SUPPLEMENTARY INFORMATION

A Genomic and Historical Synthesis of Plague in 18th Century Eurasia.

Meriam Guellil, Oliver Kersten, Amine Namouchi, Stefania Luciani, Isolina Marota, Caroline A. Arcini, Elisabeth Iregren, Robert A. Lindemann, Gunnar Warfvinge, Lela Bakanidze, Lia Bitadze, Mauro Rubini, Paola Zaio, Monica Zaio, Damiano Neri, Nils C. Stenseth, Barbara Bramanti.

www.pnas.org/cgi/doi/10.1073/pnas.2009677117

LABORATORY METHODS

Sample Preparation

For this study, twelve teeth from six individuals excavated at the site of Pestbacken, 24 teeth/individuals from the site of San Procolo A Naturno and four individuals/teeth from Maist were analysed via aDNA analysis. The samples were introduced and processed in a dedicated ancient DNA (aDNA) laboratory at the University of Oslo, Department of Biosciences.

Upon introduction into the facility, the samples were UV irradiated on all sides in a dedicated UV cupboard, followed by the mechanical removal of the outer 1-2 mm surface of the teeth with a dental sandblaster (Renfert Basic IS). The teeth were subsequently UV irradiated on all sides for a second time and milled to powder using a Retsch Oscillating Mill.

The 24 samples from San Procolo a Naturno were prepared and powdered in the aDNA dedicated Laboratory of Molecular Archaeo- Anthropology/Ancient DNA at the University of Camerino. Before and after the mechanical removal of the outer surface of the teeth using a rotary hand-drill, the samples were UV-irradiated on each side for two minutes. They were then powdered using a mortar and pestle as well as liquid nitrogen. Powder aliquots of approximatelly 120 mg were sent to the University of Oslo in UV-irradiated lobind tubes.

aDNA extraction

We used two different protocols for aDNA extraction (see Table S1 for sample specific information):

Protocol A: Ancient DNA was extracted from 0.1-0.26 g of tooth powder using a modified protocol version from Brotherton et al.(1). Powder was incubated under rotation overnight in 4.31 ml of lysis buffer (0.5 M EDTA, pH 0.8; 0.5% N.-Laurylsarcosine; 0.25 mg/mL Proteinase K). A silica suspension was prepared as detailed in Brotherton et al.(1). The lysates were pelleted, and supernatants were transferred into a 50ml falcon tube, and mixed with 125 µl silica suspension and 16 ml binding buffer (13.5 ml Qiagen QG Buffer and 2.86 ml of a solution made up of 1x Triton X100, 20 mM NaCl and 0.2 M acetic acid). The samples were then incubated under rotation for 2 hours at room temperature. Subsequently, the samples were centrifuged for 2 minutes at 13,000 rpm and most of the supernatant was discarded. The silica pellet was transferred into 2 ml safe-lock tubes in the remaining supernatant and pelleted again, prior to discarding the remaining supernatant. The pellets were then washed three times with 1 ml ethanol 80% and dried at 37°C for approximately 30 minutes. Dried pellets were eluted in 150 µl of pre-warmed (50°C) Qiagen EB buffer, incubated for 10 minutes in a thermomixer at 37°C, and finally centrifuged for 1 minute at 10,000 rpm. Eluates were kept at -20°C until further use.

Protocol B: Ancient DNA was extracted from 0.120 g of tooth powder using a modified version of the extraction protocols published by Dabney et al.(2) and Bennett et al.(3). 1 ml of lysis buffer (0.25

mg/ml Proteinase K, 0.45M EDTA) was added to each powder aliquot and incubated on a nutator at 38 °C overnight. The lysate was then added to a binding buffer (6X QG Buffer, 4X Isopropanol) and passed through a Qiagen MinElute Spin Column using a Zymo-SpinV 15 ml reservoir extension and a Qiagen Qiavac Vacuum Manifold. The column was washed twice with 750 µl of Qiagen PE Buffer. After a one minute dry spin, the DNA was eluted in two steps in a total of 50 - 60 µl of pre-heated (50°C) Qiagen EB Buffer.

qPCR Screenings

All extracts including milling and extraction blanks were screened via qPCR for human and *Y. pestis* DNA using previously published *Y. pestis* (pla and caf1M primers as published in Schuenemann et al.(4) and human mitochondrial (HVR1 L16209/H16348(5, 6)) primers. qPCR reactions (Maist and San Procolo a Naturno) were prepared with Roche LifeScience FastStart Essential DNA Green Master Mix, 0.5 µM per primer and 3 µl of template, and were run on a Roche Lightcycler 96 with the following conditions: 95°C for 600 sec, 50-55 cycles of 95°C for 40 sec, 61°C(pla)/63°C(caf1M) for 40 sec and 72°C for 15 sec, two melting steps at 95°C/10 sec and 65°C/60 sec, and a final step at 97°C for 1 sec. The samples from Pestbacken were screened via standard PCR using protocols described in Namouchi et al.(7). All of the blanks were negative and showed no sign of contamination.

qPCR reactions screening for human DNA were prepared with 2 µl of extract in a final volume of 20 µl using 1.2x AmpliTaq Gold Buffer, 2.5 U AmpliTaq Gold Polymerase, 3 mM MgCl2, 0.2 mM dNTP mix (Qiagen), 0.2 µM of each primer, 0.4 mg/ml BSA (Roche), 0.8 µl of a custom SYBR Green I Master Mix (SYBR Green I, ROX and DMSO) and Nuclease-Free Water. The cycle program was identical to the *Y. pestis* DNA amplifications except for the annealing temperature of 58°C and using only 45 amplification cycles.

Library preparation and Sequencing

Libraries from all positive extracts were built according to Meyer and Kircher(8) with the following modifications: a) a Qiagen MinElute PCR purification kit with Qiagen PB buffer (5X) and one wash with Qiagen PE Buffer was used for all purification steps; b) following the adapter fill-in step and after the initial 37°C incubation, samples were incubated at 80°C for 20 min in order to denature Bst, and were used for the indexing PCR setup without prior purification; c) 1.25 µM of adaptermix was used in the adapter ligation master mix; d) Amplitaq Gold Polymerase was used for the indexing PCR setup. 40 µl of denatured adapter fill-in reactions were split into three reactions and added to 20 µl of indexing PCR master mix with 7bp Illumina indices (1.2x AmpliTaq Gold Buffer, 3 mM MgCl2, 0.05 U/µl AmpliTaq Polymerase, 0.4 mg/mL BSA, 200 µM Qiagen dNTPs, 200 µM primer IS4/indexing primers).

Indexing PCR conditions were as follows: 95° for 6 min, 12 cycles at 95°C/40 sec, 60°C/40 sec and 72°C/40 sec and a final elongation step at 72°C for 10 min. All indexed libraries were purified using Stratec MSB Spin PCRapace® columns and Agencourt AmpureXP® and quantified using an Agilent High Sensitivity Bioanalyzer Chip and a Qubit dsDNA HS Assay. Whenever re-amplification was necessary, we used IS5/IS6 primers(8). Libraries were sequenced on an Illumina Hiseq 2500 (125bp PE) system at the Norwegian Sequencing Centre (NSC) at the University of Oslo.

Target enrichment

Yersinia pestis DNA in the libraries from plague-positive individuals, confirmed via qPCR and shotgun metagenomics, was targeted and enriched with the MYBaits® kit from MYcroarray using RNA probes at 3-5x tiling density and the following bait design: *Y. pestis* CO92 was set as a reference genome and most of the highly repetitive, as well as ribosomal, DNA regions were excluded from the design. In total, 215,512 RNA (100bp) baits were created based on the chromosomal assembly of *Y. pestis* CO92 (NCBI accession number NC_003143), 4388 based on plasmid pMT1 (NC_003134), 3197 based on plasmid pCD1 (NC_003131) and 363 based on plasmid pPCP (NC_003132), with 5x tiling density over 20 bp intervals. Some of the baits were moved up- or downstream to achieve a smooth overall coverage, particularly preceding and following excluded regions and specific SNP positions.

Prior to target enrichment, double stranded single indexed libraries were concentrated to 7 µl using a SpeedVac. All libraries were enriched individually according to the manufacturer's instructions of a modified version of the MYBaits® kit (3.01) using half-volume aliquots of MYBaits® baits and Block 1-3 (Conditions detailed in Table S1).

Subsequently, captured DNA was amplified in duplicate 50 µl PCR reactions using Herculase II Fusion Polymerase, 5x Herculase II Reaction Buffer, dNTPs (250 µM), primers IS5/IS6 (0.3 µM each) and 15 µl of template under the following conditions: 98°C for 2 min, 14 cycles of 98°C for 20 s, 60°C for 30 sec and 72°C for 30 sec, and 72°C for 5 min. Amplified samples were purified with AmpureXP® beads and quantified on a Bioanalyzer 2100 expert chip and with a Qubit® ds High Sensitivity Assay.

High-throughput sequencing (125bp PE) of purified libraries was performed on an Illumina HiSeq 2500 system at the National Sequencing Centre (NSC) at the University of Oslo.

ARCHAEOLOGICAL INFORMATION

A. The site of San Procolo a Naturno (South Tyrol, Italy):

The site of San Procolo a Naturno is situated in the village of Naturns (South Tyrol, Italy). It is renowned for its small church, which dates to the 7th century, and its well-preserved frescos. The cemetery was excavated in 1985/86 and yielded graves dating between the 6th and the beginning of the 17th century CE. A total of 195 individuals were recovered from single and multiple burials, which were dated numismatically (9).

We screened 24 individuals recovered from multiple burials associated with the last phase of the cemetery, the beginning of the 17th century. This phase was linked with an epidemic outbreak dated to 1636, which was speculated to have been caused by typhus (9). However, our metagenomic analysis showed that there were no signs for the presence of Rickettsia prowazekii (See Table S3). Instead, we discovered reads matching the plague bacterium Yersinia pestis, a pathogen which has also been associated with the 1636 epidemic in the past (10). Six individuals tested positive for the plague pathogen.

B. The site of Pestbacken (Sweden):

Sample PEB10/A975 was recovered from a skeleton excavated at the burial ground of Pestbacken (Holje, Jämshög parish, Blekinge County, Sweden). The site was affected by groundworks in connection with the construction of water and sewer pipes in 1932, during which human bones were discovered. Due to plans for the construction of an electrical power installation in 1952, the first archaeological excavation was carried out and led to the discovery of the three burials (11).

The second and main archaeological excavation took place in 2000 and covered two separate areas of the site. However, the site has yet to be fully investigated. The National Heritage Board, Archaeological Excavation Department, UV Syd carried out the excavation. The documented graves were not part of a traditional churchyard, but instead, seem to have been part of a burial ground set-up to deal with increased mortality rates in the region caused by a plague outbreak (12).

Ninety graves containing a total of 115 individuals were excavated in Pestbacken. The individuals were found in single, double and a few triple burials. The graves were arranged in proper, undisturbed rows. Children, women and men were interred in wooden coffins. The excavators expect the full undisturbed burial ground to have contained approx. 300-350 individuals. The skeletons are stored at the Lund University Historical Museum (LUHM inv. no. 31 185) (12).

The site was dated based on archaeological finds from the burials. Amongst the artefacts recovered in the graves were personal items such as toothpicks and small spoons used to remove earwax, as well as

necklaces with pearls. People were buried in their everyday clothes; details found on textile remains seem to indicate the burials date to the 18th century. Furthermore, coins were excavated together with the skeletons. All in all fifteen coins have been found at the site. They were minted from 1667 to 1710 and give clear-cut evidence of the dating. The site can be dated to the outbreak of plague in South Sweden in 1710-1711 (11).

Unfortunately, fires in the vicarage and in the county archives around 1800 have destroyed the records of the parish church from this period, so no written information of the deceased or exact dates of the outbreak can be found.

Six individuals from the site of Pestbacken, aptly named the "plague hill", were tested for the presence of *Y. pestis* aDNA. Only one individual, PEB10 (A975 G975) yielded positive results. The individual was a male, aged between 20-25 years at the time of his death. He was recovered from a single burial. His height was estimated at 170 cm. With regards to his health we have noted severe caries on the teeth of his upper jaw. We sampled caries free canines for aDNA extraction. Furthermore, marks on his teeth show that he was smoking a clay pipe, which resulted in tooth wear typically associated with long term pipe smoking (personal communication Arcini C. and Iregren E. 2019).

C. The site of Maist (Chechnya, Russia):

The tooth CHE1 stems from a female individual (522) (Fig. S5), which was excavated in Maist (Chechnya, Russia), an ancient Chechnyan capital in a high-mountainous region situated along the Georgian border, in 1962. The only available skeletal element is a skull, which is part of a cranial collection made up of 25 individuals approximately dated between the 16th and the 18th century CE, curated by author Leila Sharashidze. The skulls were collected from the crypts of Checheno-Ingushetia in 1963 by Leila Sharashidze. Further information on the excavation of 1962 or the original archaeologist is not available. The skull and the teeth have been treated with an ethanol solution of mastic making further attempts at dating difficult.

The cranium exhibited signs of cribra orbitalia and caries. The frontal bone exhibits signs of large healed trauma, potentially resulting from a non-deadly surgical procedure (personal communication Sharashidze L, 2019).

D. The site of Collalto Sabino (Italy):

Sample COL001 stems from the archaeological site of Collalto Sabino (Rieti, Lazio, Italy). Excavations in the area surrounding the Medieval church of S. Giovani in Fistola were started in 2012. We sampled a tooth from individual 16a, an adult male, which was found in the "Area 2" (sector III)

of the excavation area, located between the Medieval church and the perimeter of a monastery, which seems to have been in use until the mid-12th century according to historical sources(13).

Sources document the use of the site as a Church maintained burial place beyond the 12th century(13, 14). However, it is unknown how long the church itself remained in use. Individual 16a was recovered from a multiple burial, which it shared with another adult (female) and a sub-adult individual. A total of 63 skeletons were recovered from 46 burials at Collalto Sabino, of these 43 were adults and 20 subadults. Of note was the excavation of a mass burial (tomb 31, trench) close to "Area 2" that yielded the remains of ten adult individuals covered by a thick layer of lime. The absence of skeletal trauma and the presence of lime among the remains prompted speculation of an epidemic event at the site (13). The multiple burials are all situated in the uppermost layers of the site stratigraphy, just below the modern walking horizon, and therefore form the last and youngest layer of burials on site. The region surrounding the excavation site was known to have been affected by plague from the Black Death to the 17th century (14–16). Of particular note is an outbreak in 1363, which coincides chronologically with the site and the multiple burials(14), placing the strain to the second half of the 14th century.

REFERENCES

1. P. Brotherton, *et al.*, Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat. Commun.* **4**, 1764 (2013).

2. J. Dabney, *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15758–15763 (2013).

3. E. A. Bennett, *et al.*, Library construction for ancient genomics: single strand or double strand? *Biotechniques* **56**, 289–90, 292–6, 298, passim (2014).

4. V. J. Schuenemann, *et al.*, Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of Yersinia pestis from victims of the Black Death. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E746–52 (2011).

5. W. Haak, *et al.*, Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science* **310**, 1016–1018 (2005).

6. O. Handt, M. Krings, R. H. Ward, S. Pääbo, The retrieval of ancient human DNA sequences. *Am. J. Hum. Genet.* **59**, 368–376 (1996).

7. A. Namouchi, *et al.*, Integrative approach using Yersinia pestis genomes to revisit the historical landscape of plague during the Medieval Period. *Proc. Natl. Acad. Sci. U. S. A.* (2018) https:/doi.org/10.1073/pnas.1812865115.

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

9. S. Renhart, Antropologia: sulle orme dell'uomo di San Procolo. *S. Procolo a Naturno-La Storia attraverso gli Scavi, Mostra straordinaria a cura del Museo Archeologico Provinciale di Castel Tirolo* **12**, 153–190 (1991).

10. W. Kofler, H. Nothdurfter, U. Rupp, *San Procolo a Naturno* (Tappeiner, 2000).

11. C. Arcini, B. Jacobsson, B. Persson, *Pestbacken* (Riksantikvarieämbetet, 2006) (May 4, 2018).

12. C. Arcini, A plague on all your houses. Famine, tobacco and pestilence in 18th century Sweden. *Current World Archaeology*, 26–31 (2011).

13. Alvino G, Colasanti B, Porzi L, Rotellini A, Nuovi rinvenimenti nella valle del Turano: le recenti indagini presso la chiesa di S. Giovanni in Fistola (Collalto Sabino) in *Atti Dei Convegni, Incontri Di Studi Sul Lazio E La Sabina. Lavori E Studi Della Soprintendenza Archeologica, Belle Arti E Paesaggio per L'area Metropolitana Di Roma, La Provincia Di Viterbo E l'Etruria Meridionale. Lazio E Sabinia 10, Atti Del Convegno (Roma 4-6 Giugno 2013).*, Russo Tagliente Aghini G Mari, Ed. (Edizioni Quasar, 2016).

14. É. Hubert, *L'"Incastellamento" en Italie centrale: Pouvoirs, territoire et peuplement dans la vallée du Turano au Moyen Âge* (École française de Rome, 2002).

15. G. Mosca, *Posta nell'Alta Valle del Velino: raccolta di dati per una storia civile e religiosa* (Amministrazione provinciale, 1998).

16. P. Carrozzoni, *Collepiccolo e la valle del Turano* (Il Velino, 1986).

Figure S1: Multiple burial from which individual COL001 was recovered (Collalto Sabino).

Photographs by and courtesy of Paola Zaio and Mauro Rubini.

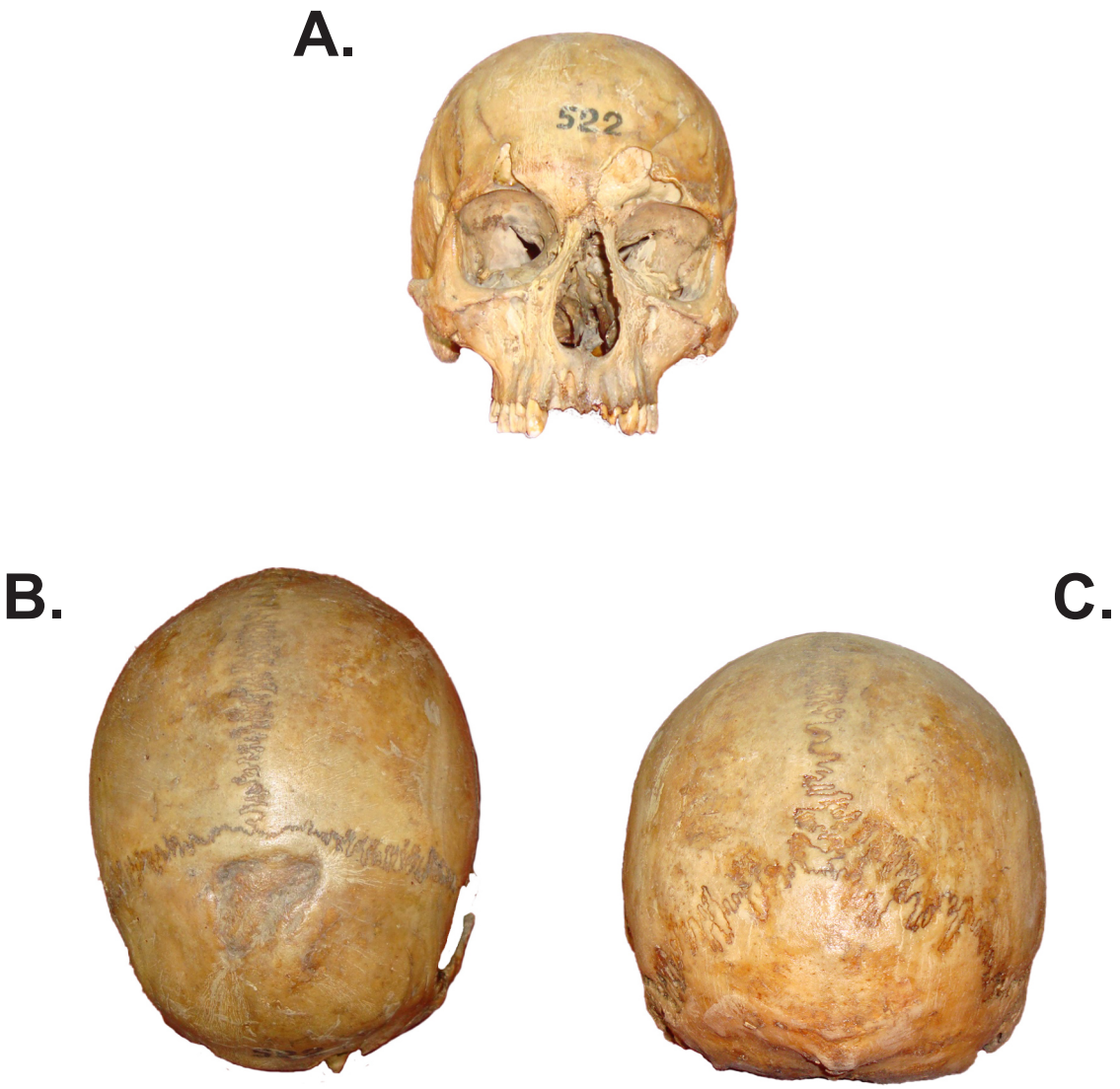

Figure S2: Skull sampled from Maist (Individual 522, sample CHE1). (A) Anterior view of the cranium. (B) Superior view of the cranium with visible healed trauma on the o*s frontale*. (C) Posterior view of the cranium.

Photographs by and courtesy of Shorena Laliashvili (Laboratory of Anthropological Studies of the Institute of History and Ethnology, Iv. Javakhishvili Tbilisi State University)

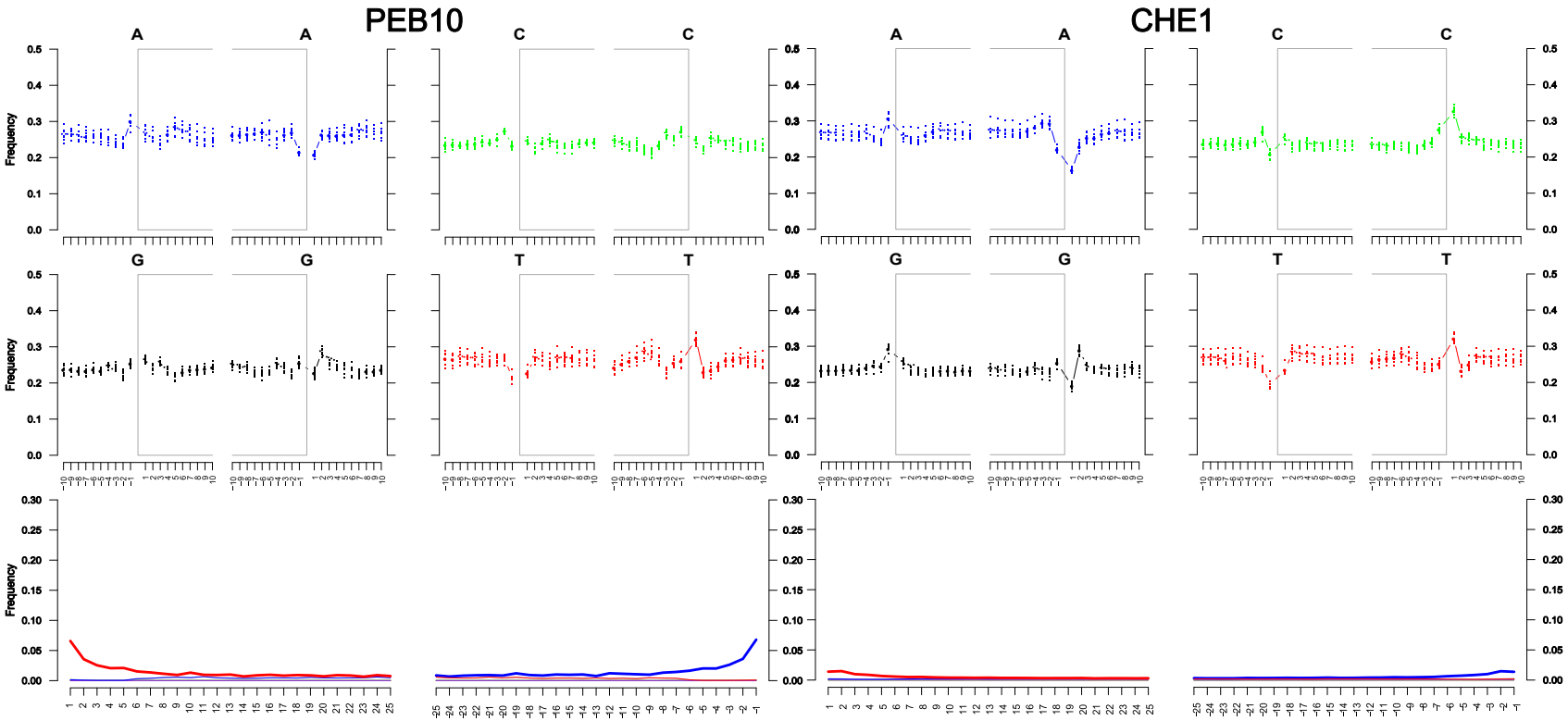

Figure S3: Ancient DNA damage plots for the mapping of *Y. pestis* reads of samples CHE1 and PEB10 against the reference genome *Y. pestis* CO92. The four upper mini-plots show the base frequency outside and within the mapped reads (the open grey box corresponds to the read). The bottom plots show position and frequency of typical aDNA nucleotide substitutions from C to T and G to A, which are reported in red and blue, respectively, and were calculated and visualized using MapDamage2.0 (Jónsson et al. 2013).

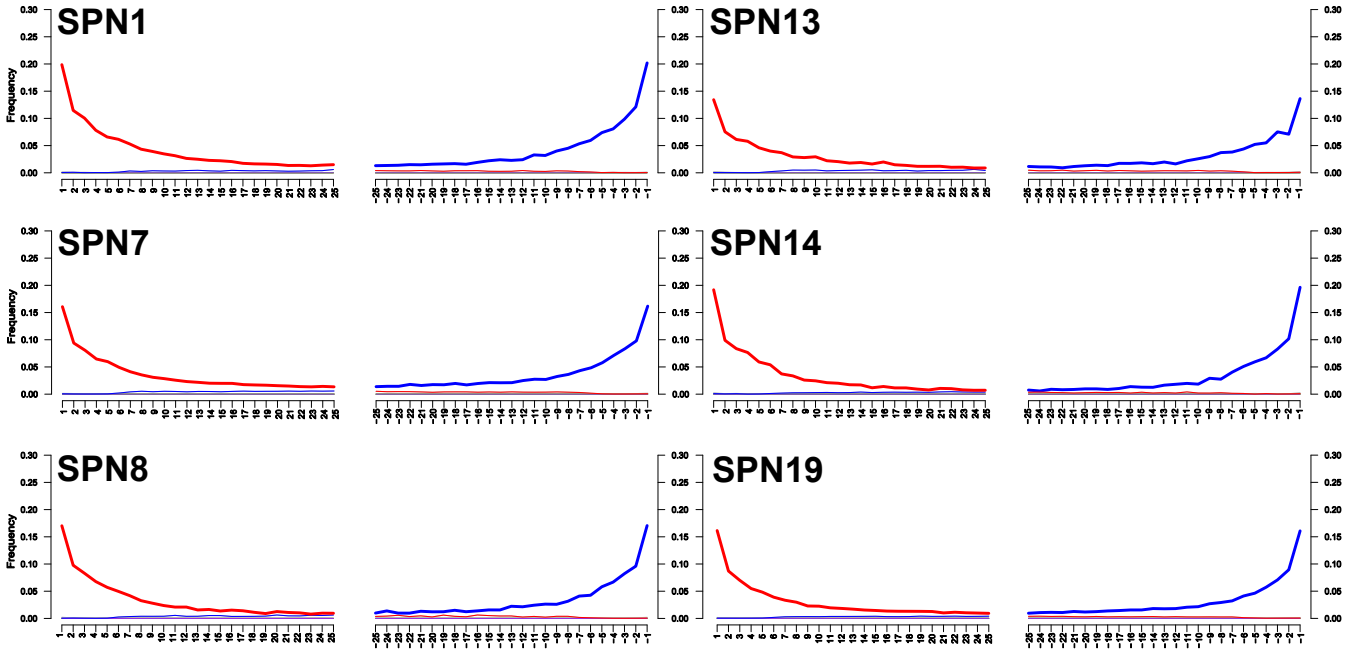

Figure S4: Ancient DNA damage plots for the mapping of *Y. pestis* reads of SPN samples against the reference genome *Y. pestis* CO92. Plots show position and frequency of typical aDNA nucleotide substitutions from C to T and G to A, which are reported in red and blue, respectively, and were calculated and visualized using MapDamage2.0 (Jónsson et al. 2013).

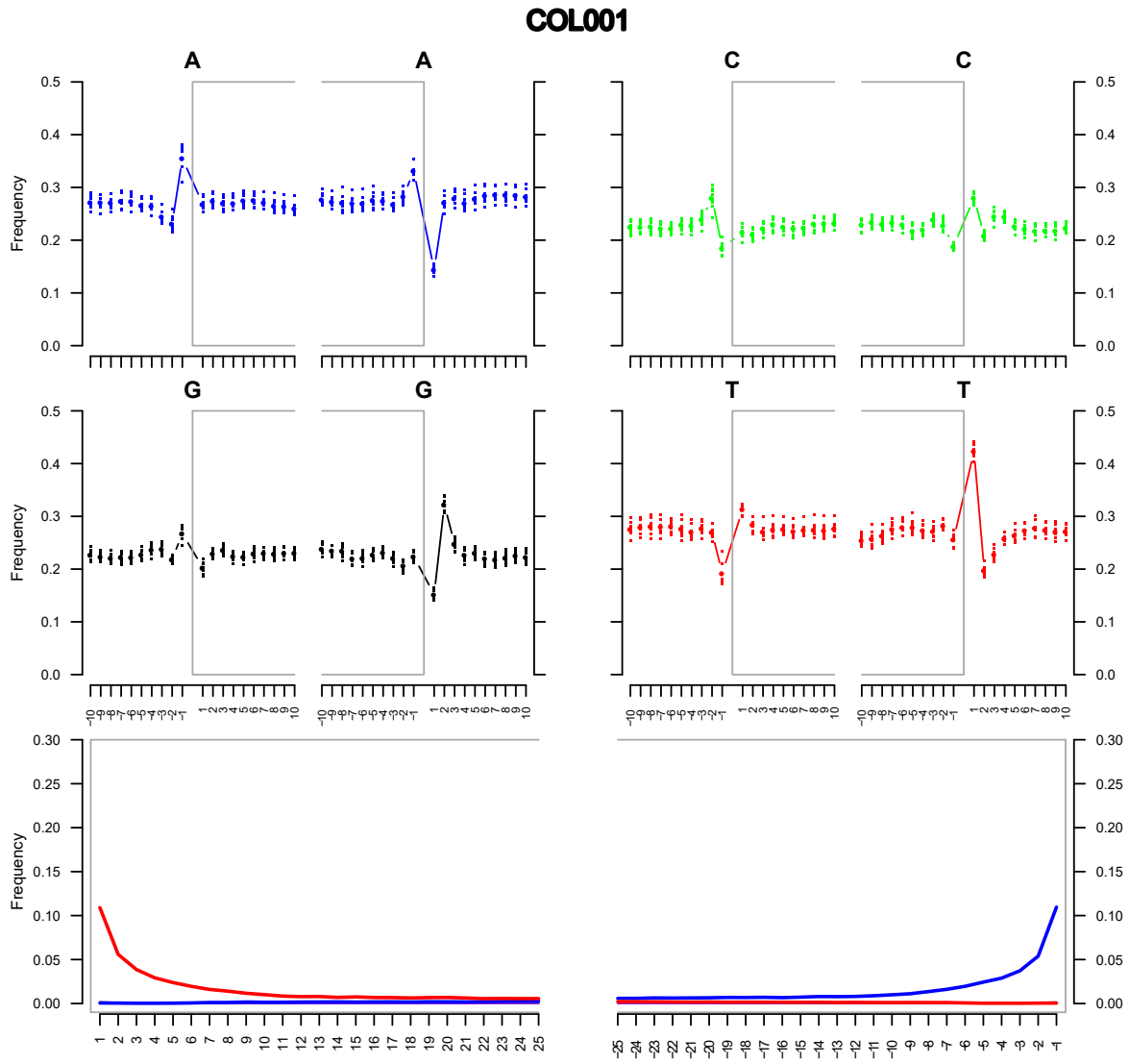

Figure S5: Ancient DNA damage plots for the mapping of *Y. pestis* reads of COL001 against the reference genome *Y. pestis* CO92. The four upper mini-plots show the base frequency outside and within the mapped reads (the open grey box corresponds to the read). The bottom plots show position and frequency of typical aDNA nucleotide substitutions from C to T and G to A, which are reported in red and blue, respectively, and were calculated and visualized using MapDamage2.0 (Jónsson et al. 2013).

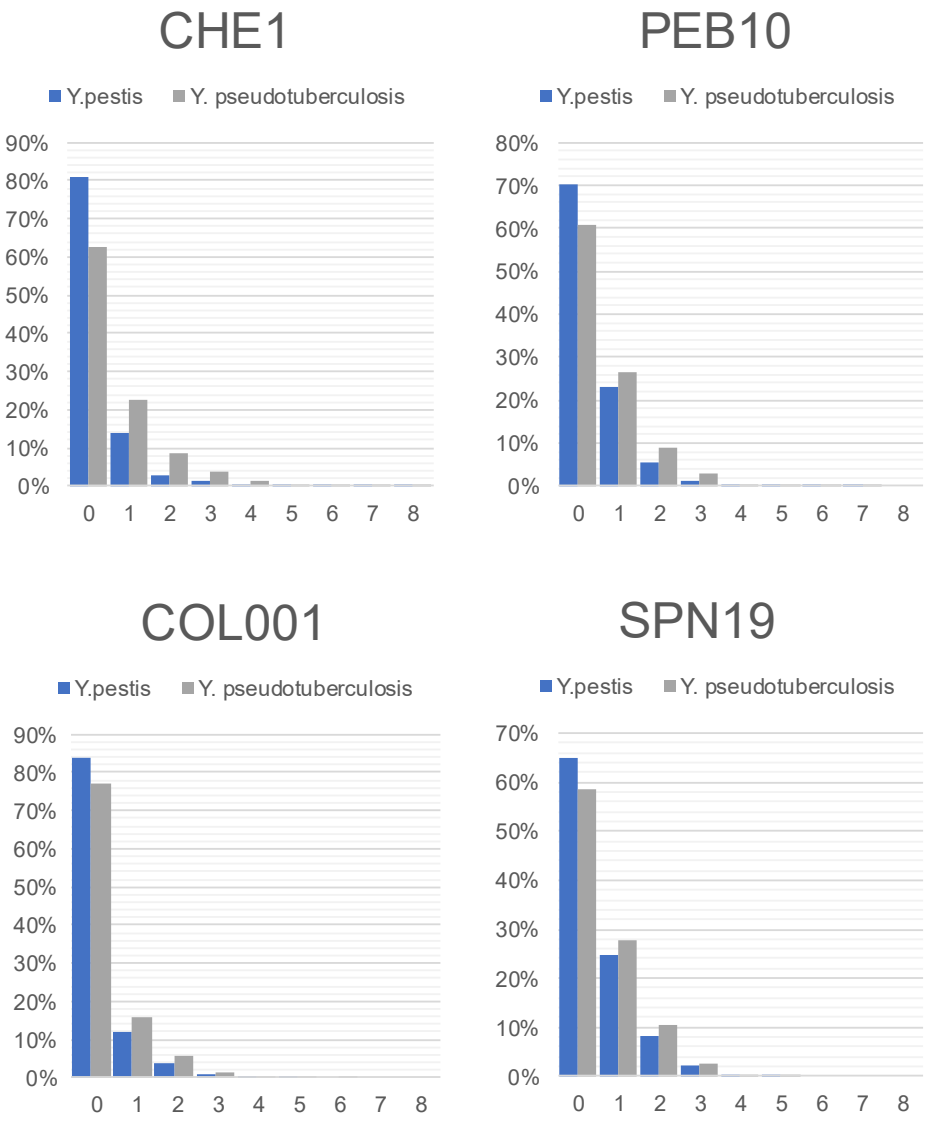

Figure S6: Edit distance plots for genomes represented in the phylogeny (Figure 3). For the *Y. pestis* mappings we used the reference strain CO92 and for the *Y. pseudotuberculosis* mapping the reference strain IP32953. Edit distances were computed exclusively for the chromosomal assemblies and without taking into account rRNA genes. The plots show the mean percentages per edit distance across the computed intervals.

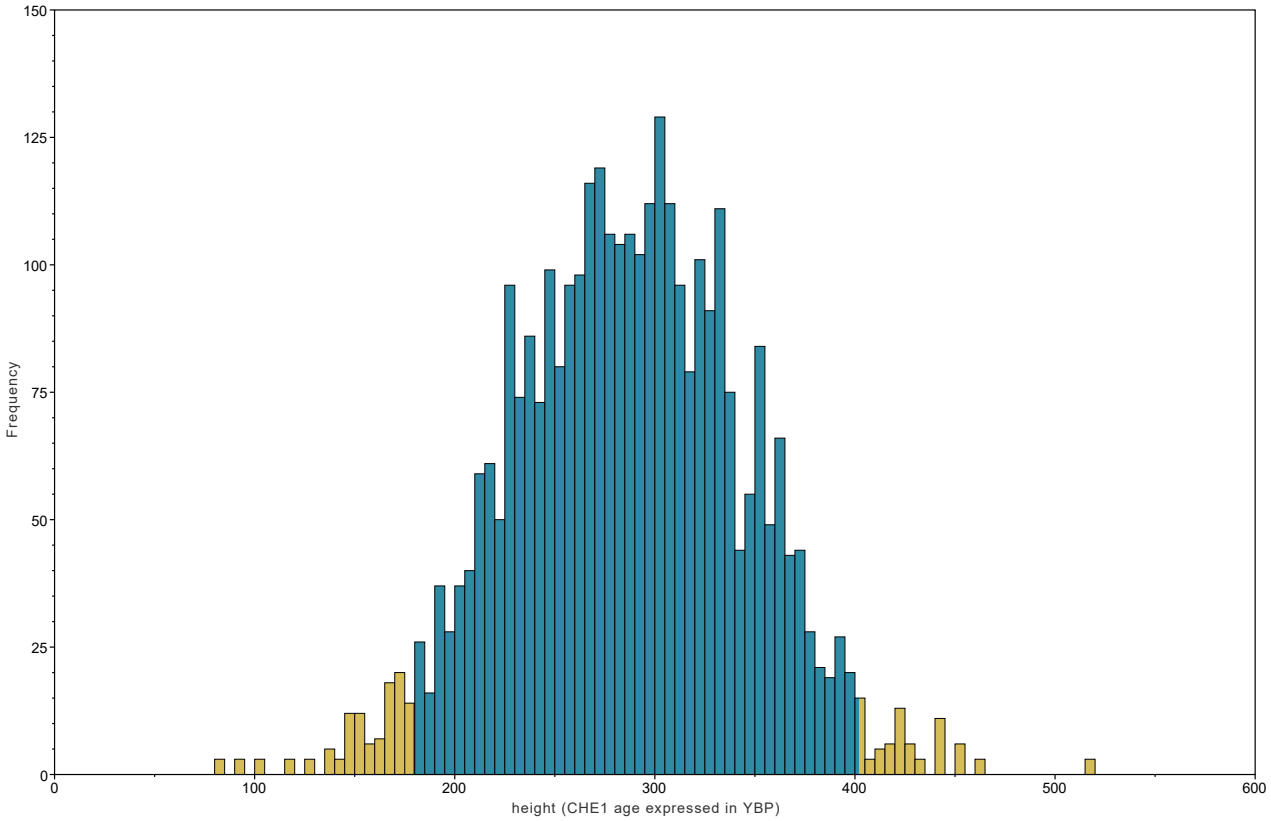

Figure S7: Posterior distribution of CHE1 height. The height expresses the age of the CHE1 sample in number of years before present (YBP). 2015 is the age of the most recent isolate.

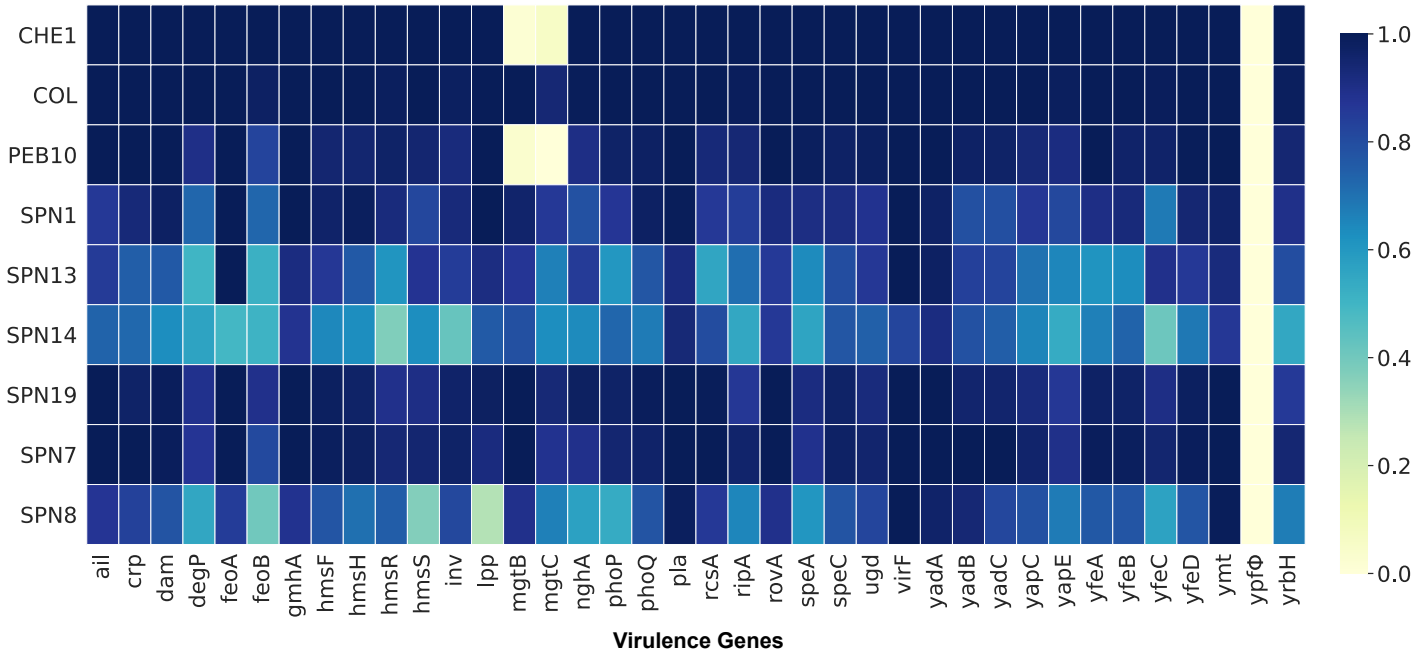

Figure S8: Virulence profile for all novel genomes from this study. The color scale indicates coverage within the gene intervals from 0 (pale yellow; no coverage) to 1 (dark blue; full coverage) based on the mappings to the reference genome Yersinia pestis CO92. The heatmap was generated using seaborn (82) in python3.