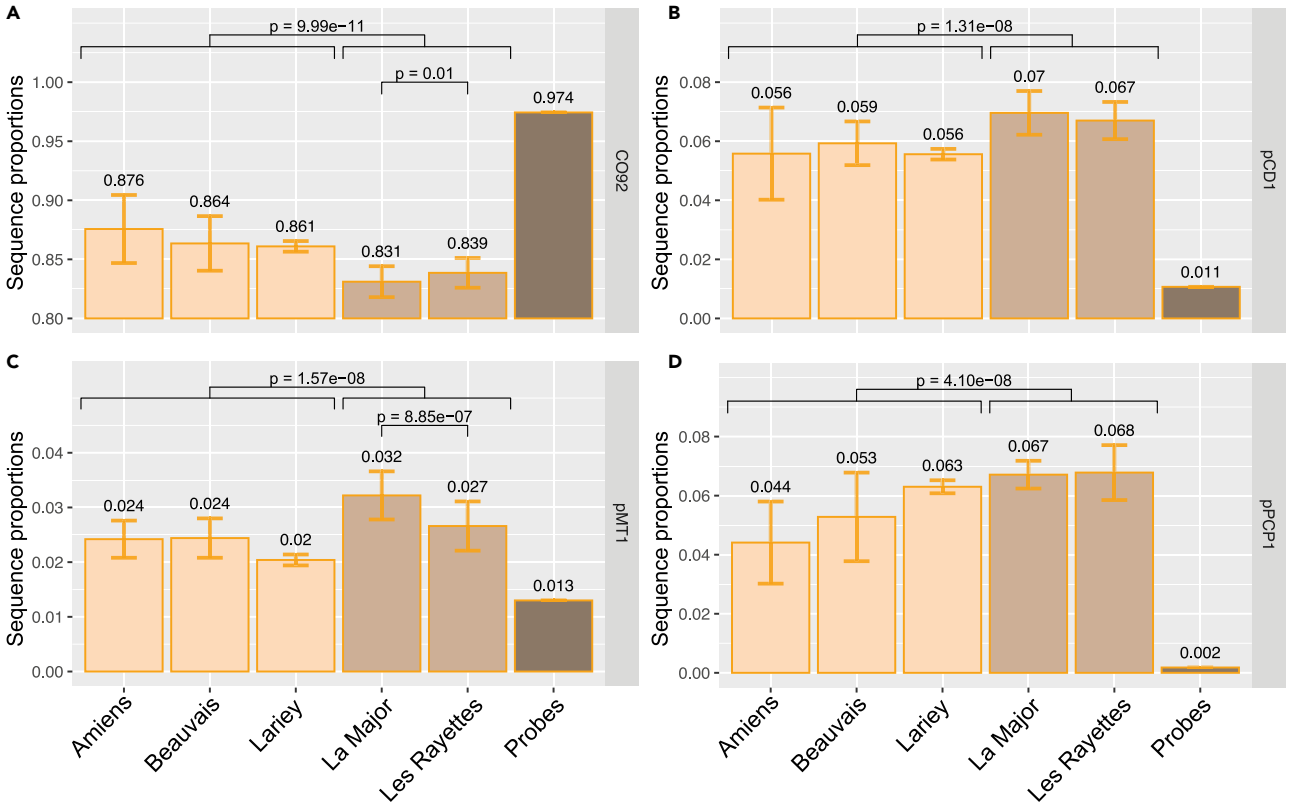


Figure 6. Plasmid copy numbers per bacterial cell

(A–D)Average proportions of sequences mapping against the CO92 circular chromosome (A) and the plasmids (B, C, and D) for the different archaeological sites from the 17th c. CE (orange, Amiens, Beauvais, and Lariey) and the 18th c. CE (light brown, La Major, and Les Rayettes). Probe proportions targeting the CO92 genome and the three plasmids are indicated by the brown bars. Data are represented as mean ± SEM. Statistically significant values (Wilcoxon test) for inter- and intra-century comparisons are indicated. Only those libraries prepared on the DNA extracted following 1-h pre-digestion were considered to avoid technical batch effects, disregarding those for which an insufficient number of high-quality read alignments could be identified (*i.e.* < 500).

virulence and pathogenicity factors (inv_2040296–2042350, fyuA_2140840–2142861, ybtE_2142992–2144569, ybtT_2144573–2145376, ybtU_2145373–2146473, irp1_2146470–2155961, irp2_2156049–2162156, ybtA_2162347–2163306, ybtP_2163563–2165275, ybtQ_2165262–2167064, ybtX_2167057–2168337) was largely under-covered in all the genomes from La Major, Beauvais, Lariey and Amiens reported here (Figure 7A). They were, however, covered in two genomes from Lariey that were previously characterized using shotgun sequencing. We, thus, consider that the limited coverage observed most likely reflects under-performing probe-template annealing in those regions than genuine genomic rearrangements.

We next used snpToolkit[56] to characterize the single nucleotide mutations found among the new 17th and 18th c. CE strains sequenced here, and their closest phylogenetic relatives. This analysis revealed a total of 375 SNPs, comprising 148 non-synonymous, 97 synonymous, and 130 intergenic mutations (Figure S7, Table S3). Interestingly, the iron ABC transporter encoded by the *YbtP* gene, which is essential for the bacterial physiology[65] and can deplete *Y. pestis* avirulence in mice,[66] was found to be polymorphic in both strains from the 17th and 18th c. CE. No mutations were otherwise found in the 235 virulence and pathogenicity loci considered in the analyses above. The genomes from Amiens and Beauvais *Y. pestis* strains (1630–1670 CE) showed remarkable genetic similarity with those from Lariey (1629–1630 CE), as they shared only eight non-synonymous variants not found in the latter (Table S3). One such variant affected the autotransporter *YapC* gene, which mediates adhesion to human cells and, thus, may affect host innate immune response.[67] In contrast, the number of non-synonymous differences found between the 1720–1722 CE genomes from La Major, Les Rayettes, and Observance was larger. Whether this is indicative of currently unidentified minor mis-mapping from other bacteria than *Y. pestis*, or faster mutation rates in this specific lineage requires further clarification through additional sequencing of complete genomes from *Y. pestis* strains responsible for the plague epidemic that spreads to the Provence region in the early 18th c. CE.

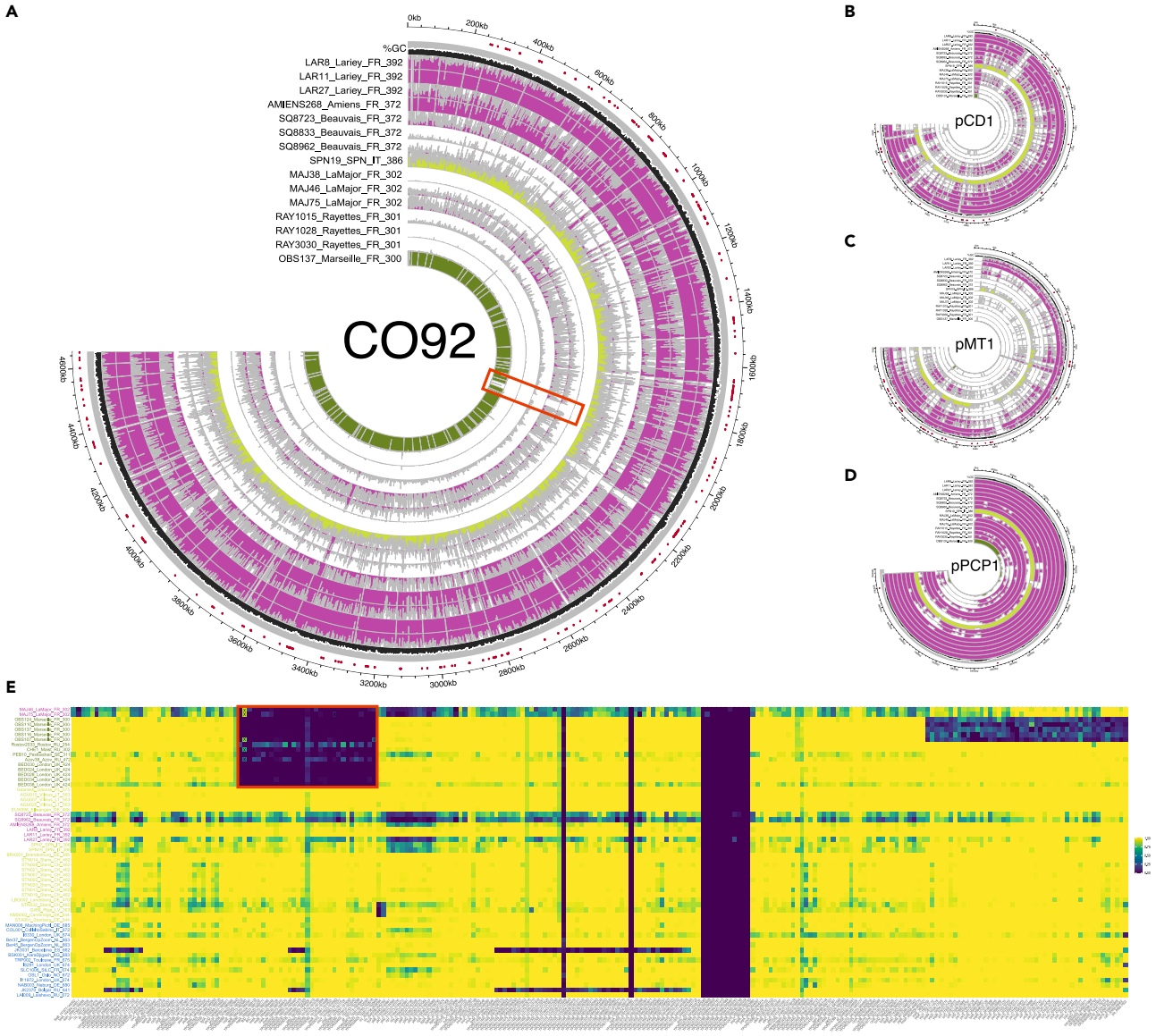


Figure 7. Genome coverage and %GC variation

(A) CO92 *Y. pestis* reference genome.

(B) pCD1 plasmid.

(C) pMT1 plasmid.

(D) pPCP1 plasmid. A total of 13 samples sequenced in this study are shown (purple) together with two close relatives from the 17th c. CE (yellow) and the 18th c. CE (green) for comparison.

(E) Fraction of the gene covered at least once for 235 virulence and pathogenicity loci. Second pandemic *Y. pestis* genomes were considered as long as they were sequenced to a minimum 2-fold average depth-of-coverage. The red box highlights the 49-kb deletion identified, with "X" indicating those virulence and pathogenicity loci showing best Blast hits against environmental *Yersinia* species, and, thus, likely reflecting false positive alignments against *Y. pestis*. See also Figures S3, S4, and S7 and Table S3.

Conclusion

In this study, we report that ancient *Y. pestis* DNA is almost entirely released during a 60 min pre-digestion wash of the dental pulp. This finding might have important methodological implications for the workflow of ancient pathogen studies, as DNA extracts obtained following pre-digestion appear to maximize the ratio of ancient pathogen DNA to human DNA. This contrasts to the extracts obtained following complete pellet digestion, which are generally used for the study of ancient plagues, despite containing a much greater

ratio of human to pathogen DNA. In the experimental conditions described herein, a pre-digestion of the dental pulp powder for 60 min at 37°C appears sufficient to release at least 99% of the ancient *Y. pestis* DNA generally preserved. Applying targeted capture to these DNA extracts allowed us to characterize 12 ancient *Y. pestis* genomes from five French archaeological sites dated to the 17th and 18th c. CE. We find that the number of plasmid copies within *Y. pestis* bacteria from the 17th c. CE was lower than in those from the 18th c. CE. The Thirty Years' War likely facilitated the spread of a plague lineage throughout Germany, Italy, and France, all of which were eventually dead-end emergences. The last plague outbreak in France appears to follow a single emergence in Marseille, before spreading to the Provence region.

Limitations of the study

In this study, genome variation among ancient *Y. pestis* strains was only investigated from sequence alignments against one single reference genome. The full extent of structural variation and their possible consequences on transmission, pathogenicity and virulence remains to be characterized. Moreover, the quality of the data generated for most of the ancient genomes presented was limited, which precluded the examination of fine-scale strain evolution and divergence dates. Only a small number of samples were subjected to pre-digestion tests, they were also confined to a single geographic area and, only encompassed preservation conditions from a relatively recent period. Even if our results highlight a strong pattern for the DNA release of *Y. pestis*, experiments should be extended to various areas, periods and epidemics. Finally, whether the recommended methodology shows similar performance on other types of non-dental tissues, and will equally apply to the entire range of pathogenic bacteria that can circulate in the blood, remains untested.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Samples and archaeological sites
- METHOD DETAILS
  - Ancient DNA extraction
  - Library construction and PCR amplification
  - Capture and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106787.

ACKNOWLEDGMENTS

We thank all the members of the AGES group at CAGT for fruitful discussion and Alice Rabasse for proofreading the manuscript. This research was funded by the CNRS MITI (*Mission pour les Initiatives Transverses et Interdisciplinaires*) IndigenousHealth program; the ANR LifeChange; the ANR GenIn; the Simone and Cino Del Duca Foundation (Subventions scientifiques 2020, HealthTime Travel) and the European Research Council (ERC), under the European Union's Horizon 2020 research and innovation program (grant agreement 681605).

AUTHOR CONTRIBUTIONS

L.O. conceived the project and designed the research. N.T., B.C., R.J., S.T., L.B., J.M.F., M.S., and C.C. provided samples and information about historical and archaeological context. P.C., C.D.S., C.T., C.G., and A.S.O. carried out ancient DNA laboratory work with input from L.C., G.T., S.S., and L.C.T., who managed the sequencing runs. P.C., L.L., C.D.S., and L.O. carried out computational analyses, with input from A.S.O. and L.O. provided reagents and material. J.K. and H.P. provided the probe design. P.C., C.D.S., A.S.O., and L.O. wrote the paper with input from J.K., H.P., and M.A.S., and all the coauthors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 22, 2022
Revised: February 11, 2023
Accepted: April 26, 2023
Published: May 2, 2023

REFERENCES

1. Llamas, B., Willerslev, E., and Orlando, L. (2017). Human evolution: a tale from ancient genomes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20150484. https://doi.org/10.1098/rstb.2015.0484.

2. Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., and Warinner, C. (2021). Ancient DNA analysis. Nat. Rev. Methods Primers *1*, 14–26. https://doi.org/10.1038/s43586-020-00011-0.

3. 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. https://doi.org/10.1371/journal.pone.0129102.

4. Sirak, K., Fernandes, D., Cheronet, O., Harney, E., Mah, M., Mallick, S., Rohland, N., Adamski, N., Broomandkhoshbacht, N., Callan, K., et al. (2020). Human auditory ossicles as an alternative optimal source of ancient DNA. Genome Res. *30*, 427–436. https://doi.org/10.1101/gr.260141.119.

5. Harney, É., Cheronet, O., Fernandes, D.M., Sirak, K., Mah, M., Bernardos, R., Adamski, N., Broomandkhoshbacht, N., Callan, K., Lawson, A.M., et al. (2021). A minimally destructive protocol for DNA extraction from ancient teeth. Genome Res. *31*, 472–483. https://doi.org/10.1101/gr.267534.120.

6. Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., and Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques *59*, 87–93. https://doi.org/10.2144/000114320.

7. Boessenkool, S., Hanghøj, K., Nistelberger, H.M., Der Sarkissian, C., Gondek, A.T., Orlando, L., Barrett, J.H., and Star, B. (2017). Combining bleach and mild predigestion improves ancient DNA recovery from bones. Mol. Ecol. Resour. *17*, 742–751. https://doi.org/10.1111/1755-0998.12623.

8. Ginolhac, A., Vilstrup, J., Stenderup, J., Rasmussen, M., Stiller, M., Shapiro, B., Zazula, G., Froese, D., Steinmann, K.E., Thompson, J.F., et al. (2012). Improving the performance of true single molecule sequencing for ancient DNA. BMC Genom. *13*, 177. https://doi.org/10.1186/1471-2164-13-177.

9. Damgaard, P.B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., and Allentoft, M.E. (2015). Improving access to endogenous DNA in ancient bones and teeth. Sci. Rep. *5*, 11184. https://doi.org/10.1038/srep11184.

10. Essel, E., Korlević, P., and Meyer, M. (2021). A method for the temperature-controlled extraction of DNA from ancient bones. Biotechniques *71*, 382–386. https://doi.org/10.2144/btn-2021-0025.

11. Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503. https://doi.org/10.1038/nature16152.

12. Rohland, N., Mallick, S., Mah, M., Maier, R., Patterson, N., and Reich, D. (2022). Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. Genome Res. *32*, 2068–2078. https://doi.org/10.1101/gr.276728.122.

13. Fu, Q., Meyer, M., Gao, X., Stenzel, U., Burbano, H.A., Kelso, J., and Pääbo, S. (2013). DNA analysis of an early modern human from Tianyuan Cave, China. Proc. Natl. Acad. Sci. USA *110*, 2223–2227. https://doi.org/10.1073/pnas.1221359110.

14. Carpenter, M.L., Buenrostro, J.D., Valdiosera, C., Schroeder, H., Allentoft, M.E., Sikora, M., Rasmussen, M., Gravel, S., Guillén, S., Nekhrizov, G., et al. (2013). Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. Am. J. Hum. Genet. *93*, 852–864. https://doi.org/10.1016/j.ajhg.2013.10.002.

15. Enk, J.M., Devault, A.M., Kuch, M., Murgha, Y.E., Rouillard, J.-M., and Poinar, H.N. (2014). Ancient whole genome enrichment using baits built from modern DNA. Mol. Biol. Evol. *31*, 1292–1294. https://doi.org/10.1093/molbev/msu074.

16. Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., et al. (2010). A draft sequence of the neandertal genome. Science *328*, 710–722. https://doi.org/10.1126/science.1188021.

17. Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., et al. (2010). Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature *463*, 757–762. https://doi.org/10.1038/nature08835.

18. Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., et al. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature *468*, 1053–1060. https://doi.org/10.1038/nature09710.

19. Allentoft, M.E., Sikora, M., Refoyo-Martínez, A., Irving-Pease, E.K., Fischer, A., Barrie, W., Ingason, A., Stenderup, J., Sjögren, K.-G., Pearson, A., et al. (2022). Population genomics of stone Age eurasia. Preprint at bioRxiv 2022. https://doi.org/10.1101/2022.05.04.490594.

20. Lazaridis, I., Alpaslan-Roodenberg, S., Acar, A., Açıkkol, A., Agelarakis, A., Aghikyan, L., Akyüz, U., Andreeva, D., Andrijašević, G., Antonović, D., et al. (2022). The genetic history of the southern arc: a bridge between west asia and Europe. Science *377*, eabm4247. https://doi.org/10.1126/science.abm4247.

21. Spyrou, M.A., Bos, K.I., Herbig, A., and Krause, J. (2019). Ancient pathogen genomics as an emerging tool for infectious disease research. Nat. Rev. Genet. *20*, 323–340. https://doi.org/10.1038/s41576-019-0119-1.

22. Pajer, P., Dresler, J., Kabíckova, H., Písa, L., Aganov, P., Fucik, K., Elleder, D., Hron, T., Kuzelka, V., Velemínsky, P., et al. (2017). Characterization of two historic smallpox specimens from a Czech museum. Viruses *9*, E200. https://doi.org/10.3390/v9080200.

23. Ferrari, G., Neukamm, J., Baalsrud, H.T., Breidenstein, A.M., Ravinet, M., Phillips, C., Rühli, F., Bouwman, A., and Schuenemann, V.J. (2020). Variola virus genome sequenced from an eighteenth-century museum specimen supports the recent origin of smallpox. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190572. https://doi.org/10.1098/rstb.2019.0572.

24. Kahila Bar-Gal, G., Kim, M.J., Klein, A., Shin, D.H., Oh, C.S., Kim, J.W., Kim, T.-H., Kim, S.B., Grant, P.R., Pappo, O., et al. (2012). Tracing hepatitis B virus to the 16th century in a Korean mummy. Hepatology *56*, 1671–1680. https://doi.org/10.1002/hep.25852.

25. Krause-Kyora, B., Susat, J., Key, F.M., Kühnert, D., Bosse, E., Immel, A., Rinne, C., Kornell, S.-C., Yepes, D., Franzenburg, S., et al. (2018). Neolithic and medieval virus

genomes reveal complex evolution of hepatitis B. Elife *7*, e36666. https://doi.org/10.7554/eLife.36666.

26. Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jäger, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., et al. (2013). Genome-wide comparison of medieval and modern Mycobacterium leprae. Science *341*, 179–183. https://doi.org/10.1126/science.1238286.

27. van Dorp, L., Gelabert, P., Rieux, A., de Manuel, M., de-Dios, T., Gopalakrishnan, S., Carøe, C., Sandoval-Velasco, M., Fregel, R., Olalde, I., et al. (2020). Plasmodium vivax malaria viewed through the lens of an eradicated European strain. Mol. Biol. Evol. *37*, 773–785. https://doi.org/10.1093/molbev/msz264.

28. Marciniak, S., and Poinar, H.N. (2019). Ancient pathogens through human history: a paleogenomic perspective. In Paleogenomics: Genome-Scale Analysis of Ancient DNA Population Genomics, C. Lindqvist and O.P. Rajora, eds. (Springer International Publishing), pp. 115–138. https://doi.org/10.1007/13836_2018_52.

29. 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. https://doi.org/10.1002/ece3.3924.

30. Duchêne, S., Ho, S.Y.W., Carmichael, A.G., Holmes, E.C., and Poinar, H. (2020). The recovery, interpretation and use of ancient pathogen genomes. Curr. Biol. *30*, R1215–R1231. https://doi.org/10.1016/j.cub.2020.08.081.

31. Fellows Yates, J.A., Andrades Valtueña, A., Vågene, Å.J., Cribdon, B., Velsko, I.M., Borry, M., Bravo-Lopez, M.J., Fernandez-Guerra, A., Green, E.J., Ramachandran, S.L., et al. (2021). Community-curated and standardised metadata of published ancient metagenomic samples with AncientMetagenomeDir. Sci. Data *8*, 31. https://doi.org/10.1038/s41597-021-00816-y.

32. Malyarchuk, A.B., Andreeva, T.V., Kuznetsova, I.L., Kunizheva, S.S., Protasova, M.S., Uralsky, L.I., Tyazhelova, T.V., Gusev, F.E., Manakhov, A.D., and Rogaev, E.I. (2022). Genomics of ancient pathogens: first advances and prospects. Biochemistry *87*, 242–258. https://doi.org/10.1134/S0006297922030051.

33. 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. https://doi.org/10.1016/S1473-3099(13)70323-2.

34. 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. USA *116*, 12363–12372. https://doi.org/10.1073/pnas.1820447116.

35. 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. https://doi.org/10.1038/nature10549.

36. Rasmussen, S., Allentoft, M.E., Nielsen, K., Orlando, L., Sikora, M., Sjögren, K.G., Pedersen, A.G., Schubert, M., Van Dam, A., Kapel, C.M.O., et al. (2015). Early divergent strains of Yersinia pestis in eurasia 5,000 Years ago. Cell *163*, 571–582. https://doi.org/10.1016/j.cell.2015.10.009.

37. 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. https://doi.org/10.1016/j.cub.2017.10.025.

38. 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. https://doi.org/10.1038/s41467-018-04550-9.

39. 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. https://doi.org/10.1016/j.cell.2018.11.005.

40. Yu, H., Spyrou, M.A., Karapetian, M., Shnaider, S., Radzevičiūtė, R., Nägele, K., Neumann, G.U., Penske, S., Zech, J., Lucas, M., et al. (2020). Paleolithic to Bronze Age siberians reveal connections with first Americans and across eurasia. Cell *181*, 1232–1245.e20. https://doi.org/10.1016/j.cell.2020.04.037.

41. Susat, J., Lübke, H., Immel, A., Brinker, U., Macăne, A., Meadows, J., Steer, B., Tholey, A., Zagorska, I., Gerhards, G., et al. (2021). A 5,000-year-old hunter-gatherer already plagued by Yersinia pestis. Cell Rep. *35*, 109278. https://doi.org/10.1016/j.celrep.2021.109278.

42. 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*. e000444-19. https://doi.org/10.1128/CMR.00044-19.

43. Spyrou, M.A., Musralina, L., Gnecchi Ruscone, G.A., Kocher, A., Borbone, P.-G., Khartanovich, V.I., Buzhilova, A., Djansugurova, L., Bos, K.I., Kühnert, D., et al. (2022). The source of the Black Death in fourteenth-century central Eurasia. Nature *606*, 718–724. https://doi.org/10.1038/s41586-022-04800-3.

44. 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. USA *110*, 577–582. https://doi.org/10.1073/pnas.1205750110.

45. Seguin-Orlando, A., Costedoat, C., Der Sarkissian, C., Tzortzis, S., Kamel, C., Telmon, N., Dalén, L., Thèves, C., Signoli, M., and Orlando, L. (2021). No particular genomic features underpin the dramatic economic consequences of 17th century plague epidemics in Italy. iScience *24*, 102383. https://doi.org/10.1016/j.isci.2021.102383.

46. Cruz-Dávalos, D.I., Llamas, B., Gaunitz, C., Fages, A., Gamba, C., Soubrier, J., Librado, P., Seguin-Orlando, A., Pruvost, M., Alfarhan, A.H., et al. (2017). Experimental conditions improving in-solution target enrichment for ancient DNA. Mol Ecol Res *17*, 508–522. https://doi.org/10.1111/1755-0998.12595.

47. Ziesemer, K.A., Ramos-Madrigal, J., Mann, A.E., Brandt, B.W., Sankaranarayanan, K., Ozga, A.T., Hoogland, M., Hofman, C.A., Salazar-García, D.C., Frohlich, B., et al. (2019). The efficacy of whole human genome capture on ancient dental calculus and dentin. Am. J. Phys. Anthropol. *168*, 496–509. https://doi.org/10.1002/ajpa.23763.

48. Suchan, T., Kusliy, M.A., Khan, N., Chauvey, L., Tonasso-Calvière, L., Schiavinato, S., Southon, J., Keller, M., Kitagawa, K., Krause, J., et al. (2022). Performance and automation of ancient DNA capture with RNA hyRAD probes. Mol Ecol Res *22*, 891–907. https://doi.org/10.1111/1755-0998.13518.

49. Signoli, M. (2022). History of the plague of 1720-1722, in Marseille. Presse Med. *51*, 104138. https://doi.org/10.1016/j.lpm.2022.104138.

50. Gamba, C., Hanghøj, K., Gaunitz, C., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Bradley, D.G., and Orlando, L. (2016). Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. Mol Ecol Res *16*, 459–469. https://doi.org/10.1111/1755-0998.12470.

51. Spyrou, M.A., 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. (2019). Phylogeography of the second plague pandemic revealed through analysis of historical Yersinia pestis genomes. Nat. Commun. *10*, 4470. https://doi.org/10.1038/s41467-019-12154-0.

52. Blanco-Miguez, A., Beghini, F., Cumbo, F., Mclver, L.J., Thompson, K.N., Zolfo, M., Manghi, P., Dubois, L., Huang, K.D., Maltez Thomas, A., et al. (2022). Extending and improving metagenomic taxonomic profiling with uncharacterized species with MetaPhlAn 4. Preprint at bioRxiv. https://doi.org/10.1101/2022.08.22.504593.

53. Der Sarkissian, C., Ermini, L., Jónsson, H., Alekseev, A.N., Crubezy, E., Shapiro, B., and Orlando, L. (2014). Shotgun microbial

profiling of fossil remains. Mol. Ecol. *23*, 1780–1798. https://doi.org/10.1111/mec.12690.

54. Andrades Valtueña, A., Neumann, G.U., Spyrou, M.A., Musralina, L., Aron, F., Beisenov, A., Belinskiy, A.B., Bos, K.I., Buzhilova, A., Conrad, M., et al. (2022). Stone Age Yersinia pestis genomes shed light on the early evolution, diversity, and ecology of plague. Proc. Natl. Acad. Sci. USA *119*. e2116722119. https://doi.org/10.1073/pnas.2116722119.

55. 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. https://doi.org/10.1016/j.chom.2016.05.012.

56. 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. USA *115*, E11790–E11797. https://doi.org/10.1073/pnas.1812865115.

57. 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. https://doi.org/10.7554/eLife.12994.

58. Bos, K.I., Kühnert, D., Herbig, A., Esquivel-Gomez, L.R., Andrades Valtueña, A., Barquera, R., Giffin, K., Kumar Lankapalli, A., Nelson, E.A., Sabin, S., et al. (2019). Paleomicrobiology: diagnosis and evolution of ancient pathogens. Annu. Rev. Microbiol. *73*, 639–666. https://doi.org/10.1146/annurev-micro-090817-062436.

59. Susko, E., and Roger, A.J. (2021). Long branch attraction biases in phylogenetics. Syst. Biol. *70*, 838–843. https://doi.org/10.1093/sysbio/syab001.

60. Eaton, K., Featherstone, L., Duchene, S., Carmichael, A.G., Varlık, N., Golding, G.B., Holmes, E.C., and Poinar, H.N. (2023). Plagued by a cryptic clock: insight and issues from the global phylogeny of Yersinia pestis. Commun. Biol. *6*, 23. https://doi.org/10.1038/s42003-022-04394-6.

61. Wang, H., Avican, K., Fahlgren, A., Erttmann, S.F., Nuss, A.M., Dersch, P., Fallman, M., Edgren, T., and Wolf-Watz, H. (2016). Increased plasmid copy number is essential for Yersinia T3SS function and virulence. Science *353*, 492–495. https://doi.org/10.1126/science.aaf7501.

62. Demeure, C.E., Dussurget, O., Mas Fiol, G., Le Guern, A.-S., Savin, C., and Pizarro-Cerdá, J. (2019). Yersinia pestis and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. Gene Immun. *20*, 357–370. https://doi.org/10.1038/s41435-019-0065-0.

63. Yu, T., Keto-Timonen, R., Jiang, X., Virtanen, J.-P., and Korkeala, H. (2019). Insights into the phylogeny and evolution of cold shock proteins: from enteropathogenic yersinia and escherichia coli to eubacteria. Int. J. Mol. Sci. *20*, 4059. https://doi.org/10.3390/ijms20164059.

64. Groisman, E.A., Hollands, K., Kriner, M.A., Lee, E.-J., Park, S.-Y., and Pontes, M.H. (2013). Bacterial Mg2+ homeostasis, transport, and virulence. Annu. Rev. Genet. *47*, 625–646. https://doi.org/10.1146/annurev-genet-051313-051025.

65. Delepelaire, P. (2019). Bacterial ABC transporters of iron containing compounds. Res. Microbiol. *170*, 345–357. https://doi.org/10.1016/j.resmic.2019.10.008.

66. Fetherston, J.D., Bertolino, V.J., and Perry, R.D. (1999). YbtP and YbtQ: two ABC transporters required for iron uptake in Yersinia pestis. Mol. Microbiol. *32*, 289–299. https://doi.org/10.1046/j.1365-2958.1999.01348.x.

67. Felek, S., Lawrenz, M.B., and Krukonis, E.S. (2008). The Yersinia pestis autotransporter YapC mediates host cell binding, autoaggregation and biofilm formation. Microbiology *154*, 1802–1812. https://doi.org/10.1099/mic.0.2007/010918-0.

68. 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. https://doi.org/10.1093/bioinformatics/btt193.

69. Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernández, 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. https://doi.org/10.1038/nprot.2014.063.

70. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res. Notes *9*, 88. https://doi.org/10.1186/s13104-016-1900-2.

71. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

72. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274. https://doi.org/10.1093/molbev/msu300.

73. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

74. Dolle, D., Fages, A., Mata, X., Schiavinato, S., Tonasso-Calvière, L., Chauvey, L., Wagner, S., Der Sarkissian, C., Fromentier, A., Seguin-Orlando, A., and Orlando, L. (2020). CASCADE: a custom-made archiving system for the conservation of ancient DNA experimental data. Front. Ecol. Evol. *8*. https://doi.org/10.3389/fevo.2020.00185.

75. 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. USA *117*, 28328–28335. https://doi.org/10.1073/pnas.2009677117.

76. Tzortzis, S., and Signoli, M. (2016). Characterization of the funeral groups associated with plague epidemics. Microbiol. Spectr. *4*. https://doi.org/10.1128/microbiolspec.PoH-0011-2015.

77. 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.

78. Gaunitz, C., Fages, A., Hanghøj, K., Albrechtsen, A., Khan, N., Schubert, M., Seguin-Orlando, A., Owens, I.J., Felkel, S., Bignon-Lau, O., et al. (2018). Ancient genomes revisit the ancestry of domestic and Przewalski's horses. Science *360*, 111–114. https://doi.org/10.1126/science.aao3297.

79. 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. https://doi.org/10.1098/rstb.2013.0624.

80. 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. https://doi.org/10.1016/j.cell.2019.03.049.

81. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*. pdb.prot5448. https://doi.org/10.1101/pdb.prot5448.

82. Poullet, M., and Orlando, L. (2020). Assessing DNA sequence alignment methods for characterizing ancient genomes and methylomes. Front. Ecol. Evol. *8*.

83. 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. https://doi.org/10.1038/35097083.

84. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R.

Bioinformatics *30*, 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.

85. Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. J. Comput. Biol. *7*, 203–214. https://doi.org/10.1089/10665270050081478.

86. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). Genome project data processing subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

87. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

88. 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. https://doi.org/10.1093/molbev/msx281.

89. 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. https://doi.org/10.1093/sysbio/syq010.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Osteological remain	This study	MAJ6t
Osteological remain	This study	MAJ17t
Osteological remain	This study	MAJ19t
Osteological remain	This study	MAJ33t
Osteological remain	This study	MAJ36t
Osteological remain	This study	MAJ38t
Osteological remain	This study	MAJ38t2
Osteological remain	This study	MAJ38t3
Osteological remain	This study	MAJ38t4
Osteological remain	This study	MAJ38t5
Osteological remain	This study	MAJ38t6
Osteological remain	This study	MAJ43t
Osteological remain	This study	MAJ44t
Osteological remain	This study	MAJ46t
Osteological remain	This study	MAJ46t2
Osteological remain	This study	MAJ46t3
Osteological remain	This study	MAJ46t4
Osteological remain	This study	MAJ46t5
Osteological remain	This study	MAJ46t6
Osteological remain	This study	MAJ46t7
Osteological remain	This study	MAJ46t8
Osteological remain	This study	MAJ46t9
Osteological remain	This study	MAJ46t11
Osteological remain	This study	MAJ48t
Osteological remain	This study	MAJ55t2
Osteological remain	This study	MAJ70t
Osteological remain	This study	MAJ72t
Osteological remain	This study	MAJ75t
Osteological remain	This study	MAJ75t3
Osteological remain	This study	MAJ75t4
Osteological remain	This study	MAJ75t5
Osteological remain	This study	MAJ75t6
Osteological remain	This study	MAJ75t7
Osteological remain	This study	MAJ75t8
Osteological remain	This study	MAJ75t9
Osteological remain	This study	MAJ86t
Osteological remain	This study	MAJ87t
Osteological remain	This study	MAJ89t
Osteological remain	This study	MAJ104t
Osteological remain	This study	MAJ1001t

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Osteological remain	This study	RAY1002t
Osteological remain	This study	RAY1005t
Osteological remain	This study	RAY1009t
Osteological remain	This study	RAY1015t
Osteological remain	This study	RAY1015t2
Osteological remain	This study	RAY1015t3
Osteological remain	This study	RAY1028t
Osteological remain	This study	RAY1028t2
Osteological remain	This study	RAY3009t
Osteological remain	This study	RAY3010t
Osteological remain	This study	RAY3015t
Osteological remain	This study	RAY3023t
Osteological remain	This study	RAY3024t
Osteological remain	This study	RAY3027t
Osteological remain	This study	RAY3030t
Osteological remain	This study	RAY3030t2
Osteological remain	This study	LAR8t
Osteological remain	This study	LAR9t
Osteological remain	This study	LAR11t
Osteological remain	This study	LAR14t
Osteological remain	This study	LAR23t
Osteological remain	This study	LAR24t
Osteological remain	This study	LAR26t
Osteological remain	This study	LAR27t
Osteological remain	This study	LAR27t2
Osteological remain	This study	LAR27t4
Osteological remain	This study	LAR30t
Osteological remain	This study	LAR31t
Osteological remain	This study	SQ7044t
Osteological remain	This study	SQ8504t
Osteological remain	This study	SQ8579t
Osteological remain	This study	SQ8833t
Osteological remain	This study	SQ8844t
Osteological remain	This study	SQ8950t
Osteological remain	This study	SQ8519t
Osteological remain	This study	SQ8637t
Osteological remain	This study	SQ8670t
Osteological remain	This study	SQ8772t
Osteological remain	This study	SQ8893t
Osteological remain	This study	SQ8918t
Osteological remain	This study	SQ8962t
Osteological remain	This study	SQ8723t
Osteological remain	This study	SQ8723t2
Osteological remain	This study	SQ8723t4
Osteological remain	This study	AMIENS109t

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Osteological remain	This study	AMIENS121t
Osteological remain	This study	AMIENS122t
Osteological remain	This study	AMIENS133t
Osteological remain	This study	AMIENS134t
Osteological remain	This study	AMIENS146t
Osteological remain	This study	AMIENS148t
Osteological remain	This study	AMIENS177t
Osteological remain	This study	AMIENS183t
Osteological remain	This study	AMIENS209t
Osteological remain	This study	AMIENS222t
Osteological remain	This study	AMIENS255t
Osteological remain	This study	AMIENS260t
Osteological remain	This study	AMIENS266t
Osteological remain	This study	AMIENS268t
Osteological remain	This study	AMIENS268t2
Osteological remain	This study	AMIENS268t4
Osteological remain	This study	AMIENS32t
Osteological remain	This study	AMIENS41t
Osteological remain	This study	AMIENS46t
Osteological remain	This study	AMIENS54t
Osteological remain	This study	AMIENS55t
Osteological remain	This study	AMIENS63t
Osteological remain	This study	AMIENS69t
Osteological remain	This study	AMIENS74t
Osteological remain	This study	AMIENS75t
Osteological remain	This study	AMIENS78t
Osteological remain	This study	AMIENS81t
Osteological remain	This study	AMIENS82t
Osteological remain	This study	AMIENS85t
Osteological remain	This study	AMIENS86t
Osteological remain	This study	AMIENS91t
Osteological remain	This study	AMIENS93t
Osteological remain	This study	AMIENS94t
Osteological remain	This study	AMIENS97t
Osteological remain	This study	AMIENS99t
Deposited data		
Sequencing dataset	This study	ENA: PRJEB59233
Chemicals, peptides, and recombinant proteins		
N-Lauroylsarcosine solution 30% 500ml	Dutscher	Cat#348533
Proteinase K 10MG	Thermo Fisher Scientific	Cat#AM2542
H2O, Molecular Biology Grade, Fisher BioReagents	Thermo Fisher Scientific	Cat#10977015
Tween 20 100ML	Thermo Fisher Scientific	Cat#13464259
Ethanol, Absolute, Mol Biology Grade	Thermo Fisher Scientific	Cat#16606002
USER Enzyme	New England Biolabs	Cat#M5505L
NEBNext End Repair Module	New England Biolabs	Cat#E6050L

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bst DNA Polymerase	New England Biolabs	Cat#M0275L
NEBNext Quick Ligation Module	New England Biolabs	Cat#E6056L
BSA Molecular Biology Grade	New England Biolabs	Cat#B9000S
ACCUPRIME PFX DNA POLYMERASE 100mL	Thermo Fisher Scientific	Cat#10472482
Agencourt AMPure XP - 60ml	Beckman Coulter	Cat#A63881
Buffer PE	QIAGEN	Cat#19065
Buffer PB	QIAGEN	Cat#19066
Buffer EB	QIAGEN	Cat#19086
dNTP Set 100mM 100mL	Thermo Fisher Scientific	Cat#10297018
EDTA 0.5M pH 8.0 Fisher Bioreagents 500ML	Thermo Fisher Scientific	Cat#AM9261
Critical commercial assays		
MinElute PCR Purification kit	QIAGEN	Cat#28006
Tapestation screenTape D1000 HS	Agilent	Cat#5067-5584
Qubit dsDNA HS Assay Kit	Thermo Fisher Scientific	Cat#Q32854
MiniSeq High Output Reagent Kit (150-cycles)	Illumina	Cat#FC-420-1002
myBaits Custom DNA-Seq	Arbor Biosciences	Cat#300596.v5
Deposited data		
Raw and analyzed data	This study	ENA: PRJEB59233
Software and algorithms		
mapDamage2	Jónsson et al.,[68]	https://ginolhac.github.io/mapDamage
PALEOMIX	Schubert et al.[69]	https://github.com/MikkelSchubert/paleomix
AdapterRemoval	Schubert et al.[70]	https://github.com/MikkelSchubert/adapterremoval
Samtools	Li and Durbin,[71]	https://github.com/samtools/samtools
snpToolkit v2.3.0	Namouchi et al.[56]	https://github.com/Amine-Namouchi/snpToolkit
IQ-Tree v1.6.12	Nguyen et al.[72]	https://github.com/iqtree/iqtree1
Bowtie2	Langmead and Salzberg,[73]	https://github.com/BenLangmead/bowtie2
BWA	Li and Durbin,[71]	https://github.com/lh3/bwa
seqtk	-	https://github.com/lh3/seqtk
MetaPhlAn 4	Blanco-Miguez et al.[52]	https://github.com/biobakery/MetaPhlAn

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and materials may be directed to lead contact, Prof. Ludovic Orlando (ludovic.orlando@univ-tlse3.fr).

Materials availability

This study did not generate new unique reagents or materials.

Data and code availability

- Sequencing data can be found on the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home) and are publicly available as of the date of publication under project accession number ENA: PRJEB59233.
- This paper does not report original code.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Samples and archaeological sites

The 120 teeth analyzed in this study originate from 89 individual skeletons buried in five archaeological sites (Figure 1, Table S1; see Data S1 for further description of the archaeological contexts). Prior to destructive sub-sampling, precise virtual 3D models were constructed for the majority of teeth investigated (94/120 teeth), using the ArtecMicro© surface scanner (Artec 3D, 1 mm voxel size). Samples and laboratory procedures were registered in the Laboratory Information Management System CASCADE.[74]

The plague cemetery of Lariey Puy-Saint-Pierre dating to 1629-1630 CE is located in the French Alps, and was excavated in 2002 CE (Minimal number of individuals, MNI = 34). Two out of 10 individuals (LAR8 and LAR11) investigated in this study were previously sequenced by Seguin-Orlando and colleagues in 2021,[45] who demonstrated that they were infected with a *Y. pestis* strain genetically close to those infecting 1636 CE individuals from San Procolo a Naturno, Italy (1636 CE),[75] in the context of the Thirty Year's War.

The excavations at the Hôtel-Dieu Saint-Jean-Baptiste of Amiens, France, were carried out in 2017 CE by the *Service archéologie préventive d'Amiens Métropole*, under the supervision of Richard Jonvel, who provided access to the city's ancient plague cemetery (MNI = 259). The 34 individuals studied here come from multiple graves dating to 1630-1670 CE, when Amiens experienced notable military and epidemic turmoil.

The site of the Maladrerie Saint-Lazare (Beauvais, France) was excavated between 2002 and 2014 CE by the *Service archéologique municipal de Beauvais* (MNI = 490). The discovery of a plague cemetery established since 1623 CE allowed the analysis of 14 individuals coming from the mass grave labelled 8559.

The parish cemetery of La Major cathedral (La Major), excavated in 2008 CE[76] was created at the beginning of September 1720 CE during the last great wave of plague in Marseille (1720-1722 CE). A total of 19 individuals were DNA tested, but showed poor macroscopic preservation, consistent with that observed across the many other individuals buried at the same site (MNI = 106).

The site of Les Rayettes (Capucins de Ferrières trenches, Martigues, France, MNI = 208), was excavated in 2002 CE. It is located less than 40 kilometers away from La Major. The 12 individuals investigated here were found in multiple mass graves dated to 1720-1721 CE and characteristic of the epidemic ongoing locally at the time.

METHOD DETAILS

Ancient DNA extraction

All molecular work, from dental pulp sampling to library construction and PCR mix preparation, was carried out in the ancient DNA facilities of the CAGT (CNRS UMR 5288/Université Paul Sabatier), following strict experimental standards to avoid and monitor contamination, including use of disposable personal protection equipment, positive air pressure, bleach/UV surface and instrument decontamination, and negative control blanks. PCR amplification, purification, quantitation as well as DNA capture and sequencing were carried out in post-PCR laboratories that are located in a building physically isolated from the ancient DNA facilities.

Between 15 and 80 mg of dental pulp were collected for each of the 120 teeth analyzed (incisors, canines, premolars and molars), following the method developed by Neumann and colleagues,[77] which is routinely used at CAGT.[45] The extraction was performed following the protocol described by Seguin-Orlando and colleagues,[45] which corresponds to a modified version of the Y2 protocol described by Gamba and colleagues.[50] The standard protocol makes use of 963 μL of lysis buffer (0.45M EDTA, 0.25 mg/mL proteinase K and 0.5% sodium lauroyl sarcosinate) to carry out a first pre-digestion step of 1 hour at 37°C. The supernatant is then collected following centrifugation at 12,000 rotation per minute (rpm) for 30 seconds and a 963 μL volume of a fresh, pre-heated, lysis buffer is added to proceed to full digestion of the remaining pellets at 42°C overnight. After centrifugation at 12,000 rpm for 30 seconds, 200 μL of the lysis buffer is purified through QIAgen© Minelute columns and concentrated in a final elution buffer consisting of 23 μL EB

supplemented with 0.05 % Tween 20 (later referred to as EB-Tween buffer). A total of 22.8 μL purified extract was treated with 7 μL of USER® enzyme (NEB®) for 3 hours at 37°C, following,[78] to reduce spurious nucleotide mis-incorporation at those deaminated cytosine residues formed by post-mortem degradation. The pre-digestion supernatant was also purified and processed in parallel to the overnight digestion fraction, to contrast the DNA content following shotgun DNA sequencing and/or DNA capture.

The standard extraction protocol was modified to assess the kinetics of DNA release in the lysis buffer. This was achieved through two main experiments, each focusing on two different teeth from four individuals that were previously tested positive for *Y. pestis* DNA (AMIENS268, tooth 2 (t2) and t4; LAR27 t2 and t4; SQ8723 t2 and t4, and; MAJ46 t9 and t11). The first experiment followed the standard extraction protocol, with the exception that the whole supernatant was collected every 10 min and replaced by 963 μL of a fresh lysis buffer. Supernatant collection was carried every 10 min until 1 hour of pre-digestion was completed. The second experiment followed the same procedure as the first, except that the supernatants were collected directly after suspending the dental pulp powder in the lysis buffer (0 min), and 0.5, 1, 2.5, 5, 7.5 and 10 min after. Each collected supernatant was further processed following the methodology described above to obtain 29.8 μL of USER-treated DNA extract.

Library construction and PCR amplification

Triple-indexed double-stranded DNA libraries compatible with Illumina sequencing were prepared following the protocol originally described by Rohland et al., 2015,[79] as modified by Fages and colleagues.[80] This protocol includes a unique combination of two barcodes, each located 3′ of the P5 and P7 library adapters. Due to their location immediately downstream of the sequencing primers, these "internal" barcodes form the first seven nucleotide positions within each sequencing read. Each library was amplified in eight parallel PCRs for eight cycles in the same conditions as described by Seguin-Orlando and colleagues[45]: in a total reaction volume of 25 μL, using 0.4 μL of AccuPrime™ Pfx DNA polymerase (one unit), 2.5 μL 10X Accuprime™ Pfx reaction mix, 2 μL of DNA library, 1 μL of BSA (20 mg/mL), 17.9 μL of H2O, 0.2 μL of inPE1 primer (25 μM) and 1 μL indexing primer (5 μM) containing a unique 6-bp "external" barcode. The latter corresponds to a third barcode located within the PCR primer.[81] The same external index was used in each of the eight parallel amplifications carried out on the same original library. Libraries were then amplified through a second additional PCR round to obtain sufficient DNA mass as input for capture enrichment with a custom myBaits® hybridization capture kit (Daicel Arbor Biosciences). To achieve this, two amplification products for a given library were merged and purified together using Agencourt Ampure XP beads (1:1.2 or 1:1.4 as DNA:beads ratio), and eluted in 11 μL of EB-Tween buffer. This provided four purified amplification products per library, which were further split and processed through two parallel PCR amplifications, using 1 μL of the eluate for each PCR reaction, for 10 cycles to obtain sufficient DNA material for DNA capture. These amplifications formed the second PCR round and were carried out in 25 μL reaction volumes, following the same conditions as for the first amplification round, except that IS5 and IS6 were used as PCR primers.[82] The resulting eight PCR products were co-purified on a single MinElute column (QIAGEN©), and eluted in 12 μL of EB-Tween buffer. Library concentration and/or size distribution were measured on a Tapestation 4200 instrument using a High sensitivity D1000 ScreenTapeAssay (Agilent technologies) as well as on a Qubit HS dsDNA assay (Invitrogen).

Capture and sequencing

A total of four to eight libraries from different extracts with unique internal and external barcodes were pooled and concentrated together on a MinElute column (QIAGEN©). This resulted in pools of 7 μL, containing between 2-12 μg of DNA, as recommended in the myBaits® High Sensitivity Protocol Version 5.00. DNA capture was performed according to the aforementioned protocol, consisting of two rounds of 16 to 24 hour hybridization with the probes described by Wagner and colleagues.[33] Captured libraries were sequenced using the Paired-End mode on an Illumina MiniSeq instrument (2x81 cycles) at CAGT. A similar procedure was followed for shotgun sequencing, except that only one amplification round of 8-12 PCR cycles was performed on each library, before purification, quantification, pooling and DNA sequencing.

QUANTIFICATION AND STATISTICAL ANALYSIS

Read demultiplexing and collapsing, as well as adapter and poor-quality end trimming were performed using AdapterRemoval2 version 2.3.0,[70] tolerating at most one mismatch in each internal barcode (–barcode-mm-r[12] 1 –minadapteroverlap 3 –mm 5). Read mapping against the human reference genome

(hg19, GRCh37) and the revised Cambridge Reference Sequence (rCRS, Accession Number = NC_012920.1) were carried out using Bowtie2 version 2.3.4.1,[73] and the recommended parameters from Poullet and Orlando, 2020.[82] Read mapping against the *Y. pseutotuberculosis* reference genome (Accession Number NC_006155.1), the *Y. pestis* reference genome (strain CO92, Accession Number = NC_003143.1)[83] and the three plasmids pCD1 (Accession Number NC_003131.1), pMT1 (Accession Number NC_003134.1) and pPCP1 (Accession Number AL109969.1) were performed using BWA version 0.7.17 "backtrack" mode,[71] and the "stringent" parameters defined by Spyrou and colleagues.[51] Read (re-)alignment, PCR duplicate removal, quality filtering and mapDamage profiling[68] were carried out using the automated PALEOMIX pipeline v1.2.13.2.[69] The same procedures were followed using previously published read sequence data (Table S2) to obtain the comparative panel of *Y. pestis* genomes and plasmids used for phylogenetic reconstructions and/or for assessing the presence/absence of specific virulence genes and lineage-specific mutations, which was summarized using snpToolkit v2.3[56] (see below).

Basic alignment statistics for each sample, DNA library and experimental conditions are provided in Table S1. When comparing the respective performance of each experimental condition, *Y. pestis* and human DNA content of each library were estimated for normalized sequencing efforts. This was achieved by considering the fraction of collapsed reads only, and down-sampling the sequencing data 10 times for each library to the number collected for the least sequenced library.

The investigation of pathogen and mitochondrial unique reads released through time was performed on the first set of extraction tests on four teeth, which were digested in lysis buffer during one hour, sampling and replacing the supernatant with fresh lysis buffer every 10 min. Results were normalized by down-sampling the number of reads for each digestion time by the lowest number of reads sequenced across the different digestion times for each sample. As the percentage of unique reads mapping against the pathogen genomes (CO92, pCD1, pMT1 and pPCP1) over digestion time approximatively fit an inverse exponential function, data have been log-linearized. Linear models of the logarithm of the percentage of unique reads mapping against pathogen and mitochondrial DNA (mtDNA) were performed using the R *lm* command. The remaining *Y. pestis* and human mtDNA to be released at a given time could be predicted using *predict* command in R, according to these models.

The proportions of unique reads mapping against the circular chromosome and the different plasmids of the pathogen were calculated for each site to compare the plasmid load between 17th and 18th c. CE strains. These were contrasted to the proportion of probes targeting these different genetic materials. Calculations were based on libraries subjected to a one-hour digestion, as long as at least 500 high-quality alignments against the CO92 genome were found. Inter and intra-century differences in these proportions were compared for 17th and 18th c. CE samples using non-parametric Wilcoxon and Dunn tests in R.

The read-to-reference edit distance distributions (Figure S2) were generated for each of the new genomes characterized in this study from the NM:I field obtained while running the *samtools view* command, considering either the *Y. pestis* (CO92, NC_003143.1), or the *Y. pseudotuberculosis* (IP 32953, NC_006155.1) as a reference. Depth-of-coverage variation was calculated using the PALEOMIX[69] *coverage* command and non-overlapping sliding windows of 1,000 bp for the CO92 chromosome, or 100 bp for the pCD1, pMT1 and pPCP1 plasmids. GC content was estimated for the same windows using seqtk (https://github.com/lh3/seqtk), and visualized together with coverage estimates through circular plots produced using the R *circlize* library.[84] The fraction of the positions that are covered at least once within a given gene was also calculated using the PALEOMIX *depth* command to assess the presence or absence of 235 plague virulence and pathogenicity loci, 190 of which located on the CO92 chromosome, and 45 on the three bacterial plasmids.[43,45,54] Gene coverage plots for virulence genes were generated using the heatmap function from the R ggplot2 library (https://cran.r-project.org/web/packages/ggplot2/index.html). A 49-kb region of CO92 was found absent in most of the phylogenetically-close *Y. pestis* strains from New Churchyard (BED038, BED034, BED028, BED024, BED030), Azov (Azov38), CHE1, Rostov2033, Observance (OBS107, OBS116, OBS137, OBS110, OBS124), and La Major (MAJ75 and MAJ46). However, the coverage of some genes within this region was sporadically greater than zero in a subset of strains, suggesting the possible presence of spurious alignments, despite stringent alignment parameters. This was further confirmed using BlastN[85] against the nucleotide collection (nt), showing best hits against environmental *Yersinia* strains than *Y. pestis*. Those cases are highlighted on the heatmap presented on Figure 7E by crossing the cell presenting the corresponding coverage estimate. We used snpToolkit v2.3[56] and the individual vcf files

generated as described above to annotate the new *Y. pestis* genome and plasmid sequences characterized in this study for synonymous and non-synonymous sequence polymorphisms present in at least 90% of the reads. The minimal depth threshold was set to 3 (-d 3) for bases showing Phred scores above or equal to 30. The samples considered for this analysis were merged by archaeological site using samtools to provide sufficient coverage for the detection of strain-specific SNPs.

Microbial taxonomic profiles were obtained using MetaPhlAn 4[52] on collapsed sequencing unique reads after filtering for those mapping against the human nuclear and mitochondrial genomes following the procedure presented above. To accommodate for variable sequencing efforts across DNA libraries, we random-sampled a number of 100,000 reads. This procedure was repeated 10 times to assess the robustness of the taxonomic assignments, given the limited of reads considered. Reads were mapped against the MetaPhlAn 4 biomarker database using Bowtie2 v.2.3.4.1[73] with default parameters, and further filtered for minimal mapping quality of 25, and PCR duplicates using the *view* and *rmdup* modules from samtools v1.11,[86] respectively. Those species showing relative abundances <1% in a given dataset were defined as "low abundance".

Maximum Likelihood phylogenetic trees were constructed using IQ-TREE v1.6.12[72] and a panel of *Y. pestis* genomes, including 124 modern accessions and 61 second pandemic samples, 12 of which were sequenced here for the first time. Individual base calling was carried out using BCFtools[87] (–ploidy 1), requiring minimum base, mapping and genotype Phred quality scores of 30, and filtering those variants located within 10 bp of indels. The best mutational model was selected based on AICc and BIC values with the Model Finder Plus module (MFP), as part of IQ-TREE. Node support was estimated from a total of 1,000 pseudo-replicates using the ultrafast bootstrap (UFBoot) approximation,[88] optimized by nearest neighbor interchange (NNI), as well as the SH-like approximate likelihood ratio (aLRT) test.[89] The final sequence alignments corresponded, on the first hand, to the entire CO92 chromosome, considering polymorphic sites only, or, on the other hand, were restricted to 29,609 genome-wide variant positions previously characterized to provide robust phylogenetic signal.[43,54] Two sequence alignments were considered, requiring a conservative minimum 3-fold average depth-of-coverage per sample. Accessions for the sequence data underlying the comparative panel of 185 *Yersinia pestis* and 1 *Yersinia pseudotuberculosis* genomes are provided in Table S2.