Integrative approach using *Yersinia pestis* genomes to revisit the historical landscape of plague during the Medieval Period

Amine Namouchi\(^a\), Meriam Guellil\(^ab\), Oliver Kersten\(^a\), Stephanie Hänsch\(^h\), Claudio Ottoni\(^b\), Boris V. Schmid\(^d\), Elsa Pacciani\(^b\), Luisa Quaglia\(^b\), Marco Vermunt\(^c\), Egil L. Bauer\(^d\), Michael Derrick\(^d\), Anne O. Jensen\(^d\), Sacha Kacki\(^ef\), Samuel K. Cohn Jr.\(^g\), Nils C. Stenseth\(^ah\), and Barbara Bramanti\(^a\)

*Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, N-0316 Oslo, Norway; \(^b\)Soprintendenza Archeologia, Belle Arti e Paesaggi di Firenze, Pistoia e Prato, 50125 Florence, Italy; \(^c\)Department of Monuments and Archaeology, Municipality of Bergen op Zoom, 4611 BT-59 Bergen op Zoom, The Netherlands; \(^d\)Norwegian Institute for Cultural Heritage Research, N-0515 Oslo, Norway; \(^e\)Department of Archaeology, Durham University, DH1 3LE Durham, United Kingdom; \(^f\)UMR 5199 De la Préhistoire à l’Actuel: Culture, Environnement et Anthropologie, Centre National de la Recherche Scientifique, University of Bordeaux, 33615 Pessac, France; \(^g\)School of Humanities, University of Glasgow, G12 8QO Glasgow, United Kingdom; \(^h\)Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Ministry of Education, 100084 Beijing, China; and \(^i\)Department of Biomedical and Specialty Surgical Sciences, Faculty of Medicine, Pharmacy, and Prevention, University of Ferrara, 35-44122 Ferrara, Italy

Contributed by Nils C. Stenseth, October 19, 2018 (sent for review July 26, 2018; reviewed by Bruce M. S. Campbell and Ludovic Orlando)

Over the last few years, genomic studies on *Yersinia pestis*, the causative agent of all known plague epidemics, have considerably increased in numbers, spanning a period of about 5,000 y. Nonetheless, questions concerning historical reservoirs and routes of transmission remain open. Here, we present and describe five genomes from the second half of the 14th century and reconstruct the evolutionary history of *Y. pestis* by reanalyzing previously published genomes and by building a comprehensive phylogeny focused on strains attributed to the Second Plague Pandemic (14th to 18th century). Corroborated by historical and ecological evidence, the presented phylogeny, which includes our *Y. pestis* genomes, could support the hypothesis of an entry of plague into Western European ports through distinct waves of introduction during the Medieval Period, possibly by means of fur trade routes, as well as the recirculation of plague within the human population via trade routes and human movement.

**Significance**

While our knowledge of modern plague reservoirs and their hosts is extensive, we have little to no knowledge about the origin of the Medieval plague pandemics or the routes of transmission involved in their spread. Prior genomic data provide a patchy low-resolution picture of the transmission dynamics involved during the Second Plague Pandemic, with only five distinct genomes. We have reevaluated all Medieval strains under the light of archaeological and historical evidence to carefully discuss the involvement of different transmission routes during the Second Plague Pandemic. Our interpretation showcases the importance of trade routes and human movements and further supports the identification of *Yersinia pestis* as the pathogenic agent of the so-called *pestis secunda* (1357–1366).

**Author contributions:** A.N., S.H., N.C.S., and B.B. designed research; A.N., M.G., O.K., S.H., A.O.J., and S.K.C. performed research; E.P., L.O., M.V., E.L.B., M.D., A.O.D., S.K., S.K.C., and B.B. contributed samples; B.V.S., E.P., L.O., M.V., E.L.B., M.D., A.O.D., S.K., S.K.C., and B.B. provided advice; A.N., M.G., O.K., S.H., and C.O. analyzed data; A.N., M.G., O.K., N.C.S., and B.B. wrote the paper with contributions of all authors; and S.H., C.O., B.V.S., M.V., and S.K.C. edited versions of the paper.

**Reviewers:** B.M.S.C., The Queen’s University of Belfast; and L.O., Natural History Museum of Denmark and University of Copenhagen.

The authors declare no conflict of interest.

This open access article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: Raw sequencing reads have been deposited in the European Nucleotide Archive (ENA) (accession no. PRJEB24499).

1To whom correspondence may be addressed. Email: amine.namouchi@ibv.uio.no, n.c.stenseth@ibv.uio.no, or barbara.bramanti@ibv.uio.no.
2M.G., O.K., and S.H. contributed equally to this work.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1812865115/-DCSupplemental.

Published online November 26, 2018.
study can improve the resolution of the *Yersinia pestis* phylogeny and can contribute to the discussion on transmission routes and reservoir establishments during historical pandemics. Using state-of-the-art bioinformatics methods, we evaluated the evolutionary history of *Y. pestis* by incorporating the data reported in this study into a revised phylogeny. We evaluated our results in light of historical, epidemiological, and ecological studies to improve our understanding of the fully documented spectrum of the dynamics at work during the early decades of the Second Plague Pandemic.

**Results**

**Screening and Authentication of *Y. pestis* DNA in Human Skeletons.** The presence of *Y. pestis* in the samples from Saint-Laurent-de-la-Cabrèrissete (SLC) and Bergen op Zoom (BER) has been confirmed by PCR in a previous study by Haensch et al. (13). A similar approach was used to screen 36 individuals from Abbadia San Salvatore (BSS) and nine individuals from Oslo (OSL) (see Methods for more details on the experimental work). Positive samples from Abbadia San Salvatore and Oslo were enriched without initial metagenomic screening.

**Genomic Data Analyses.** We systematically reanalyzed all available raw sequencing data from ancient *Y. pestis* metagenomic datasets from the Second Plague Pandemic. Additionally, we reanalyzed the raw data from two Bronze Age samples and two Justinian Plague samples from the First Plague Pandemic. Most sequencing data stemming from modern samples were included as well. By systematically reanalyzing the available raw data, we were able to avoid any incongruity caused by differences in data filtering and SNP calling. Our custom data workflow (SI Appendix, Fig. S5) also accounted for deamination damage patterns typical for ancient DNA (aDNA). For all generated ancient genomes, reads mapping to the genome of *Y. pestis*, strain CO92, showed a mean read length between 50 and 60 bp, which is consistent with aDNA. The aligned reads also showed a clear deamination profile typical of aDNA misincorporations (SI Appendix, Fig. S6). Moreover, by aligning all generated reads from each sample to *Yersinia pseudotuberculosis*, strain IP-32953, the edit distance analysis clearly showed that the reads belong to *Y. pestis* rather than to *Y. pseudotuberculosis* (SI Appendix, Fig. S7). In fact, for all aDNA reported in this study, the percentage of reads with an edit distance of 0 is higher when aligned against *Y. pestis* vs. *Y. pseudotuberculosis*.

After read quality rescaling, the transitions to transversion ratios of all aDNA assemblies were comparable with those of modern genomes (SI Appendix, Fig. S8), which emphasizes the consistency of our methods and data. In addition, to investigate possible cross-contamination or multiple infection, for each sample, the heterozygous profile analysis (SI Appendix, Fig. S9 and Table S4) showed that, for almost all ancient samples, the...
The heterozygous ratio is low (<0.3), as for modern data. Nevertheless, this ratio is high (>0.4) in some modern data.

The samples from Bergen op Zoom, Ber37 and Ber45, and the sample from Oslo, OSL1, covered more than 87% of the reference genome, with a minimum of $3\times$ depth of coverage (Fig. 2 and SI Appendix, Table S3). For the samples SLC1006 (Saint-Laurent-de-la-Cabrerisse) and BSS31 (Abbadia San Salvatore), we were able to recover 73% and 44% of the genome at $3\times$ depth of coverage, respectively (Fig. 2 and SI Appendix, Table S3).

Overall, after having considered all possible patterns of deamination derived from diagenetic processes, along with sequencing inconsistencies, and having rescaled the data accordingly, the comparison of 126 modern Y. pestis genomes with 15 ancient genomes from the Bronze Age and the First and the Second Pandemic (SI Appendix, Table S1) yielded 2,826 polymorphic sites (SI Appendix, Table S5). These SNPs were grouped as follows: 1,456 nonsynonymous SNPs, 625 synonymous SNPs, and 745 intergenic SNPs. Among the 2,826 identified polymorphic sites, 112 were homoplastic (3.9%) (SI Appendix, Table S6).

**Dating of the Y. pestis Strain from Abbadia San Salvatore.** Attributed to a major epidemic of the second half of the 14th century by archaeological and stratigraphic evidence, the age of the mass grave found at Abbadia San Salvatore (SI Appendix), as well as the Y. pestis genome retrieved, needed to be refined to help with the interpretation of the phylogenetic tree. The consultation of notorial records from the region, with an impressive increase of death-bed testaments and a decrease of nontestamentary property contracts, revealed the presence of a major dramatic event from late June to early September 1348 (SI Appendix, Fig. S3 and Table S2). The analysis of the data confidently demonstrated that such a catastrophic event, which has no equivalent in the course of the century, was responsible for the death of the large number of victims retrieved in Abbadia San Salvatore.

**Evolutionary History Reconstruction.** We built our phylogenetic tree using maximum likelihood, as implemented in IQ-TREE (version 1.5.5) (14). Yersinia pseudotuberculosis, strain IP32953, was used as an outgroup. The adequate substitution model was identified and applied using ModelFinder (15) before building the phylogeny. After testing 484 models, the best-fitting model, according to the Bayesian information criteria (BIC), was the general time reversible (GTR) model (16), with unequal rates and base frequencies with ascertainment (ASC) bias correction and free rate heterogeneity across sites with four rate categories. The topology of the generated tree (Fig. 3 and SI Appendix, Fig. S10) was supported by three different tests (bootstrap, SH-alrt, and local bootstrap probability test) and showed a star-like radiation at the end of branch 0, giving rise to branches 1, 2, 3, and 4. As previously reported, two Bronze Age samples from Bateni and Kytmanovo (Russia) (17), as well as the Justinian samples (9, 18), can be found on branch 0 (Fig. 3A). As expected, the five described ancient genomes cluster with all other known Second Plague Pandemic genomes (Fig. 3).

**Divergence Dates and Tip Dating.** The reconstructed maximum likelihood tree shows a high correlation between the root-to-tip distance and the sampling time (correlation coefficient $r^2 = 0.74$).
The date-randomization test (DRT) showed a different result. In fact, the outcome of the DRT showed that the 95% highest posterior density (HPD) of the substitution rate estimated for both the original and 20 date-randomized datasets is overlapping (SI Appendix, Fig. S11).

Discussion

As mentioned in the Introduction, two scenarios have been proposed in past studies for the persistence of plague during the Second Plague Pandemic. However, the few currently available ancient Y. pestis genomes have proven insufficient to explain the mechanisms underlying the recurrent plague outbreaks in Europe during the Second Plague Pandemic. To provide further insights into the events driving the Second Plague Pandemic and to properly address which of the two proposed scenarios has the highest likelihood, we have evaluated the available data along with five sequenced Y. pestis strains stemming from four European archaeological sites of the 14th century, using an integrative...
approach. By combining historical, archaeological, and genomic data, we attempted to reconstruct the spread of plague in a multifaceted context and to define the routes of transmission following the initial introduction of plague during the Black Death, accordingly. The need of an integrative approach is also supported by the lack of temporal signal in the data, which could lead to inaccurate estimates of substitution rates and timescale, as shown by the date-randomization test, which, so far, represents one of the most robust methods to estimate temporal signal.

Two of our *Y. pestis* strains, SLC1006 and OS1L, had genomes identical to the clones of London3148 and Barcelona3031 (Fig. 3B), both attributed to the Black Death (7, 19). The strains were extracted from skeletons excavated in Saint-Laurent-de-la-Cabrerisse (France) and Oslo (Norway). These results allowed us to place the outbreak in Saint-Laurent-de-la-Cabrerisse into the Black Death period, as had been proposed by Haensch et al. (13). While no direct historical description of the arrival of plague in Saint-Laurent-de-la-Cabrerisse is available, reports of plague epidemics in the surrounding region (Southern France) in the second half of the 14th century exist (13, 20). Historically, the regions surrounding the site of Saint-Laurent-de-la-Cabrerisse (Aude, Occitania) are known to have been ravaged by plague from November 1347 onwards; specifically, plague was present in Narbonne and Carcassonne, cities close to Saint-Laurent-de-la-Cabrerisse, from March to April 1348. Interestingly, studies have proposed that plague reached Barcelona from the very same region by April–May 1348 (3, 21). The genomic similarity between the strains SLC1006 and Barcelona3031 seems to corroborate the hypothesis of a common origin. The strain from the St. Nicola's Church's cemetery in Oslo (Norway) also clusters with Black Death strains. The time of insurrection of the Medieval plague epidemic in Oslo is debated. However, radiocarbon dating and archaeological data (see Appendix) place the strain OS1L at around 1350, which is in line with time lines proposed by some historical studies (22, 23), but in disagreement with others, which had given the Black Death in Oslo an earlier origin (1348) (24, 25). It is important to note that, although the samples identified in Saint-Laurent-de-la-Cabrerisse, Barcelona, London, and Oslo seem to be identical, this similarity is based solely on the regions of the chromosome covered across all of these samples. Regions not covered, as well as structural and plasmid variation, which could account for additional differences among the strains, have not been evaluated or considered in this study or any other previous whole genome aDNA plague phylogeny.

In contrast to the above-mentioned strains, the clone recovered from Abbadia San Salvatore (Siena, Italy) accumulated two additional, specific SNPs (T3529404C and A3897987T), compared with the other Black Death samples (Fig. 3B), and might therefore be interpreted as a more recent derivate. While no direct historical sources are available to confirm the presence of plague in Abbadia San Salvatore in the 14th century, historical data collected and evaluated for this study have revealed a dramatic increase in death-bed testaments in the region from June to September 1348. These data support an account by Agnolo di Tura (26), who recorded the desolation of the city of Siena and its surrounding countryside by the plague in 1348 (27). Starting from the port city of Pisa in mid-January 1348 (4), plague made its way through Tuscany and was subsequently reported in Florence (March–April 1348) before reaching Siena and its countryside in May 1348, likely over the Via Francigena, a well-frequented route by pilgrims (4). Considering this chain of events, it is reasonable to assume that the Abbadia San Salvatore strain reached Pisa over sea and continued to spread over land by infected individuals or goods. Given the number of subsequent outbreaks recorded over this short period of time, it is plausible that the additional two SNPs were acquired through the large transmission chain in Italy, rather than having been gained before the strain's arrival in Pisa or within a newly established local wildlife reservoir.

While the Black Death branches exhibit low variability, the subbranch departing from the Black Death clones, which gave rise to the outbreak in Ellwangen (1485–1627 by radiocarbon dating) (7) and culminated in the outbreak of Marseille (1720–1721), sees a considerable increase in variability (Fig. 3B). Previous literature (7) has attributed this branch to a novel stable reservoir established at the time of the Black Death, in which favorable environmental conditions have enabled a permanent presence of *Y. pestis* in a host population of wild rodents. The study argued that this ancient reservoir was probably situated in the Alpine region (28). However, ecological and historical–epidemiological studies have concluded that the presence of a plague reservoir in Western Europe during the Second Plague Pandemic was highly improbable (10, 12). Taking a closer look at historical sources, it becomes clear that at least the Great Plague of Marseille (1720–1722) (29, 30) had an extra-Western European source. In fact, historical records clearly associate the start of the outbreak with the successive arrival of several ships from Sidon (Lebanon) (31). While no additional information regarding the geographic location of the wildlife reservoir feeding plague to Europe during the 17th and 18th century is available, studies have argued that, before being introduced to the Mediterranean basin in 1347, plague spread to the territories around the Black Sea and to the Middle East (32, 33) and thus had an initial opportunity to be disseminated in this region.

On the *Y. pestis* phylogenetic tree, specifically on branch 1, we found two additional genomes isolated from skeletons recovered in Bergen op Zoom. The two identical isolates show four additional SNPs, compared with the Black Death strains [T6999494C, G2218046T (homoplasic), T2262577G, and A2894703T] (Fig. 3B) and are basal to the clone found in the St. Mary Grace Cemetery in London (London-ind6330) and in Bolgar (BolgarCity2370). Although no direct historical evidence exists (13), the large number of victims retrieved in Bergen op Zoom likely died during documented plague outbreaks, which occurred in the Belgian Hainault region and the region of Flanders and Antwerp between 1358 and 1363 (34). Based on the phylogeny, the strain London-ind6330 (with three specific SNPs G1130135T, C1159914A, and G4301295T) could be attributed to the wave of plague outbreaks that hit Medieval Europe in 1361, known as *"pestis secunda,“* and might have been imported directly from Bergen op Zoom, which supports the fact that St. Mary Grace is in general considered a post-Black Death cemetery. The strain isolated from Bolgar (7) was placed one SNP (T3806677C) further on branch 1 from the clones of Bergen op Zoom and is characterized by one additional specific SNP (G3643837T). The *Y. pestis* clone found in Bolgar, which had previously been attributed to 1362+ (7), could now be more precisely attributed to the outbreaks, which started in 1346 in Nizhnii Novgorod (Fig. 4) and swept throughout all towns in Russia until 1366, killing up to 100 people per day (35). During an excavation of the marketplace of Bolgar, which was active from the 1340s/1350s and was destroyed by a fire in the 1360s/1370s, archaeologists recovered a large range of trade goods, in particular artifacts originating from Flemish towns (Tournai, Ypres, etc.) (36, 37), which constitutes direct evidence for trading contacts between the Low Countries and the Volga region. Therefore, it cannot be excluded, that, as reported in previous studies (7), plague was imported to Bolgar from Western Europe, where it had established in a reservoir from the Black Death. Nevertheless, the possibility that the three places (London, Bergen op Zoom, and Bolgar) were infected independently from extra-Western European reservoirs in the 1360s should not be ignored.

Here, we considered historical evidence and propose that the phylogeny of the plague strains of the 14th century (branch 1) may be also explained by independent introductions on the fur
trade routes, a complex network of interconnected maritime, riverine, and overland routes into Western Europe. During the second half of the 14th century, Bolgar and Novgorod (Russia) had played a central role in fur export to cities like Hamburg and Lübeck. The regions highlighted in orange represent known modern plague reservoirs. The darker orange delimits the region in which we believe secondary plague reservoirs were established before the Black Death. The red dots represent the locations of all known ancient Y. pestis genomes. Black lines, Silk Road; dark-blue lines, maritime trade routes.

Fig. 4. Schematic representation of the connection between the fur trade routes and the spread of plague during the early stage of the Second Plague Pandemic (14th century). This simplified map shows cities strategically situated along the fur trade routes (indicated as a white line). The city of Novgorod (Russia) has played a central role in fur export to cities like Hamburg and Lübeck. The regions highlighted in orange represent known modern plague reservoirs. The darker orange delimits the region in which we believe secondary plague reservoirs were established before the Black Death. The red dots represent the locations of all known ancient Y. pestis genomes. Black lines, Silk Road; dark-blue lines, maritime trade routes.

Overall, evidence from different disciplines increasingly suggests successive and independent introductions of plague to Western Europe via the transport of infected individuals and goods on trade routes during the Second Plague Pandemic. By analyzing five ancient genomes from the Second Plague Pandemic, adding them to the established Y. pestis phylotype, and evaluating them with a historical background, our study aimed to answer some of the most debated questions on plague in Western Europe during the Second Plague Pandemic. All things considered, the hypothesis that Y. pestis reached Europe through multiple introductions during the Middle Ages through different routes, including the fur trade, appears very plausible, at least during the Second Plague Pandemic.

Methods

Samples. The Y. pestis strains presented in this study stem from four European sites dating to the second half of the 14th century (Fig. 1): Saint-Laurent-de-la-Cabrerisse (France) (13), Bergen op Zoom (The Netherlands) (13), Oslo (Norway) (42), and Abbadia San Salvatore (Italy). For this study, a total of 126 Y. pestis strains (SI Appendix, Table S1) and 15 published ancient DNA samples were used, including 2 Bronze Age samples (17), 2 Justinian plague samples (9, 43), and 10 samples from the Second Plague Pandemic (7, 19).

Dating the Strain of Abbadia San Salvatore with Historical Data. We collected data of testaments and contracts from 1340 to 1381 (Table S2), which summarize parchments from Monte Amiata, the region where Abbadia San Salvatore is placed.

Sample Preparation. We screened a total of 6 teeth from the site of Saint-Laurent-de-la-Cabrerisse, 41 teeth from the site of Abbadia San Salvatore, 3 teeth from the site of Bergen op Zoom, and 18 teeth from the site of Oslo. Laboratory work was performed at the Paleogenetics Laboratories at the University of Mainz, (Germany) and at the Ancient DNA Laboratory at the University of Oslo (Norway). Both laboratories are solely dedicated to the analysis of ancient samples and are subjected to strict anticontamination protocols, including full overnight UV irradiation. Target enrichment was performed at the post-PCR capture laboratory of the Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Oslo.

Extraction and PCR Screening. The teeth were decontaminated, sandblasted, and milled to fine powder, as previously described (13). aDNA was extracted using either a previously published phenol chloroform protocol (13) or modified versions of silica-based protocols based on Brotherton et al. (44) and Dabney et al. (45). We used 0.2 to 0.5 g of tooth powder for phenol chloroform extractions and 0.1 to 0.25 g for the silica-based extraction methods. Two samples from Bergen op Zoom (Ber37 and Ber45), 8 samples from Saint-Laurent-de-la-Cabrerisse, and 39 samples from Abbadia San Salvatore were extracted by phenol chloroform extraction (protocol A) (SI Appendix). Three teeth from Saint-Laurent-de-la-Cabrerisse, 2 teeth from Abbadia San

with the onset of the Black Death, whose origins remain unclear before its arrival from Caffa. Suggestively, Ibn al-Wardi, an Arab historian, who died of plague in Aleppo (Syria) in 1349, described how returning merchants from Crimea, attributed the beginning of the “contagion” to the “Land of Darkness,” in which the pestilence had been present for 15 years before reaching the West (3). Without having visited the West-Ural region from which fur was import to the Golden Horde, the Arab merchants likely just meant that the source of the Black Death was associated with the fur routes, placed somewhere before Sarai (seen from Caffa), where plague was present before 1347 (3), and hence probably a region next to the Caspian Sea—an area with well-established plague reservoirs today (Fig. 4). From this region, plague may have been transported on the fur trade network to the whole of Europe, not only during the Black Death, but also later, with (possibly) independent goods deliveries to Bergen op Zoom, London, and Bolgar. Extra-SNPs, which developed on subbranches lateral to branch 1, might also have been accumulated during chains of epidemics on the long journey to Western Europe.

Overall, evidence from different disciplines increasingly suggests successive and independent introductions of plague to Western Europe via the transport of infected individuals and goods on trade routes during the Second Plague Pandemic. By analyzing five ancient genomes from the Second Plague Pandemic, adding them to the established Y. pestis phylotype, and evaluating them with a historical background, our study aimed to answer some of the most debated questions on plague in Western Europe during the Second Plague Pandemic. All things considered, the hypothesis that Y. pestis reached Europe through multiple introductions during the Middle Ages through different routes, including the fur trade, appears very plausible, at least during the Second Plague Pandemic.

For this study, a total of 126 Y. pestis strains (SI Appendix, Table S1) and 15 published ancient DNA samples were used, including 2 Bronze Age samples (17), 2 Justinian plague samples (9, 43), and 10 samples from the Second Plague Pandemic (7, 19).

Dating the Strain of Abbadia San Salvatore with Historical Data. We collected data of testaments and contracts from 1340 to 1381 (SI Appendix, Fig. 53 and Table S2), which summarize parchments from Monte Amiata, the region where Abbadia San Salvatore is placed.

Sample Preparation. We screened a total of 6 teeth from the site of Saint-Laurent-de-la-Cabrerisse, 41 teeth from the site of Abbadia San Salvatore, 3 teeth from the site of Bergen op Zoom, and 18 teeth from the site of Oslo. Laboratory work was performed at the Paleogenetics Laboratories at the University of Mainz, (Germany) and at the Ancient DNA Laboratory at the University of Oslo (Norway). Both laboratories are solely dedicated to the analysis of ancient samples and are subjected to strict anticontamination protocols, including full overnight UV irradiation. Target enrichment was performed at the post-PCR capture laboratory of the Centre for Ecological and Evolutionary Synthesis (CEES), University of Oslo, Oslo.

Extraction and PCR Screening. The teeth were decontaminated, sandblasted, and milled to fine powder, as previously described (13). aDNA was extracted using either a previously published phenol chloroform protocol (13) or modified versions of silica-based protocols based on Brotherton et al. (44) and Dabney et al. (45). We used 0.2 to 0.5 g of tooth powder for phenol chloroform extractions and 0.1 to 0.25 g for the silica-based extraction methods. Two samples from Bergen op Zoom (Ber37 and Ber45), 8 samples from Saint-Laurent-de-la-Cabrerisse, and 39 samples from Abbadia San Salvatore were extracted by phenol chloroform extraction (protocol A) (SI Appendix). Three teeth from Saint-Laurent-de-la-Cabrerisse, 2 teeth from Abbadia San
Salvatore, and 18 teeth from Oslo were extracted by silica extraction based in Haensch et al. (13), were screened for human and Y. pestis DNA using previously published primers: ypl11D/ypl10R as published in Raoult et al. (46), caf1 caf1U2L2 as published in Haensch et al. (13), and human mitochondrial HVR1 primers L16209 (47) and H16248 (48). PCR conditions were as described in Haensch et al. (13). Positive samples were shotgun sequenced on an Illumina HiSeq2500 system (125 bp paired end PEI) at the Norwegian Sequencing Centre (NSC) at the University of Oslo.

Library Preparation and Target Enrichment. Library preparation was done following a modified Meyer and Kircher (49) protocol (for more details, see SI Appendix). Amplified libraries were purified using commercial kits (Stratagene PCRapace or Qiagen MinElute PCR purification kits, followed by an AmpPureXP beads purification) and subsequently quantified on a Bioanalyzer 2100 expert dSNP Sensor High Sensitivity Chip and using a Qubit HS assay kit. When necessary, re amplifications were performed with 15S and IS6 primers following the original protocol by Meyer and Kircher (49). Positive samples, screened via standard PCR and/ or shotgun metagenomics, were enriched for Y. pestis DNA. Over the course of this study, we used two different custom bait kits from different manufacturers for in-solution target enrichment (for more details, see SI Appendix).

Sequencing. High throughput sequencing (125 bp PE) of the captured libraries was performed on an Illumina HiSeq2500 system at the NSC (University of Oslo). Capture products were pooled for sequencing and split over two lanes (single capture and double capture products were sequenced separately), and Bss 31d and Ber37c products were each sequenced and pooled with other samples on different lanes and flow cells. Libraries for OS1A were sequenced and multiplexed with other libraries over two lanes.

Shotgun Metagenomics Data Analysis. The presence of Y. pestis in shotgun metagenomics datasets was investigated using two methods. The first method assessed the abundance of specific genes and was implemented in Metaphlan (50). The second method is based on the exact alignments of k-mer sequences and was implemented in CountK (51). Amplified libraries were prepared using a protocol based on Dabney et al. (45) (protocol C, SI Appendix). All extractions included negative milli- and control extractions.

All extracts, which had not previously been screened for Y. pestis in Haensch et al. (13), were screened for human and Y. pestis DNA using previously published primers: ypl11D/ypl10R as published in Raoult et al. (46), caf1 caf1U2L2 as published in Haensch et al. (13), and human mitochondrial HVR1 primers L16209 (47) and H16248 (48). PCR conditions were as described in Haensch et al. (13). Positive samples were shotgun sequenced on an Illumina HiSeq2500 system (125 bp paired end PEI) at the Norwegian Sequencing Centre (NSC) at the University of Oslo.

Library Preparation and Target Enrichment. Library preparation was done following a modified Meyer and Kircher (49) protocol (for more details, see SI Appendix). Amplified libraries were purified using commercial kits (Stratagene PCRapace or Qiagen MinElute PCR purification kits, followed by an AmpPureXP beads purification) and subsequently quantified on a Bioanalyzer 2100 expert dSNP Sensor High Sensitivity Chip and using a Qubit HS assay kit. When necessary, reamplifications were performed with 15S and IS6 primers following the original protocol by Meyer and Kircher (49). Positive samples, screened via standard PCR and/or shotgun metagenomics, were enriched for Y. pestis DNA. Over the course of this study, we used two different custom bait kits from different manufacturers for in-solution target enrichment (for more details, see SI Appendix).

Sequence alignment, assembly, and polymorphism identification. The reads were identified and marked using picard-tools (https://broadinstitute.github.io/gatk/) option -n to 0.1 (for modern data, the -n option was set to 0.01). Duplicated reads were trimmed and merged using ClipAndMerge (https://github.com/apel/ClipAndMerge). Thereafter, we used bwa-mem (53) to align the merged reads against Y. pestis strain CO92 by disabling the seed option (-i) and setting the option -n to 0.1 (for modern data, the -n option was set to 0.01). Duplicated reads were identified and marked using picard-tools (https://broadinstitute.github.io/picard/). Local alignment of reads was performed around gapped regions before SNP calling using the GATK IndelRealigner module (https://software. broadinstitution.org/gatk/). DeDup (https://github.com/apel/DeDup) was also used to remove duplicated sequences as it has been specifically designed for merged reads. Aligned Binary Alignment Map (BAM) files of ancient DNA samples were analyzed using MapDamage2 (54) to assess and recalibrate aDNA damage patterns in the form of C-to-T or G-to-A conversions.

Edit Distance. To verify that the generated reads belong to Y. pestis rather than Y. pseudotuberculosis, all reads from each sample were mapped to the genome of Y. pseudotuberculosis, strain IP-32953, as described above. The edit distance, which defines the minimal number of substitutions, insertions, and deletions to transform the read sequence to its homologous sequence in the reference genome, was calculated using BAMStats (bamstats.sourceforge.net).

SNP Call. SNP calling was performed using samtools (55, 56) and bcftools (https://samtools.github.io/bcftools/). SNPs located within a frame of 10 bp from indels were excluded with samtools. For each sample, all identified SNPs were filtered and annotated using a freely available tool, snp Toolkit (https://bitbucket.org/Amine-Namouchi/snp toolkit/). snp Toolkit allows one to filter and annotate SNPs from vcf files by providing the genbank file of the reference sequence used during reads alignment. SNPs were filtered according to three criteria: quality score (≥30), depth of coverage (≥3), and allele frequency (≥0%). In addition, SNPs that are close to each other by less than 10 bp were excluded during the annotation process using snp Toolkit with option -f. All generated annotation output files where thereafter compared against each other using the snp Toolkit option combine. For each polymorphic site, snp Toolkit checks the corresponding alignment region in the bam file to accurately determine the distribution of each polymorphic site by considering the same region in all other aDNA strains of the Second Plague Pandemic. The efficiency of the method implemented in snp Toolkit was assessed after manual inspection of these polymorphic sites using Integrative Genomics Viewer (IGV) (57, 58). The option combine produces two output files. The first one represents the distribution of all identified polymorphic sites of all analyzed samples. The second file is a concatenation of all polymorphic sites in fasta format.

Phylogenetic Analysis. Phylogenies were inferred using IQ-TREE (14). IQ-TREE was run using ModelFinder (15), with the option -m MPF to infer the best substitution model for building the maximum likelihood phylogenetic tree. A total number of 484 models were tested. One thousand fast bootstrap replicates were performed to assess statistical support (59). In addition, branch supports were also assessed through the single branch test SHA-LRT (60) with 1,000 replicated as well as local bootstrap probability tests (61). As the concatenated SNPs include missing information indicated by an exclamation mark in the fasta file generated using snp Toolkit, we also used the ASC option to account for ascertainment bias correction. The generated tree was visualized using FigTree (version 1.2.1, tree.bio.ed.ac.uk/software/figtree), and the alignment of each to each SNP were mapped in the phylogenetic tree using Fastml v3.1 (62).

Divergence Time and Tip Dating Analysis. Temporal signal detection in the sequence data was performed using TempEst (63). As previously described (17, 64, 65), the lognormal relaxed clock model and constant population size models implemented in BEAST (66) were applied to the alignment sequence and phylogenetic tree to determine the divergence dates at each node. To evaluate the reliability of the Bayesian inference, we applied the date randomization test (DRT) as implemented in the R package tipdatingBEAST (67) to generate 20 BEAST input files in xml format with randomized dates. As the out dataset included a large number of modern sequences, the randomization process was only applied to ancient sequences, as previously recommended (68). BEAST was run on each file for 50 million iterations.

Data Availability. Raw sequencing reads produced for this study have been deposited in the European Nucleotide Archive (ENA) under accession number ENA: PRJEB24499.

ACKNOWLEDGMENTS. We thank Lars Wallace for useful comments and discussion; and Joachim Burger for providing access to the aDNA facilities at the University of Mainz, Germany. This project was funded by the European Research Council (ERC) under the FP7-IDEAS-ERC Program (Grant 324249). Part of the data analysis was performed on the Abel Cluster, owned by the University of Oslo and the Norwegian metacentre for High Performance Computing (NOTUR) and operated by the Department for Research Computing at the University Center for Information Technology Services (USIT), University of Oslo. C.O. is funded by Satsningsmiljø Grant Comparative Infection Biology (COMBI), Department of Biosciences, University of Oslo.

1. Stenseth NC, et al. (2008) Plague: Past, present, and future. PLoS Med 5:e13. 2. Bramanti B, Stenseth NC, Wallace L, Lei X (2016) Plague: A disease which changed the path of human civilization. Adv Exp Med Biol 918:1-20. 3. Benedictow OJ (2004) The Black Death 1346-1352: The Complete History (Boydell Press, Woodbridge, UK). 4. Michele da Piazza (1980) Cronaca, ed Giuffrida A (ILA Palma, Palermo, Italy). 5. Campbell BMS (2016) The Great Transition: Climate, Disease and Society in the Late-Medieval World (Cambridge Univ Press, Cambridge, UK).

6. Vogler AJ, et al. (2013) A decade of plague in Madagascar, Madagascar: Insights into the global maritime spread of pandemic plague. MBio 4:e00623-12.

7. Spyrou MA, et al. (2016) Historical Y. pestis genomes reveal the European Black Death as the source of ancient and modern plague pandemics. Cell Microbe 19:874-881. 8. Buranand J, et al. (2011) A draft genome of Yersinia pestis from victims of the Black Death. Nature 478:506-510. 9. Wagner DM, et al. (2014) Yersinia pestis and the plague of Justinian 541-543 AD: A genomic analysis. Lancet Infect Dis 14:319-326. 10. Schmid BV, et al. (2015) Climate-driven introduction of the Black Death and successive plague reintroductions into Europe. Proc Natl Acad Sci USA 112:3020–3025. 11. Yue RPH, Lee HF, Wu CYH (2016) Navigable rivers facilitated the spread and reoccurrence of plague in pre-industrial Europe. Sci Rep 6:34867.
12. Yue RPH, Lee HF, Wu CYH (2017) Trade routes and plague transmission in pre-industrial Europe. Sci Rep 7:12973.

13. Haensch S, et al. (2010) Distinct clones of Yersinia pestis caused the Black Death. PLoS Pathog 6:1001134.

14. Nguyen L-T, Schmidt HA, van Haeseler A, Minh BQ (2015) IQ-TREE: A fast and ef
cient bootstrap algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268-274.

15. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat Methods 14:587-589.

16. Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. American Mathematical Society: Lectures on Mathematics in the Life Sciences (American Mathematical Society, Fort Collins, CO), Vol 17, pp 57-86.

17. Rasmussen S, et al. (2015) Eighteenth century Yersinia pestis genomes reveal the long-term persistence of an historical plague focus. eLife 5:e12994.

18. Harbeck M, et al. (2013) Yersinia pestis DNA from skeletal remains from the 6th century AD reveals insights into Justinianic plague. PLoS Pathog 9:e10033349.

19. Bos KI, et al. (2016) Eighteenth century Yersinia pestis genomes reveal the long-term persistence of an historical plague focus. eLife 5:e12994.

20. Kacki S, Rahalison L, Rajerison M, Ferroglio E, Bianucci R (2011) Black Death in the rural cemetery of Saint-Laurent-de-la-Cabrerisse Aude-Languedoc, southern France, 14th century: Immunological evidence. J ArchaEOl Sci 38:581-587.

21. Oeding P (1990) [The Black Death in Norway]. Namouchi et al.

22. Balme D (2012) Black Death in the Volga River Region: Archaeological and Historical Evidence. Volgad Rev Archaeol 4:9-33.

23. Veale EM (2003) The English Fur Trade in the Later Middle Ages (London Record Society, London), 2nd Ed, Vol 38. Available at https://www.british-history.ac.uk/london-record-soc/vol38. Accessed July 3, 2018.

24. Suhm PF, Nyerges R (2012) Historie AF Danmark: Fra AAR 1340 Til 1375 (Nabu Press, Charleston, SC). Available at https://www.betterworldbooks.com/product/detail/ historie-of-danmark-fra-aar-1340-til-1375-volume-13-1274946565. Accessed July 3, 2018.

25. Varlik N (2015) Plague and Empire in the Early Modern Mediterranean World: The Ottoman Experience, 1347-1600 (Cambridge Univ Press, Cambridge, UK).

26. Martin J (1978) The Land of Darkness and the Golden Horde. The fur trade under the Mongols XIII-XVIIth centuries. Cah Monde Russe Sov 19:401-421.

27. Guellil M, et al. (2018) Genomic blueprint of a relapsing fever pathogen in 15th century Scandinavia. Proc Natl Acad Sci USA 115:10422-10427.

28. Feldman M, et al. (2016) A high-coverage Yersinia pestis genome from a sixth-century Justinianic plague victim. Mol Biol Evol 33:2911-2923.

29. Brotherton P, et al.; Genographic Consortium (2013) Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat Commun 4:1764.

30. Dabney J, et al. (2013) Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc Natl Acad Sci USA 110:15758-15763.

31. Raoult D, et al. (2000) Molecular identification by “suicide PCR” of Yersinia pestis as the agent of medieval Black Death. Proc Natl Acad Sci USA 97:12800-12803.

32. Handt O, Krings M, Ward RH, Pääbo S (1996) The retrieval of ancient human DNA from skeletal remains. Am J Hum Genet 59:368.

33. Haak W, et al. (2005) Ancient DNA from the first European farmers in 5000 years ago. Cell 122:557-566.

34. Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010:db.prot5448.

35. Truong DT, et al. (2015) MetaPhlAn2: An enhanced metagenomic taxonomy profiling. Nat Methods 12:902-903.

36. Wood DE, Salzberg SL (2014) Kraken: Ultrafast metagenomic sequence classification using exact alignments. Genome Biol 15:466.

37. Köster J, Rahmann S (2012) Snakemake-A scalable bioinformatics workflow engine. Bioinformatics 28:2520-2522.

38. Li H, Dunbrin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-1760.

39. Johnson G, Gilmohac A, Schubert M, Johnson PLF, Orlando L (2013) mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29:1682-1684.

40. Li H, et al.; 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMTools. Mol BioI Evol 26:2078-2079.

41. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. Bioinformatics 27:2987-2993.

42. Thévenodt H, Robinson JT, Mesirov JP (2013) Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. Brief Bioinform 14:178-192.

43. Robinson JT, et al. (2011) Integrative genomics viewer. Nat Biotechnol 29:24-26.

44. Nguyen L-T, Minh BQ, van Haeseler A (2013) Ultrafast approximation for phylo-
genetic bootstrap. Mol Biol Evol 30:1188-1195.

45. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-
likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol 59:307-321.

46. Adachi J, Hasegawa M (1996) MOLPHY version 3.2: Programs for molecular phylo-
genetics based on maximum likelihood. Available at citeseerx.ist.psu.edu/viewdoc/
summary?doi=10.1.1.476.8552. Accessed September 5, 2018.

47. Ashkenazy H, et al. (2012) FastML: A web server for probabilistic reconstruction of ancestral sequences. Nucleic Acids Res 40:W580-W584.

48. Rambaut A, Lam TT, Max Carvalho L, Pybus OG (2016) Exploring the temporal structure of hetrochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol 2:wve007.

49. Cui Y, Song Y (2016) Genome and evolution of Yersinia pestis. Yersinia pestis: Retrospective and Perspective, eds Yang R, Anisimov A (Springer, Dordrecht, The Netherlands), pp 171–192.

50. Andrades Valveria A, et al. (2017) The stone age plague and its persistence in Eurasia. Curr Biol 27:3683-3691.el8.

51. Drummond AJ, Bouckaert RR (2015) Bayesian Evolutionary Analysis with BEAST 2 (Cambridge Univ Press, Cambridge, UK).

52. Rieux A, Khatkhchian CE (2017) tipdateing: An R package to assist the im-
plementation of phylogenetic tip-dating tests using beast. Mol Ecol Resour 17:
606-613.

53. Ho SYW, et al. (2011) Bayesian estimation of substitution rates from ancient DNA sequences with low information content. Syst Biol 60:366-375.

54. Eren AM, et al. (2015) An advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319.