Report

Current Biology

The Stone Age Plague and Its Persistence in Eurasia

Highlights

- Six Late Neolithic–Early Bronze Age European *Y. pestis* genomes were reconstructed
- All Late Neolithic and Early Bronze Age *Y. pestis* form a single phylogenetic branch

Authors

Aida Andrades Valtueña, Alissa Mittnik, Felix M. Key, ..., Philipp W. Stockhammer, Alexander Herbig, Johannes Krause

Correspondence

herbig@shh.mpg.de (A.H.), krause@shh.mpg.de (J.K.)

In Brief

Andrades Valtueña et al. present the first six European *Y. pestis* genomes dating from the Late Neolithic and the Early Bronze Age. These data suggest that *Y. pestis* entered Europe during a human migration around 4800 BP, persisted in Europe, and traveled back to Central Eurasia.

Andrades Valtueña et al., 2017, Current Biology *27*, 3683–3691
December 4, 2017
https://doi.org/10.1016/j.cub.2017.10.025

CellPress

Current Biology
Report

The Stone Age Plague and Its Persistence in Eurasia

Aida Andrades Valtueña,[1] Alissa Mittnik,[1,2] Felix M. Key,[1] Wolfgang Haak,[1,3] Raili Allmäe,[4] Andrej Belinskij,[5] Mantas Daubaras,[6] Michal Feldman,[1,2] Rimantas Jankauskas,[7] Ivor Janković,[8,9] Ken Massy,[10,11] Mario Novak,[8] Saskia Pfrengle,[2] Sabine Reinhold,[12] Mario Šlaus,[13] Maria A. Spyrou,[1,2] Anna Szécsényi-Nagy,[14] Mari Tõrv,[15] Svend Hansen,[12] Kirsten I. Bos,[1,2] Philipp W. Stockhammer,[1,10] Alexander Herbig,[1,2,*] and Johannes Krause[1,2,16,*]
[1]Max Planck Institute for the Science of Human History, Jena, Germany
[2]Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of Tübingen, Tübingen, Germany
[3]School of Biological Sciences, The University of Adelaide, Adelaide SA 5005, South Australia, Australia
[4]Archaeological Research Collection, Tallinn University, Tallinn, Estonia
[5]"Nasledie" Cultural Heritage Unit, Stavropol, Russia
[6]Department of Archaeology, Lithuanian Institute of History, Vilnius, Lithuania
[7]Department of Anatomy, Histology and Anthropology, Vilnius University, Vilnius, Lithuania
[8]Institute for Anthropological Research, Zagreb, Croatia
[9]Department of Anthropology, University of Wyoming, Laramie, WY, USA
[10]Institute for Pre- and Protohistoric Archaeology and Archaeology of the Roman Provinces, Ludwig-Maximilians-University Munich, Munich, Germany
[11]Heidelberg Academy of Sciences, Heidelberg, Germany
[12]Eurasia Department, German Archaeological Institute, Berlin, Germany
[13]Anthropological Center, Croatian Academy of Sciences and Arts, Zagreb, Croatia
[14]Institute of Archaeology, Research Centre for the Humanities, Hungarian Academy of Sciences, Budapest 1097, Hungary
[15]Department of Archaeology, Institute of History and Archaeology, University of Tartu, Tartu, Estonia
[16]Lead Contact
*Correspondence: herbig@shh.mpg.de (A.H.), krause@shh.mpg.de (J.K.)
https://doi.org/10.1016/j.cub.2017.10.025

SUMMARY

Yersinia pestis, the etiologic agent of plague, is a bacterium associated with wild rodents and their fleas. Historically it was responsible for three pandemics: the Plague of Justinian in the 6th century AD, which persisted until the 8th century [1]; the renowned Black Death of the 14th century [2, 3], with recurrent outbreaks until the 18th century [4]; and the most recent 19th century pandemic, in which *Y. pestis* spread worldwide [5] and became endemic in several regions [6]. The discovery of molecular signatures of *Y. pestis* in prehistoric Eurasian individuals and two genomes from Southern Siberia suggest that *Y. pestis* caused some form of disease in humans prior to the first historically documented pandemic [7]. Here, we present six new European *Y. pestis* genomes spanning the Late Neolithic to the Bronze Age (LNBA; 4,800 to 3,700 calibrated years before present). This time period is characterized by major transformative cultural and social changes that led to cross-European networks of contact and exchange [8, 9]. We show that all known LNBA strains form a single putatively extinct clade in the *Y. pestis* phylogeny. Interpreting our data within the context of recent ancient human genomic evidence that suggests an increase in human mobility during the LNBA, we propose a possible scenario for the early spread of *Y. pestis*: the pathogen may have entered Europe from Central Eurasia following an expansion of people from the steppe, persisted within Europe until the mid-Bronze Age, and moved back toward Central Eurasia in parallel with human populations.

RESULTS

Screening

A total of 563 tooth and bone samples dating from the Late Neolithic to the Bronze Age (LNBA) from Russia (n = 122), Hungary and Croatia (n = 139), Lithuania (n = 27), Estonia (n = 45), Latvia (n = 10), and Germany (n = 220) were screened for *Y. pestis* by mapping reads ranging from 700,000 to 21,000,000 against a multi-fasta reference of 12 different *Yersinia* (Table 1).

To evaluate whether an individual was potentially *Y. pestis*-positive, we calculated a score based on the number of specific reads mapping to *Y. pestis* compared to other *Yersinia* (see STAR Methods). Individuals with a positive score were deemed potential candidates. Those with scores > 0.005 and reads mapping to all three *Y. pestis* plasmids were considered "strong" positives. We identified five "strong" candidates: one individual from Rasshevatskiy (RK1001; North Caucasus, Russia), one from Gyvakarai (Gyvakarai1; Lithuania), one from Kunila (Kunila II; Estonia), and two from Augsburg, Germany (Haunstetten, Unterer Talweg 85 Feature 1343 [1343UnTal85]; Haunstetten, Postillionstrasse Feature 6 [6Post]). One individual from Beli Manastir-Popova zemlja (GEN72; Croatia) did not pass the "strong" candidate threshold but was included by virtue of having the highest number of reads mapping to the *Y. pestis* chromosome and plasmids (chromosome = 993, pCD1 = 243,

Table 1. Genomes from the NCBI RefSeq/Nucleotide Database, Used in the Multi-species Reference Panel for Screening for *Y. pestis* aDNA

Species Name	Strain	NCBI Accession Number
Y. pestis	CO92	NC_003143.1
Y. pseudotuberculosis	IP 32953	NC_006155.1
Y. enterocolitca	subsp. enterocolitica 8081	NC_008800.1
Y. aldovae	ATCC 35236	NZ_ACCB01000210.1
Y. bercovieri	ATCC 43970	NZ_AALC02000229.1
Y. frederiksenii	ATCC 33641	NZ_AALE02000161.1
Y. intermedia	ATCC 29909	NZ_AALF02000123.1
Y. kristensenii	ATCC 33638	NZ_ACCA01000153.1
Y. mollaretii	ATCC 43969	NZ_AALD02000179.1
Y. rohdei	ATCC 43380	NZ_ACCD01000141.1
Y. ruckeri	ATCC 29473	NZ_ACCC01000174.1

pMT1 = 111, pPCP1 = 22). For additional archaeological information, see Table 2 and STAR Methods.

Genome Reconstruction

"Strong" positive individuals were shotgun sequenced to a depth of 379,155,741–1,529,935,532 reads. RK1001 and GEN72 were further enriched for *Y. pestis* DNA using in-solution capture (see STAR Methods). After mapping to the reference genome (*Y. pestis* CO92, NC_003143.1), we reconstructed genomes for all six candidates with a mean coverage between 3- and 12-fold, with 86%–94% of the reference covered 1-fold (Table 2, Figure S2). The reads were independently mapped to the three *Y. pestis* CO92 plasmids yielding mean coverages of 7- to 24-fold (pCD1), 3- to 14-fold (pMT1), and 18- to 43-fold (pPCP1; Table S1, Figure S2).

To authenticate the ancient origin of the bacterial genomes, we evaluated terminal deamination damage common to ancient DNA [14]. Our samples presented typical damage profiles similar to the corresponding associated human DNA (Figure S1), and multi-strain infection was not observed (Figure S1).

Phylogeny and Dating

To assess the phylogenetic position of the reconstructed genomes in comparison to modern and ancient *Y. pestis* genomes (see STAR Methods), we computed neighbor joining (Figure S3A), maximum parsimony (Figure S3B), and maximum likelihood (Figures 1 and S3C) trees. Our samples form a clade with the previously reported RISE509 and RISE505 strains [7], with a bootstrap support > 95% for all methods.

The branching point of the LNBA genomes and all other strains represents the most recent common ancestor (MRCA) of all currently available *Y. pestis* genomes, which was "tip-dated" using BEAST [15] to 6,078 years (95% highest posterior density interval: 5,036–7,494 years), in agreement with previous estimates [7]. The time to the MRCA of *Y. pestis* and *Y. pseudotuberculosis* strain IP32953 was estimated at 28,258 years (95% highest posterior density interval: 13,200–44,631 years). The maximum clade credibility tree (Figure S4A) supports the same topology as the methods described above, with high statistical support of the LNBA branching point.

Genetic Makeup

We identified 423 single-nucleotide polymorphisms (SNPs) on the LNBA branch (Data S1, sheets 1–4), including strain-specific and shared SNPs, of which 114 are synonymous and 202 are non-synonymous (see STAR Methods). The LNBA genomes share five SNPs (Data S1, sheet 2).

The percent coverage was calculated for genes related to virulence, flea transmission, colonization, and dissemination (Figure 1B). The YpfΦ prophage [16], integrated only into the chromosomes of the 1.ORI strains responsible for the third pandemic [17], was absent in all LNBA genomes. Additionally, *yapC*, possibly involved in the adhesion to mammalian cells, autoagglutination, and biofilm formation [18], was lost in the three youngest LNBA strains (1343UnTal85, Post6, and RISE505).

The only plasmid virulence factor missing in the LNBA strains is *ymt* (Figures 1B and S2). *ymt* codes for the *Yersinia* murine toxin, an important virulence factor in flea transmission [19, 20]. Expression of *ymt* protects against toxic blood digestion byproducts and permits colonization of the flea midgut [20]. Other plasmid virulence factors such as *pla* and *caf1*, absent in *Y. pseudotuberculosis*, were already present in the LNBA *Y. pestis* strains.

Urease D (*ureD*) plays an important role in flea transmission. *ureD* expression causes a toxic oral reaction killing 30%–40% [21] of infected fleas. *ureD* is a pseudogene in *Y. pestis* due to a frameshift mutation [22]. Close inspection of this gene revealed that the frameshift is not present, indicating that this gene was functional in the LNBA strains and possibly making them as toxic to fleas as their ancestor *Y. pseudotuberculosis*.

Furthermore, *Y. pseudotuberculosis*-specific regions that have been lost in *Y. pestis* were still present in the LNBA strains (Data S1, sheet 5). We also observed genome decay in the LNBA clade mostly affecting flagellin genes and membrane proteins (Figure S2; Data S1, sheet 5).

DISCUSSION

The prehistoric genomes presented here are the first to reveal *Y. pestis* diversity in the European LNBA. This complements contemporary *Y. pestis* genomes from Bronze Age individuals recovered from Southern Siberia [7] and offers higher resolution to evaluate the evolution and dissemination of prehistoric plague strains. All LNBA genomes, including those previously reconstructed from Southern Siberia [7], form a distinct clade. The strains RISE509 [7] and RK1001 occupy the most basal position of all *Y. pestis* genomes sequenced to date, formally tested with CONSEL [23] (Figure S4B). These data are compatible with two scenarios. In scenario 1, plague was introduced multiple times to Europe from a common reservoir between 5,000 and 3,000 BP. The bacterium spread to Europe from a source most likely located in Central Eurasia at least four times during a period of over 1,000 years: once to Lithuania and Croatia, once to Estonia, and twice to Germany. In this model, the phylogeny of the LNBA lineages results exclusively from their temporal relationship. In scenario 2, plague entered Europe from Central Eurasia once during the Neolithic. A reservoir was established within or close

Table 2. Statistics of the *Y. pestis* Genome Reconstruction

Individual	Tissue Sampled	Site	Country	Radiocarbon Date (14dC)	Dating (Median cal BP)	2-Sigma Interval [cal BP]; [cal BC]	In-Solution Capture	Clipped, Merged, and Quality-Filtered Reads before Mapping	Unique Reads Mapping to *Y. pestis* Reference	Endogenous DNA (%)	Mean Coverage	Coverage (%) ≥1×	Coverage (%) ≥3×	Publication
RK1001	Tooth	Rasshevatskiy	Russia	4,171 ± 22	4720	[4,828–4,622]; [2,879–2,673]	no	1,529,935,532	119,540	0.01	1.0213	58.11	10.87	This study
RK1001	Tooth	Rasshevatskiy	Russia	4,171 ± 22	4720	[4,828–4,622]; [2,879–2,673]	yes	303,148,884	383,900	0.85	3.3984	82.83	55.3	This study
RK1001	Tooth	Rasshevatskiy	Russia	4,171 ± 22	4720	[4,828–4,622]; [2,879–2,673]	combined shotgun/ capture	1,833,084,416	418,581	0.17	3.6816	86.16	59.63	This study
GEN72	Tooth	Beli Manastir-Popova zemlja	Croatia	4,176 ± 28	4721	[4,833–4,592]; [2,884–2,640]	yes	19,777,683	1,321,320	24.36	12.6549	91.65	86.61	[10]
Gyvakarai1	Tooth	Gyvakarai	Lithuania	4,030 ± 30	4485	[4,571–4,422]; [2,578–2,491]	no	1,021,452,137	473,207	0.05	5.2245	94.07	84.12	[11]
KunilaII	Tooth	Kunila	Estonia	3,960 ± 40	4427	[4,524–4,290]; [2,576–2,340]	no	379,155,741	546,243	0.16	5.5418	92.48	77.58	[12]
1343UnTal85	Tooth	Augsburg	Germany	3,819 ± 24	4203	[4,346–4,098]; [2,397–2,149]	no	1,174,989,269	1,165,435	0.14	10,5745	93.69	92.59	[13]
6Post	Tooth	Augsburg	Germany	3,574 ± 19	3873	[3,957–3,832]; [2,009–1,883]	no	419,717,299	598,030	0.17	5.3062	89.71	71.14	[13]

The radiocarbon dates were calibrated with Calib 7.1. calBP = calibrated years Before Present; cal BC = calibrated years Before Christ. All individuals were directly radiocarbon dated. See also Table S1.

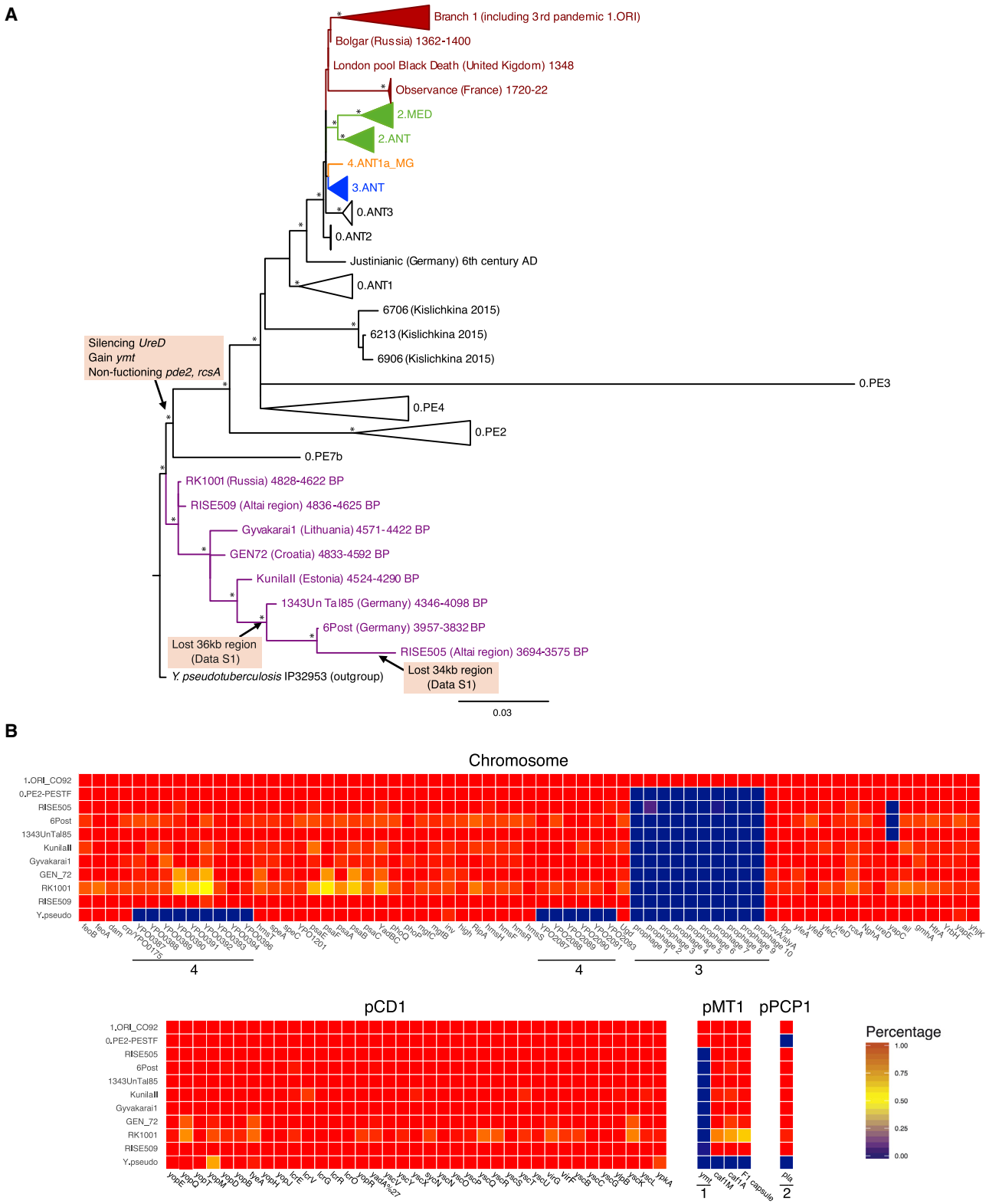

Figure 1. Maximum-Likelihood Tree and Percent Coverage Plot of Virulence Factors of *Yersinia pestis*

(A) Maximum-likelihood tree of all *Yersinia pestis* genomes, including 1,265 SNP positions with complete deletion. Nodes with support ≥ 95% are marked with an asterisk. The colors represent different branches in the *Y. pestis* phylogeny: branch 0 (black), branch 1 (red), branch 2 (green), branch 3 (blue), branch 4 (orange),

(legend continued on next page)

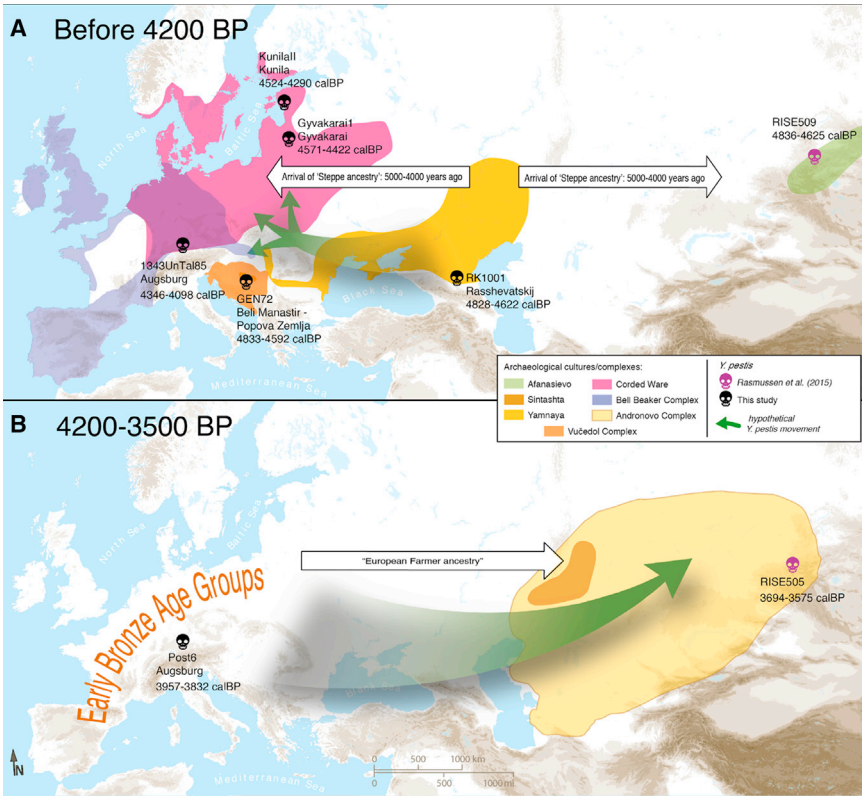

Figure 2. Map of Proposed *Yersinia pestis* Circulation throughout Eurasia

(A) Entrance of *Y. pestis* into Europe from Central Eurasia with the expansion of Yamnaya pastoralists around 4,800 years ago.

(B) Circulation of *Y. pestis* to Southern Siberia from Europe. Only complete genomes are shown.

to Europe from which it circulated, and ultimately it moved back to Central Eurasia during the Bronze Age (Figure 2).

With few genomes available, it is difficult to disentangle the two hypotheses; however, interpreting our data in the context of human genetics and archaeological data can offer some resolution. Ancient human genomic data point to a change in mobility and large-scale expansion of people from the Caspian-Pontic Steppe associated with the Yamnaya complex, both east and west starting around 4,800 BP. These people carried a distinct genetic component that is also seen in highly mobile groups associated with the Southern Siberian Afanasievo complex, the Yamnaya complex, and the Central and Eastern European Corded Ware complex [24]. In Central European individuals, it is first observed in the Corded Ware complex and then becomes part of the genetic composition of most subsequent and all modern-day European populations [24, 25].

Our earliest indication of plague in Europe is found in Croatia and the Baltic, coinciding with the arrival of "steppe ancestry" [24, 25] in human populations. The Baltic Late Neolithic *Y. pestis* genomes (Gyvakarai1 and KunilaII) were reconstructed from individuals associated with the Corded Ware complex. Along with the Croatian *Y. pestis* genome (Vučedol complex), these are derived from a common ancestor shared with the Yamnaya-derived RK1001 and Afanasievo-derived RISE509. This supports the notion of the pathogen spreading in the context of the large-scale expansion of steppe peoples from Central Eurasia to Eastern and Central Europe. Furthermore, human genomic analyses indicate that RISE509, Gyvakarai1, KunilaII, and GEN72 carry "steppe ancestry" [10, 11, 24]. Evidence for such long-distance contact is also present in the archaeological record. For example, the Gyvakarai1 burial is characterized by a specific inventory of grave items (e.g. hammer-headed pins) and distinct skeletal morphology that have no analogs in earlier local populations [26].

The younger Late Neolithic *Y. pestis* genomes from southern Germany are derived from the Baltic strains, and one of these is found in an individual associated with the Bell Beaker complex. Previous analyses have shown that Bell Beaker individuals from Germany also carry "steppe ancestry" [24, 25, 27]. This suggests that *Y. pestis* may have spread further southwest, analogous to the human "steppe" component. The youngest of the LNBA *Y. pestis* genomes (RISE505, Southern Siberia) associated with the Central Eurasian Andronovo complex, descends from the Central European strains, suggesting a spread into Southern Siberia. Interestingly, genome-wide human data show that individuals associated with the Sintashta, Srubnaya, and Andronovo cultural complexes in the Eurasian steppes (dating around 3,700–3,300 calibrated years before present) carried mixed ancestry of middle Neolithic European farmers and Bronze Age steppe people, suggesting a backflow of human genes from Europe to Central Eurasia [24]. Archaeologically, there seems to be a close connection between the Eastern European Abashevo cultural complex and Sintashta that might have included population shifts from west to east. In particular, the post-Sintashta Andronovo complex is an epoch of population shifts affecting all areas east of the Urals to the western borders of China, including populations with European origin [28, 29]. The steppe, a natural corridor connecting people and their livestock throughout Central and Western Eurasia, might have facilitated the spread of strains related to the European Early Bronze Age *Y. pestis* to Southern Siberia, where RISE505 was found. In our view, human genetic ancestry and admixture, in combination with the temporal series within the

and LNBA *Y. pestis* branch (purple). *Y. pseudotuberculosis*-specific SNPs were excluded from the tree for clarity of representation. In the light-colored boxes, discussed losses and gains of genomic regions and genes are indicated. Related to Figure S3.

(B) Percent coverage of virulence factors located on the *Yersinia pestis* chromosome and plasmids, plotted in R using the ggplot2 package. The numbers represent specific genes: (1) *ymt* gene, (2) *pla* gene, (3) filamentous prophage YpfΦ, (4) *Y. pestis*-specific genes. Related to Figure S2. See also Data S1.

LNBA *Y. pestis* branch, supports scenario 2. *Y. pestis* was possibly introduced to Europe from the steppe around 4,800 BP. Thereafter, a local reservoir was established within or in close proximity to Europe. The European *Y. pestis* strain was disseminated to Southern Siberia potentially through anthropogenic processes connected to the backflow of human genetic ancestry from Western Eurasia into Southern Siberia. The pathogen diversity mirrors the archaeological evidence, which indicates intensification of Eurasian trade networks from the beginning of the Bronze Age [8, 9].

Even though *Y. pestis* seems to have spread in patterns strikingly similar to human movements (Figure 2), the mode of transmission during this early phase of its evolution cannot be easily determined. Most contemporary cases of *Y. pestis* infection occur via a flea vector and stem from sylvatic rodent populations with resistance to the bacterium. Flea transmission is accomplished by one of two mechanisms [30]: blockage-dependent transmission [31] or early-phase transmission (EPT) [32]. In the former, *Y. pestis* obstructs the flea digestive system by producing a biofilm that blocks the flea's foregut within 1–2 weeks post-infection. This blockage prevents a blood meal from reaching the flea's midgut, and blood regurgitation during failed feeding sheds live bacteria into the host [31, 33]. The blockage-dependent transmission requires a functional *ymt* gene and *hms* locus, and non-functional *rcsA*, *pde2*, and *pde3* genes [34]. *ymt* protects *Y. pestis* within the digestive system of the flea, allowing colonization of the flea midgut. The *hms* locus is involved in biofilm formation, and *rcsA*, *pde2*, and *pde3* are biofilm downregulators. However, *Y. pestis* can be transmitted within the first 1–4 days after entering the flea prior to colonization of the midgut and biofilm formation [32, 35] (the EPT model). This model is currently less understood than blockage-dependent transmission but has been shown to be both biofilm [36] and *ymt* independent [37].

The genetic characteristics of the LNBA genomes (i.e., lack of *ymt*, functional *pde2*, and *rcsA*) were previously interpreted as evidence that early forms of *Y. pestis* were unable to cause blockage in the flea gut, thus suggesting that the bubonic form of the disease evolved later [7]. However, as none of these genes seem to be required for EPT, one cannot exclude that LNBA *Y. pestis* infections could have been acquired from fleas via this transmission mode. Under this model, transmission would have been less efficient since a functional UreD would have reduced the number of flea vectors by 30%–40%.

The presence of mammalian virulence-related genes such as *pla* and *caf1* indicates that LNBA *Y. pestis* was to some extent adapted to these hosts. The LNBA *pla* presents the ancestral I259 variant, shown to be less efficient in infiltrating the host [38, 39] than the derived T259 form [40]. Strains carrying the ancestral variant can cause pneumonic disease but are less efficient in colonizing other tissues [38]. This indicates that LNBA *Y. pestis* could have caused a pneumonic or less severe bubonic form. The genome decay we detected, affecting membrane and flagellar proteins possibly involved in interactions with the host's immune system, could indicate adaptation to new hosts or pathogenic lifestyles [41].

Modern plague is a rodent-adapted disease, in which commensal species such as *Rattus rattus* and their fleas play a central role as disease vectors for humans [42]. Although rodent-flea transmission is compatible with the genomic makeup of the LNBA strains, disease dynamics may have differed in the past. The Neolithic is considered to be a time period in which new diseases were introduced into human groups during transition from a mobile to sedentary lifestyle. Adoption of agriculture and increased population density are thought to have acted synergistically to change the disease landscape [43]. Whether commensal rodent populations were large enough to function as a plague reservoir during human migrations at this time is unknown. In Central Eurasian Bronze Age cultures, agriculture (i.e., large-scale food storage) was mostly absent [44]. However, contact between sylvatic rodents in the steppe, pastoralists, and their herds might have been frequent in these environments. Alternative models of transmission involving different host species, perhaps even humans or their livestock, might carry some traction, as the ancient disease may have behaved differently from what we know today.

The LNBA was a time of increased mobility and cultural change. The threat of *Y. pestis* infections may have been one of the causes for this increased mobility [7]. Further sampling of skeletal material could provide much-needed details about the range and frequency of *Y. pestis* infections during this transformative period. Presence of the disease in Europe could have played a role in the processes that led to the genetic turnover observed in European human populations, who may have harbored different levels of immunity against this disease. Testing these hypotheses will require extensive assessment of both human and *Y. pestis* genomes from the steppes and from Europe before and after the steppe migration.

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Ethical approvals
  - Description of samples and archaeological sites
- METHOD DETAILS
  - Sampling and extraction
  - Shotgun screening sequencing
  - *In silico* screening
  - Deep shotgun sequencing
  - *Y. pestis* in-solution capture
  - Genome reconstruction and authentication
  - Individual sample treatment due to laboratory preparation and sequencing strategies
  - SNP calling, heterozygosity, and phylogenetic analysis
  - Dating analysis
  - SNP effect analysis and virulence factors analysis
  - Indel analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - *In silico* screening
  - Phylogenetic analysis
  - Tree topology test
  - Molecular clock test
  - Dating analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, one table, and one dataset and can be found with this article online at https://doi.org/10.1016/j.cub.2017.10.025.

AUTHOR CONTRIBUTIONS

J.K., A.H., and A.A.V. conceived the study. K.M., R.A., M.D., R.J., M.T., P.W.S., A.B., I.J., M.N., S.R., M.S., A.S.-N., and S.H. provided samples and performed archaeological assessment. A.M., S.P., M.F., A.A.V., and A.S.-N. performed laboratory work. A.A.V., A.H., M.A.S., F.M.K., and J.K. analyzed the data. A.A.V., A.H., J.K., P.W.S., K.I.B., W.H., and A.M. wrote the manuscript with contributions from all authors. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Corina Knipper, Ernst Pernicka, Stephanie Metz, Fabian Wittenborn, Stephan Schiffels, Joris Peters, Michaela Harbeck, and Archaeogenetics Department members of the MPI-SHH for helpful discussion. We thank Annette Günzel for graphical support. We thank Isil Kucukkalipci, Antje Wissgott, Marta Burri, and Franziska Göhringer for technical support in the lab. We thank Joachim Wahl, Josip Burmaz and Dženi Los, and Gunita Zatina and Andrejs Vasks for kindly providing the Althausen, Croatian, and Latvian samples, respectively. We thank Natalia Berezina and Julia Gresky for the anthropological assessment of RK1001. We thank James A. Fellows Yates for proofreading. The Heidelberg Academy of Science financed the genetic and archaeological research on human individuals from the Augsburg region within the project WIN Kolleg: "Times of Upheaval: Changes of Society and Landscape at the Beginning of the Bronze Age." This work was financially supported by the Max Planck Society, European Research Council starting grant APGREID (to J.K.), and Croatian Science Foundation grant 1450 (to M.N. and I.J.).

Received: May 16, 2017
Revised: July 31, 2017
Accepted: October 9, 2017
Published: November 22, 2017

REFERENCES

1. Russell, J.C. (1968). That earlier plague. Demography *5*, 174–184.
2. Zietz, B.P., and Dunkelberg, H. (2004). The history of the plague and the research on the causative agent *Yersinia pestis*. Int. J. Hyg. Environ. Health *207*, 165–178.
3. Benedictow, O.J. (2004). The Black Death, 1346–1353: The Complete History (Boydell Press).
4. Cohn, S.K., Jr. (2008). Epidemiology of the Black Death and successive waves of plague. Med. Hist. Suppl. *2008*, 74–100.
5. Stenseth, N.C., Atshabar, B.B., Begon, M., Belmain, S.R., Bertherat, E., Carniel, E., Gage, K.L., Leirs, H., and Rahalison, L. (2008). Plague: past, present, and future. PLoS Med. *5*, e3.
6. World Health Organization (2017). Plague fact sheet. http://www.who.int/mediacentre/factsheets/fs267/en/.
7. 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.
8. Vandkilde, H. (2016). Bronzization: the Bronze Age as pre-modern globalization. Praehist. Z. *91*, 103–123.
9. Hansen, S. (2014). The 4th Millennium: a watershed in European prehistory. In Western Anatolia before Troy: Proto-Urbanisation in the 4th Millennium BC?, B. Horejs, and M. Mehofer, eds. (Vienna: Österreichische Akademie der Wissenschaften), pp. 243–260.
10. Mathieson, I., Roodenberg, S.A., Posth, C., Szécsényi-Nagy, A., Rohland, N., Mallick, S., Olade, I., Broomandkhoshbacht, N., Cheronet, O., Fernandes, D., et al. (2017). The Genomic History Of Southeastern Europe. bioRxiv. https://doi.org/10.1101/135616.
11. Mittnik, A., Wang, C.-C., Pfrengle, S., Daubaras, M., Zariņa, G., Hallgren, F., Allmäe, R., Khartanovich, V., Moiseyev, V., Furtwängler, A., et al. (2017). The Genetic History of Northern Europe. bioRxiv. https://doi.org/10.1101/113241.
12. Kriiska, A., Lõugas, L., Lõhmus, M., Mannermaa, K., and Johanson, K. (2007). New AMS dates from Estonian Stone Age burials sites. Estonian J. Archaeol. *11*, 83–121.
13. Stockhammer, P.W., Massy, K., Knipper, C., Friedrich, R., Kromer, B., Lindauer, S., Radosavljević, J., Wittenborn, F., and Krause, J. (2015). Rewriting the Central European Early Bronze Age Chronology: Evidence from Large-Scale Radiocarbon Dating. PLoS ONE *10*, e0139705.
14. Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. Proc. Natl. Acad. Sci. USA *104*, 14616–14621.
15. Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. *29*, 1969–1973.
16. Derbise, A., Chenal-Francisque, V., Pouillot, F., Fayolle, C., Prévost, M.-C., Médigue, C., Hinnebusch, B.J., and Carniel, E. (2007). A horizontally acquired filamentous phage contributes to the pathogenicity of the plague bacillus. Mol. Microbiol. *63*, 1145–1157.
17. Derbise, A., and Carniel, E. (2014). YpfΦ: a filamentous phage acquired by *Yersinia pestis*. Front. Microbiol. *5*, 701.
18. 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.
19. Hinnebusch, J., Cherepanov, P., Du, Y., Rudolph, A., Dixon, J.D., Schwan, T., and Forsberg, A. (2000). Murine toxin of *Yersinia pestis* shows phospholipase D activity but is not required for virulence in mice. Int. J. Med. Microbiol. *290*, 483–487.
20. Hinnebusch, B.J., Rudolph, A.E., Cherepanov, P., Dixon, J.E., Schwan, T.G., and Forsberg, A. (2002). Role of Yersinia murine toxin in survival of *Yersinia pestis* in the midgut of the flea vector. Science *296*, 733–735.
21. Chouikha, I., and Hinnebusch, B.J. (2014). Silencing urease: a key evolutionary step that facilitated the adaptation of *Yersinia pestis* to the flea-borne transmission route. Proc. Natl. Acad. Sci. USA *111*, 18709–18714.
22. Sebbane, F., Devalckenaere, A., Foulon, J., Carniel, E., and Simonet, M. (2001). Silencing and reactivation of urease in *Yersinia pestis* is determined by one G residue at a specific position in the ureD gene. Infect. Immun. *69*, 170–176.
23. Shimodaira, H. (2001). Multiple comparisons of log-likelihoods and combining nonnested models with applications to phylogenetic tree selection. Comm. Stat. A Theory Methods *30*, 1751–1772.
24. Allentoft, M.E., Sikora, M., Sjögren, K.-G., Rasmussen, S., Rasmussen, M., Stenderup, J., Damgaard, P.B., Schroeder, H., Ahlström, T., Vinner, L., et al. (2015). Population genomics of Bronze Age Eurasia. Nature *522*, 167–172.
25. Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., Stewardson, K., et al. (2015). Massive migration from the steppe was a source for Indo-European languages in Europe. Nature *522*, 207–211.
26. Tebelškis, P., and Jankauskas, R. (2006). The Late Neolithic grave at Gyvakarai in Lithuania in the context of current archaeological and anthropological knowledge. Archaeol. Baltica, 8–20.
27. Olalde, I., Brace, S., Allentoft, M.E., Armit, I., Kristiansen, K., Rohland, N., Mallick, S., Booth, T., Szécsényi-Nagy, A., Mittnik, A., et al. (2017). The Beaker Phenomenon and the Genomic Transformation of Northwest Europe. bioRxiv. https://doi.org/10.1101/135962.

28. Kuzmina, E.E. (2008). The Prehistory of the Silk Road (University of Pennsylvania Press).
29. Koryakova, L., and Epimakhov, A.V. (2007). The Urals and Western Siberia in the Bronze and Iron Ages (Cambridge University Press).
30. Hinnebusch, B.J. (2012). Biofilm-dependent and biofilm-independent mechanisms of transmission of *Yersinia pestis* by fleas. In Advances in Yersinia Research, A.M.P. de Almeida, and N.C. Leal, eds. (New York: Springer), pp. 237–243.
31. Hinnebusch, B.J., Fischer, E.R., and Schwan, T.G. (1998). Evaluation of the role of the *Yersinia pestis* plasminogen activator and other plasmid-encoded factors in temperature-dependent blockage of the flea. J. Infect. Dis. *178*, 1406–1415.
32. Eisen, R.J., Bearden, S.W., Wilder, A.P., Montenieri, J.A., Antolin, M.F., and Gage, K.L. (2006). Early-phase transmission of *Yersinia pestis* by unblocked fleas as a mechanism explaining rapidly spreading plague epizootics. Proc. Natl. Acad. Sci. USA *103*, 15380–15385.
33. Chouikha, I., and Hinnebusch, B.J. (2012). Yersinia–flea interactions and the evolution of the arthropod-borne transmission route of plague. Curr. Opin. Microbiol. *15*, 239–246.
34. Sun, Y.-C., Jarrett, C.O., Bosio, C.F., and Hinnebusch, B.J. (2014). Retracing the evolutionary path that led to flea-borne transmission of *Yersinia pestis*. Cell Host Microbe *15*, 578–586.
35. Eisen, R.J., Dennis, D.T., and Gage, K.L. (2015). The Role of Early-Phase Transmission in the Spread of *Yersinia pestis*. J. Med. Entomol. *52*, 1183–1192.
36. Vetter, S.M., Eisen, R.J., Schotthoefer, A.M., Montenieri, J.A., Holmes, J.L., Bobrov, A.G., Bearden, S.W., Perry, R.D., and Gage, K.L. (2010). Biofilm formation is not required for early-phase transmission of *Yersinia pestis*. Microbiology *156*, 2216–2225.
37. Johnson, T.L., Hinnebusch, B.J., Boegler, K.A., Graham, C.B., MacMillan, K., Montenieri, J.A., Bearden, S.W., Gage, K.L., and Eisen, R.J. (2014). Yersinia murine toxin is not required for early-phase transmission of *Yersinia pestis* by Oropsylla montana (Siphonaptera: Ceratophyllidae) or Xenopsylla cheopis (Siphonaptera: Pulicidae). Microbiology *160*, 2517–2525.
38. Lathem, W.W., Price, P.A., Miller, V.L., and Goldman, W.E. (2007). A plasminogen-activating protease specifically controls the development of primary pneumonic plague. Science *315*, 509–513.
39. Sebbane, F., Jarrett, C.O., Gardner, D., Long, D., and Hinnebusch, B.J. (2006). Role of the *Yersinia pestis* plasminogen activator in the incidence of distinct septicemic and bubonic forms of flea-borne plague. Proc. Natl. Acad. Sci. USA *103*, 5526–5530.
40. Zimbler, D.L., Schroeder, J.A., Eddy, J.L., and Lathem, W.W. (2015). Early emergence of *Yersinia pestis* as a severe respiratory pathogen. Nat. Commun. *6*, 7487.
41. Ochman, H., and Moran, N.A. (2001). Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science *292*, 1096–1099.
42. Perry, R.D., and Fetherston, J.D. (1997). *Yersinia pestis*–etiologic agent of plague. Clin. Microbiol. Rev. *10*, 35–66.
43. Barrett, Ronald, Kuzawa, Christopher W., McDade, Thomas, and Armelagos, G.J. (1998). Emerging and re-emerging infectious diseases: the third epidemiologic transition. Annu. Rev. Anthropol. *27*, 247–271.
44. Ryabogina, N.E., and Ivanov, S.N. (2011). Ancient agriculture in Western Siberia: problems of argumentation, paleoethnobotanic methods, and analysis of data. Archaeol. Ethnol. Anthropol. Eurasia *39*, 96–106.
45. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. *2010*, t5448.
46. Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., and Nieselt, K. (2016). EAGER: efficient ancient genome reconstruction. Genome Biol. *17*, 60.
47. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
48. 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). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
49. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.
50. 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.
51. Okonechnikov, K., Conesa, A., and García-Alcalde, F. (2016). Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics *32*, 292–294.
52. Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Curr. Protoc. Bioinformatics *11*, 11.10.1–11.10.33.
53. Bos, K.I., Harkins, K.M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S.A., Bryant, J.M., Harris, S.R., Schuenemann, V.J., et al. (2014). Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. Nature *514*, 494–497.
54. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. *30*, 2725–2729.
55. 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.
56. Schmidt, H.A., Strimmer, K., Vingron, M., and von Haeseler, A. (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics *18*, 502–504.
57. Keane, T.M., Creevey, C.J., Pentony, M.M., Naughton, T.J., and Mclnerney, J.O. (2006). Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC Evol. Biol. *6*, 29.
58. Cingolani, P., Platts, A., Wang, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92.
59. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.
60. Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis (Springer).
61. R Development Core Team (2008). R: A Language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing), http://www.R-project.org.
62. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192.
63. 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.
64. Zhgenti, E., Johnson, S.L., Davenport, K.W., Chanturia, G., Daligault, H.E., Chain, P.S., and Nikolich, M.P. (2015). Genome Assemblies for 11 *Yersinia pestis* Strains Isolated in the Caucasus Region. Genome Announc. *3*, e01030-15.
65. Kislichkina, A.A., Bogun, A.G., Kadnikova, L.A., Maiskaya, N.V., Platonov, M.E., Anisimov, N.V., Galkina, E.V., Dentovskaya, S.V., and Anisimov, A.P. (2015). Nineteen Whole-Genome Assemblies of *Yersinia pestis* subsp. *microtus*, Including Representatives of Biovars caucasica, talassica,

hissarica, altaica, xilingolensis, and ulegeica. Genome Announc. *3*, e01342-15.

66. 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.
67. Feldman, M., Harbeck, M., Keller, M., Spyrou, M.A., Rott, A., Trautmann, B., Scholz, H.C., Päffgen, B., Peters, J., McCormick, M., et al. (2016). A High-Coverage *Yersinia pestis* Genome from a Sixth-Century Justinianic Plague Victim. Mol. Biol. Evol. *33*, 2911–2923.
68. 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.
69. 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.
70. Rostunov, V.L. (2006). Opyt prekonstrukcii sakral'nogo postranstva rannykh kurganov Evropy I Severogo Kavkaza (Vladikavkaz: Severo-Oseinskiy institute gumannitarnykh I sozial'nykh issledovanny).
71. Rostunov, V.L. (2007). Epokha eneolita - sredney bronzy Zentralnogo Kavkaza III Opyt prekonstrukcii sakral'nogo postranstva rannykh kurganov Evropy (Vladikavkaz: Severo-Oseinskiy institute gumannitarnykh I sozial'nykh issledovanny).
72. Jaanits, L. (1948). Aruanne kaevamistest Kursi khk-s ja vallas Kunila külas Mäe-Jaaniantsu e. Keldri talu piirides asuval Jaaniantsu mäel 5.-10. (Tallinn: Institute of History of Tallinn University).
73. Jaanits, L., Laul, S., Lõugas, V., and Tõnisson, E. (1982). Eesti esiajalugu (Tallinn: Eesti Raamat).
74. Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763.
75. Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. *40*, e3.
76. Schuenemann, V.J., Bos, K., DeWitte, S., Schmedes, S., Jamieson, J., Mittnik, A., Forrest, S., Coombes, B.K., Wood, J.W., Earn, D.J.D., et al. (2011). Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of *Yersinia pestis* from victims of the Black Death. Proc. Natl. Acad. Sci. USA *108*, E746–E752.
77. Briggs, A., and Heyn, P. (2012). Preparation of next-generation sequencing libraries from damaged DNA. In Ancient DNA Methods in Molecular Biology, B. Shapiro, and M. Hofreiter, eds. (Humana Press), pp. 143–154.
78. 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.
79. 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.
80. 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.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Human archaeological remains	This paper	ENA: PRJEB19335, https://www.ebi.ac.uk/ena/data/view/PRJEB19335
Chemicals, Peptides, and Recombinant Proteins		
0.5 M EDTA pH 8.0	Life Technologies	Cat No./ID: AM9261
1x Tris-EDTA pH 8.0	AppliChem	Cat No./ID: A8569,0500
Proteinase K	Sigma-Aldrich	Cat No./ID: P2308-100MG
Guanidine hydrochloride	Sigma-Aldrich	Cat No./ID: G3272-500 g
3M Sodium Acetate pH 5.5	Sigma-Aldrich	Cat No./ID: S7899-500ML
Ethanol	Merck	Cat No./ID: 1009832511
Isopropanol	Merck	Cat No./ID: 1070222511
ATP	New England Biosciences	Cat No./ID: P0756 S
BSA 20mg/ml	New England Biosciences	Cat No./ID: B9000 S
Bst 2.0 DNA Polymerase	New England Biosciences	Cat No./ID: M0537 S
Buffer Tango	Life Technologies	Cat No./ID: BY5
dNTPs 25 mM	Thermo Scientific	Cat No./ID: R1121
Ethanol	Merck	Cat No./ID: 1009832511
NEBuffer 2 10x	New England Biosciences	Cat No./ID: B7002 S
T4 DNA Polymerase	New England Biosciences	Cat No./ID: M0203 L
T4 Polynucleotide Kinase	New England Biosciences	Cat No./ID: M0201 L
Pfu Turbo Cx Hotstart DNA Polymerase	Agilent Technologies	Cat No./ID: 600412
Tween 20	Sigma-Aldrich	Cat No./ID: P9416-50ML
Uracil Glycosylase inhibitor (UGI)	New England Biosciences	Cat No./ID: M0281 S
User Enzyme	New England Biosciences	Cat No./ID: M5505 L
Water Chromasolv Plus	Sigma-Aldrich	Cat No./ID: 34877-2.5L
Critical Commercial Assays		
Min Elute PCR Purification Kit	QIAGEN	Cat No./ID:28006
Quick Ligation Kit	New England Biosciences	Cat No./ID: M2200 L
DyNAmo Flash SYBR Green qPCR Kit	Life Technologies	Cat No./ID: F-415L
SureSelect DNA Capture Arrays 1M	Agilent Technologies	Cat No./ID:G3358A
High Pure Viral Nucleic Acid Large Volume Kit	Roche	Cat No./ID: 5114403001
Oligonucleotides		
IS1_adapter.P5 A*C*A*C*TCTTTCCCTACACG ACGCTCTTCCG*A*T*C*T	[45]	Sigma Aldrich
IS2_adapter.P7 G*T*G*A*CTGGAGTTCAGAC GTGTGCTCTTCCG*A*T*C*T	[45]	Sigma Aldrich
IS3_adapter.P5+P7 A*G*A*T*CGGAA*G*A*G*C	[45]	Sigma Aldrich
P5 Indexing 5′-AATGATACGGCGACCACCGAGATCT ACACxxxxxxxxACACTCTTTCCCTACACGACGC-3′	[45]	Sigma Aldrich
P7 Indexing 5′-CAAGCAGAAGACGGCATACGAGA TxxxxxxxxGTGACTGGAGTTCAGACGTGTGC-3′	[45]	Sigma Aldrich
Deposited Data		
Y. pestis LNBA aDNA data	This study	ENA: PRJEB19335, https://www.ebi.ac.uk/ena/data/view/PRJEB19335

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
EAGER	[46]	https://github.com/apeltzer/EAGER-GUI
Burrows-Wheeler Aligner (BWA)	[47]	http://bio-bwa.sourceforge.net/; RRID: SCR_010910
samtools	[48]	http://samtools.sourceforge.net/; RRID: SCR_002105
MarkDuplicates (Picard)	http://broadinstitute.github.io/picard/	http://broadinstitute.github.io/picard/; RRID: SCR_006525
Dustmasker (BLAST+)	[49]	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
MapDamage	[50]	https://ginolhac.github.io/mapDamage/; RRID: SCR_001240
Qualimap	[51]	http://qualimap.bioinfo.cipf.es/; RRID: SCR_001209
GATK UnifiedGenotyper	[52]	https://software.broadinstitute.org/gatk/; RRID: SCR_001876
MultiVCFAnalyzer	[53]	https://github.com/alexherbig/MultiVCFAnalyzer
MEGA6	[54]	http://megasoftware.net/; RRID: SCR_000667
PhyML 3.0	[55]	http://www.atgc-montpellier.fr/phyml/; RRID: SCR_014629
TREE-PUZZLE	[56]	http://www.tree-puzzle.de/
CONSEL	[23]	http://stat.sys.i.kyoto-u.ac.jp/prog/consel/
BEAST	[15]	http://beast.community/; RRID: SCR_010228
Calib 7.1	http://calib.qub.ac.uk/calib/	http://calib.qub.ac.uk/calib/
ModelGenerator	[57]	http://mcinerneylab.com/software/modelgenerator/#
SnpEff	[58]	http://snpeff.sourceforge.net/; RRID: SCR_005191
BEDtools	[59]	http://bedtools.readthedocs.io/en/latest/; RRID: SCR_006646
ggplot2	[60]	http://ggplot2.org/; RRID: SCR_014601
R Project for Statistical Computing	[61]	http://www.r-project.org/; RRID: SCR_001905
Integrative Genomics Viewer (IGV)	[62]	https://www.broadinstitute.org/igv/; RRID: SCR_011793
Other		
Y. pestis Bronze Age aDNA	[7]	ENA: PRJEB10885, https://www.ebi.ac.uk/ena/data/view/PRJEB10885
130 genomes modern comparison dataset	[63]	SRA: SRA010790, https://www.ncbi.nlm.nih.gov/sra/?term=SRA010790
Georgia_1412	[64]	GenBank: CP006783, https://www.ncbi.nlm.nih.gov/nuccore/CP006783
Georgia_1413	[64]	GenBank: CP006762, https://www.ncbi.nlm.nih.gov/nuccore/CP006762
Georgia_1670	[64]	GenBank: AYLR00000000, https://www.ncbi.nlm.nih.gov/nuccore/AYLR00000000
Kyrgyzstan_790	[64]	GenBank: CP006806, https://www.ncbi.nlm.nih.gov/nuccore?term=CP006806
Armenia_1522	[64]	GenBank: CP006758, https://www.ncbi.nlm.nih.gov/nuccore/CP006758
Russia_Federation_2944	[64]	GenBank: CP006792, https://www.ncbi.nlm.nih.gov/nuccore/CP006792

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Georgia_3067	[64]	GenBank: CP006754, https://www.ncbi.nlm.nih.gov/nuccore?term=CP006754
Georgia_8787	[64]	GenBank: CP006748, https://www.ncbi.nlm.nih.gov/nuccore/CP006748
Georgia_3770	[64]	GenBank: CP006751, https://www.ncbi.nlm.nih.gov/nuccore?term=CP006751
Armenia_14735	[64]	GenBank: AYLS00000000, https://www.ncbi.nlm.nih.gov/nuccore?term=AYLS00000000
Azerbaijan_1045	[64]	GenBank: CP006794, https://www.ncbi.nlm.nih.gov/nuccore?term=CP006794
19 draft genomes *Y. pestis* subsp. *microtus* strains	[65]	BioProject: PRJNA269675, https://www.ncbi.nlm.nih.gov//bioproject/PRJNA269675
Black Death *Y. pestis* genome	[66]	SRA: SRA045745, https://www.ncbi.nlm.nih.gov/sra/?term=SRA045745
Justinian *Y. pestis* genome	[67]	ENA: PRJEB14851, https://www.ebi.ac.uk/ena/data/view/PRJEB14851
Bolgar *Y. pestis* genome	[68]	ENA: PRJEB13664, https://www.ebi.ac.uk/ena/data/view/PRJEB13664
Observance *Y. pestis* genomes	[69]	ENA: PRJEB12163, https://www.ebi.ac.uk/ena/data/view/PRJEB12163

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Johannes Krause (krause@shh.mpg.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Ethical approvals

The scope of the study was limited to prehistoric archaeological material. Therefore, no ethics approvals were required.

Description of samples and archaeological sites

Rasshevatskiy, Russia, RK1001, Yamnaya Complex

The Rasshevatskiy site is part of an agglomeration of burial mound cemeteries in linear structure situated not far from river Kuban. Here the river turns from a south-north direction to the West. The cemetery 1 consists of 27 burial mounds, with mound 21 being the largest, and stretching over approximately 2 km. Nearby mound 20 was another large mound, all others were moderate in size. With the exception of mound 24, all features were excavated during construction work in two field seasons 1998 and 2000 by J.B. Berezin and V.L. Rostunov [70, 71].

The big mound 21 had an oval shape with an annex and a size of 85 × 110 m and was 6.2 m in height. The mound was built in five construction phases of which the first is associated with the Maikop epoch, the second was built by Yamnaya groups, and the third and fourth are related to the Novotitorovskaya culture, which is a local variant succeeding Yamnaya in the Northwest Caucasus. The last mound-shell was added by groups of the Catacomb grave complex and North Caucassian cultural complexes. Thus, the mound included 22 graves in total with different cultural affiliations and the dates of the Bronze Age individuals span from before 4560 ± 60 BP to at least 3960 BP, i.e., approximately 500-600 years. Novotitorovskaya grave 7 as well as Catacomb graves 8 and 20 contained wooden wagons or wagon parts.

Grave 11 (RK1001) is one of four Yamnaya burials in the mound and was dug into the second mound shells. The grave is directly built on top of grave 13 dated to 4,447 ± 22 uncalBP (MAMS-29818). The second shell was constructed over grave 3, another probable Yamnaya grave-pit but without a skeleton or inventory. Grave 11 is a typical Yamnaya inhumation in an oval-rectangular pit lying in 5.22 m deep slightly off-center. The individual was placed in a supine position on the back with the head in a western direction, legs crouched and slightly tilted to the north. The skeleton was placed on organic bedding and remains of dark red ochre were found below the feet. Likewise, small pieces of chalk were found below the skeleton. As inventory a small silex instrument was found. The individual is a male, 30-39 years of age. He presents enamel hypoplasia and cribra orbitalia; both signs of childhood stress. There are signs of chronic inflammation of the endocrania, plural prints of small blood vessels, and newly built bone on the inner surface of the frontal and occipital bones and inside the frontal and maxillary sinuses. Furthermore, there are signs of chronic inflammation present on the legs of the individual: slight periostitis and blood vessel imprints visible on the femurs, tibias, and fibulas, and the

metatarsals are also affected by inflammation. Pathologies such as interproximal grooving, chipping, calculus, and abnormal wear of the front teeth in comparison with the posterior teeth are observed in the dentition of the individual. In addition, numerous fractures of the foot bones were recorded. The individual was directly radiocarbon dated 4171 ± 22 uncalBP (MAMS-29816).

The principle stratigraphy of mound 21 is published [70, 71], the comprehensive publication including anthropology and isotope-studies, as well as a re-evaluation of the stratigraphy of the sites is forthcoming.

Beli Manastir - Popova zemlja, Croatia, GEN72, Vučedol Complex

This site is located in eastern Croatia, approximately 2 km West from the town Beli Manastir. Two rescue excavations took place in 2014 and 2015 unearthing approximately 37,000 m^2, where two main layers were identified: a prehistoric layer with several strata from Neolithic and Chalcolithic periods; and a Roman layer where two rectangular brick furnaces were uncovered.

In the prehistoric layer, a total of 28 dwelling pits and 35 inhumations were recovered. Some of the prehistoric burials (21 in total) were found within the dwelling pits either at the bottom or at the top of the backfills of the pits. The rest of the inhumations were located at the bottom of either waste pits or a large canal at the eastern side of the settlement. The Neolithic burials were found in a crouched position on either the right or the left side with different orientation in most of the cases, and some had one or more ceramic vessels placed by the head of the deceased.

The individual GEN72 was found in the grave number 17. The skeleton was in a crouched lying on its belly on its left side. A well-healed ante-mortem depression fracture was located on the posterior part of the left parietal bone. It also exhibited mild, healed ectocranial porosity on the occipital and parietal bones, healed cribra orbitalia on its superior orbits, and mild, healed periostitis on the right tibia and fibula. The individual was recently dated by AMS dating to 4176 ± 28 uncal BP (Labno. BRAMS-1304).

Gyvakarai, Lithuania, Gyvakarai1, Corded Ware/Boat Axe Complex

This site is located in the northern part of Lithuania on the steep gravelly bank (elevation up to 79 m a. s. l.) of the rivulet Žvikė. The burial was accidently discovered in 2000 by locals and subsequently rescue excavations were conducted in the same year in the surrounding area of the highly disturbed grave resulting in discovery of a single grave, Gyvakarai 1, containing a fragmentary skeleton belonging to an adult man, 35-45 years of age, and C14 dated to the Late Neolithic (3,745 ± 70 uncal BP (right tibia, Ki-9470) and 3,710 ± 80 uncal BP (left ulna, Ki-9471) [26]. The dating, the fact that it is a singular grave of an adult individual and the set of grave goods, including a boat-shaped polished stone axe, led to the association of this individual with the Late Neolithic Corded Ware/Boat Axe cultural complex. The individual was recently dated by AMS dating to 4030 ± 30 uncal BP (Labno. Poz-61584).

Kunila, Estonia, Kunila II, Corded Ware Complex

This burial site is situated in Central Estonia, 4 km south-west of Puurmani on the western side of a small drumlin on Jaaniantsu Hill. The burial site was discovered in 1938 during gravel digging, revealing a stone axe and loose human bones. During archaeological excavations in 1948 the skeletal remains of two adult males (Kunila I and II) were uncovered (AI 3723) [72]. Kunila I was buried with a stone adze and a battle-axe; Kunila II with an adze of white flint and a grinding stone. The burials are attributed to Corded Ware Culture [73] and Kunila II has been directly dated to 3960 ± 40 uncal BP (mandible; Poz-10825) [12].

Haunstetten Unterer Talweg 85, Feature 1343 = Grave I/3, Augsburg, Germany, Bell Beaker Complex

The site of ''Unterer Talweg 85'' (due to a change of the street numbers after the excavation, the site is nowadays also known as ''Unterer Talweg 49'') is situated in Haunstetten, a quarter of Augsburg to the very south of the city and only 300 m north of Unterer Talweg 58-62, from which samples are also included in this study (see above). The cemetery consists of two small groups of burials, group I with 5 graves and group II with 2 graves, both situated roughly 20 m apart from each other. Group I, the so-called northern group, was excavated in 2001. Three single burials were radiocarbon dated and their 2 sigma ranges fall between 2,465 and 2,152 cal BCE [13].

Palaeogenetic data from the dentine of the individual from grave I/3 (Feat.no. 1343) was included in this study. Grave I/3 (feat. 1343) is dated to 2,397–2,149 cal BCE (2 sigma) and contained a male individual in contracted position with an arrowhead and several pieces of flint as grave goods.

Haunstetten Postillionstrasse 6, Feature 6 = Grave 36, Augsburg, Germany, Early Bronze Age

The site of ''Postillionstrasse'' is also situated in Haunstetten ca. 3.2 km south of the site ''Unterer Talweg 85'' and was excavated in 19. The cemetery consists at least of 41 graves, all dating to the Early Bronze Age (total span of radiocarbon dates: 2,198-1,772 BCE). Three graves were covered with a burial mound and surrounded by a ring ditch.

Grave 36 (feat. 6) is dated to 2,009–1,883 cal BCE (2 sigma) [13]. It contained a male individual in crouched position. A copper dagger and a bangle, two flint arrowheads were part of his grave goods. The burial was placed in the southern part of the cemetery surrounded by other graves.

METHOD DETAILS

Sampling and extraction

Sampling of a total of 563 tooth and bone samples (Russia (n = 122), Hungary and Croatia (n = 139), Lithuania (n = 27), Estonia (n = 45), Latvia (n = 10), and Germany (Althausen n = 4, Augsburg n = 83, Mittelelbe-Saale n = 133)) took place in the clean room facilities of the Institute for Archaeological Sciences at the University of Tübingen, the Institute of Archaeology RCH HAS in Budapest and of the MPI-SHH in Jena. After irradiation with UV light to remove surface contamination, teeth were sawed apart transversally at the border of crown and root, and dentine from inside the crown was sampled and powdered using a sterile dentistry drill. For the samples processed in Budapest, whole teeth were powdered. For bone samples, the surface layer from the sampling area was removed with a

dentistry drill prior to obtaining bone powder from the inside of the bone by drilling. For each specimen we gathered between ~30 and 120 mg of powder to be used for DNA extraction.

Extraction was performed following a protocol optimized for the recovery of small ancient DNA molecules [74] in the following manner: around 30-50 mg of powder (bone or teeth) for each individual was combined with the extraction buffer (EDTA,1.04ml; Proteinase K, 0.03ml; and UV-water 0.9ml), and rotated over night at 37°C, then it was centrifuged at 14000 rpm to pellet the bone powder and the supernatant was transferred to a 50ml Falcon tube (containing funnel and purification column) already containing 10 mL of binding buffer and 400 μg sodium acetate. The Falcon tube was then spun at 1500 rpm for 8 min with slow acceleration and fast deceleration. After which, the column was transferred into a new collection tube and liquid remaining in the funnel was transferred to the column. This was followed by 2 washing steps consisting of: adding 450 μL of wash buffer (40ml Ethanol plus 22ml wash buffer from the High Pure Viral Nucleic Acid Large Volume Kit) to the column and spun it at 8000 rpm for 1 min; plus two dry spins at 14000 rpm for 1 min. Then the DNA was eluted in a new tube by doing two elution steps by adding into the purification column 50 μL of TET and spinning for 1 min at 14000 rpm. The extraction resulted in 100 μL of DNA extract per sample. An aliquot of 20 μL of extract was used to generate double-indexed libraries as described elsewhere [45, 75]. In brief, samples were blunt-end repaired by adding the extract to 60 μL of Blunt-End Repair mastermix (NEB Buffer 2, 10.50 μl; ATP, 10.50 μl; BSA, 4.20 μl; dNTPs, 0.42 μl; T4 PNK, 4.20 μl; T4 Polymerase, 0.84 μl; and UV water, 53.34 μl); the reactions were incubated in a thermocycler at 15°C for 15 min, followed by 15 min at 25°C and thereafter purified with a MinElute kit. The blunt-end repaired extract was then ligated to Illumina adapters by adding 21 μL of Quick Ligase Buffer, 1.05 μL of adaptor mix (0.25 μM concentration) and 18 μL of the purified blunt-end repaired extract, and finally adding 1 μL of Quick ligase (0.125 U concentration) and incubated for 20 min at 20°C, followed by purification with a MinElute kit. Finally, 20 μL of each sample was adaptor filled by adding 20 μL mastermix (Themopol Buffer, 4 μl; dNTPs 0.20 μl; Bst, 2 μl; and UV-Water 13.80 μl) and incubated at 37°C for 30 min followed by 10 min at 80°C. Then, each sample was quantified by qPCR. Each sample was then indexed with dual sample-specific indexes. For each indexing reaction, each library was split to ensure that there were less than 1.5e+10 copies of DNA in each indexing reaction. For example, if the library had to be split into two, we added 18 μL of the prepared libraries and combine it with the indexing mastermix (Pfu Turbo Buffer, 10 μl; BSA, 1.5 μl; dNTPs, 1 μl, Pfu Turbo Polymerase, 1 μl; P5 and P7 sample-specific indexes, 2 μL for each; and UV-water 64.5 μl) and if it had to be split into four reactions, we would add 9 μL of the library into the indexing mastermix, only changing the amount of UV-water to 73.5 μl. Once the reactions were set up, these were transferred to a modern DNA lab for indexing PCR: initial denaturation at 95°C for 2 min and 10 cycles of: 30 s at 95°C, 30 s at 58°C, and 1 min at 72°C; and a elongation phase of 10 min at 72°C. The amplified double-indexed libraries were then purified with a MinElute kit. Negative controls were included in the extraction and library preparation and taken along for all further processing steps. Negative controls were included in the extraction and library preparation and taken along for all further processing steps.

Shotgun screening sequencing

Libraries were PCR-amplified and quantified using an Agilent 2100 Bioanalyzer DNA 1000 chip and pooled at equimolar concentrations prior to paired-end sequencing on a NextSeq500 with 2x101+8+8 and a HiSeq2500 with 2x101+8+8 cycles according to the manufacturer's instructions to a depth of ~1.5 million reads per library.

In silico screening

The sequencing data for the 170 samples was preprocessed with ClipAndMerge [46] to remove adaptors, base quality-trim (20) and merging and filtering for only merged reads. Reads were mapped using the BWA aln algorithm [47] to a multi-species reference panel, containing various representatives of the genus *Yersinia* (Table 1) and the plasmids of *Yersinia pestis*: pCD1, pMT1 and pPCP1 from *Y. pestis* CO92. The region comprising 3000-4200bp of the *Y. pestis* specific plasmid pPCP1 was masked in the reference, since it is highly similar to an expression vector used during the production of enzyme reagents [76].

Mapped files were then filtered for reads with a mapping quality higher than 20 with Samtools [48]. PCR duplicates were removed using the MarkDuplicates tool in Picard (1.140, http://broadinstitute.github.io/picard/). The number of reads mapping specifically to each genome and to the plasmids were retrieved from the bam files using Samtools [48] idxstats. An endogenous based score was used to assess the potential of the sample being 'positive' for *Y. pestis*. It was calculated as follows:

$$\frac{(YPS - \max(YS))}{M} \times 1,000$$

where YPS is the number of reads specifically mapping to *Y. pestis*; YS is the maximum number of reads mapping specifically to a *Yersinia* species with the exception of *Y. pestis* and M is the total number of merged reads in the sample. By using the maximum number of reads mapping to another species of the genus *Yersinia*, the score takes in account different source of contamination other than *Y. pseudotuberculosis*. Five samples (RK1001, Gyvakarai1, KunilaII, 6Post and 1343UnTal85) fulfilled the criteria for being considered strong candidates (score higher than 0.005 (Figure S4C) and reads mapping to all plasmids). Another samples, GEN72, was also included in further processing and analysis since it had higher numbers mapping the *Y. pestis* chromosome and plasmids even though it did not full-fill the score requirements. For a detailed description of the archaeological sites and individuals see the next section.

Deep shotgun sequencing

The five strong candidate samples detected in screening of the shotgun data were processed for deep shotgun sequencing as following: For Gyvakarai1 the screening library described above was pair-end sequenced on two lanes of a HiSeq4000 for 100 cycles, and on a full run of a NextSeq500 for 75 cycles. The screening library for KunilaII was pair-end sequenced deeper on 80% of one lane of a HiSeq4000 for 100 cycles. Additionally, 40 μL of DNA extract of KunilaII was converted in to a library treated with UDG and endonuclease VIII to remove deaminated bases [77], and pair-end sequenced on one lane of a HiSeq4000 for 75 cycles.

For RK1001, Post6 and 1343UnTal85, 60 μL of DNA extract each were converted into DNA libraries using so-called UDG-half treatment, whereby deaminated bases are partially removed and retained mostly at the ends of the molecule [78]. The library of RK1001 was deep shotgun pair-end sequenced in 8 lanes of a HiSeq4000 for 55 cycles. The libraries of 6Post and 1343UnTal85 were deep shotgun single-end sequenced on 2 and a half lanes of a HiSeq4000 for 75 cycles. Post6 was additionally pair-end sequenced on a full run of a NextSeq500 for 75 cycles.

Y. pestis in-solution capture

Y. pestis whole-genome DNA capture probes were designed using as template sequences the *Y. pestis* CO92 chromosome (NC_003143.1), *Y. pestis* CO92 plasmid pMT1 (NC_003134.1), *Y. pestis* CO92 plasmid pCD1 (NC_003131.1), *Y. pestis* KIM 10 chromosome (NC_004088.1), *Y. pestis* Pestoides F chromosome (NC_009381.1) and *Y. pseudotuberculosis* IP 32953 chromosome (NC_006155.1). We used a 6 bp tiling with a probe length of 52 bp with an additional 8 bp 3′ linker sequence as described in [79]. Low complexity regions were masked using dustmasker [49] (version 2.2.32+). Redundant probes as well as probes with more than 20% masked nucleotides were discarded. The procedure resulted in 816,413 unique probe sequences. A second probe set was created with a coordinate offset of 3 bp resulting in 827,438 unique probe sequences. In combination the two probe sets represent an effective tiling density of 3 bp. The two probe sets were ordered on two 1 million feature Agilent SureSelect DNA Capture Arrays. The full capacity of the arrays was filled up with randomly selected probes. The two arrays were turned into in-solution DNA capture libraries as described elsewhere [79].

For GEN72, 25 μL of DNA extract was converted into DNA libraries using so-called UDG-half treatment as described above [78]. The UDG-half libraries of RK1001 and GEN72 were enriched for *Y. pestis* DNA using in-solution DNA capture probes (see above) as described elsewhere [25, 79, 80]. The capture products of RK1001 and GEN72, were sequenced on 1 and 0.6 of the lane, respectevely, of the HiSeq4000 for 75 cycles.

Genome reconstruction and authentication

All samples were processed with the EAGER pipeline [46]. Sequencing quality for each sample was evaluated with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and adaptors clipped using the ClipAndMerge module in EAGER. For paired-end data, the reads were also merged with ClipAndMerge and only the merged reads were kept for further analysis. Furthermore, 7 bp in the 5 prime end were clipped in GEN72 due to the presence of an extra barcode in the DNA library.

Individual sample treatment due to laboratory preparation and sequencing strategies

Gyvakarai1: two HiSeq lanes and one Next-Seq run paired-end of the non-UDG treated library were combined and reads mapped to *Y. pestis* CO92 reference with BWA aln (-l 16, -n 0.01, hereby referred to as non-UDG parameters). Reads with mapping quality scores lower than 37 were filtered out. PCR duplicates were removed with MarkDuplicates. MapDamage (v2.0) [50] was used to calculate damage plots (Figure S1). Coverage was calculated with Qualimap (v2.2) [51].

KunilaII: UDG and the non-UDG libraries were sequenced in 2 HiSeq pair-end lanes and processed separately until calculation of the coverage. The non-UDG treated libraries were mapped with non-UDG parameters while the UDG treated library reads were mapped with more stringent parameters (-l 32, -n 0.1, referred to as UDG parameters). Reads with mapping qualities less than 37 were filtered out and duplicates were removed with MarkDuplicates as before. The non-UDG bam file was used to calculate damage plots as indicated above (Figure S1). After duplicate removal, the UDG- and non-UDG treated BAM files were merged together and used to calculate the coverage as above.

GEN72, Post6 and 1343UnTal85: the UDG-half treated libraries were sequenced in two HiSeq lanes for Post6 and 1343UnTal85 and 19,777,683 reads were generated in the HiSeq for GEN72, and two different runs were performed. For the first run, reads without clipping were used to retain miscoding lesions indicative of aDNA. BWA aln was used for mapping with non-UDG parameters (-l 16 and -n 0.01). Reads with mapping qualities lower than 37 were filtered and PCR duplicates were removed with MarkDuplicates as described above. Coverage and damage plots were calculated as above (Figure S1). After clipping the last two bases with the module ClipAndMerge in eager, potentially affected by damage, the samples were mapped with UDG parameters.

RK1001: UDG-half library was shotgun sequenced pair-end in 8 HiSeq lanes and in-solution captured and sequenced single end to a depth of 303,148,884 reads sequenced in the HiSeq. Shotgun and captured data were combined in a fastq file and processed as described above for GEN72, Post6 and 1343UnTal85.

SNP calling, heterozygosity, and phylogenetic analysis

Prior to SNP calling in order to avoid false SNP calling due to aDNA damage, the quality scores of damaged sites in the non-UDG treated samples were downscaled using MapDamage [50] (v2.0), as performed in previous analysis [7]. For the UDG-half data, the files with the two last bases clipped, hence removing the damage signal, and mapped with UDG parameters were used for SNP calling (see above).

SNP calling was performed with GATK UnifiedGenotyper [52] in EAGER[43] with default parameters and the 'EMIT_ALL_SITES' output mode.

VCF files were generated for the two complete genomes from Rasmussen et al. (2015) [7], the Black Death [66], Justinianic Plague [67], Bolgar [68] and Observance [69] genomes and a curated dataset of 130 modern genomes [63] in addition to 11 samples from the Former Soviet Union [64] and 19 draft genomes of *Y. pestis* subsp. *microtus* strains [65] in the following manner: All the modern genomes were cut into 100bp reads and map to the reference *Y. pestis* CO92 with UDG parameters as described above, including the reference genome. The raw reads from the previous Bronze Age *Y. pestis* samples [7], were mapped to the reference with non-UDG parameters and downscaled using MapDamage [50] (v2.0), as above. The rest of the ancient genomes, were mapped to the reference with UDG parameters. The SNP calling for all the modern and ancient genomes was performed as described above.

All the VCF files from the modern and ancient dataset and the samples produced in this study were combined and processed with *MultiVCFAnalyzer* (v0.85, https://github.com/alexherbig/MultiVCFAnalyzer) [53] that produced a SNP table and a fasta alignment file containing the concatenation of all variable positions in the dataset (SNPs alignment), in respect to the reference *Y. pestis* CO92. In order to call a SNP a minimum genotyping quality (GATK) of 30 was required, with a minimum coverage of 3X, and with a minimal allele frequency of 90% for a homozygous call. No heterozygous calls were included in the output files.

The SNP alignment was curated by removing all alignment columns with missing data (complete deletion). The curated SNP alignment was then used to compute Neighbor Joining (NJ) and Maximum Parsimony (MP) trees with MEGA6 [54] and a Maximum Likelihood (ML) tree using PhyML 3.0 [55] with the GTR model used in previous *Y. pestis* work [7, 63], with 4 gamma categories and the best of NNI and SPR as tree branch optimization. The specific variants of *Y. pseudotuberculosis* were removed from the analysis to improve the visual resolution of the tree.

To check for infections with multiple strains, all the VCF files from modern and ancient dataset including the samples produced in this study were combined and processed with *MultiVCFAnalyzer* (v0.85, https://github.com/alexherbig/MultiVCFAnalyzer) [53] as described above, with the exception that heterozygous calls were included in the output files. The only parameters changed were minimal allele frequency of 90% for a homozygous and of 10% for the heterozygous calls (Figure S1).

Dating analysis

The SNP alignment after complete deletion was used for molecular dating using BEAST 1.8.2 [15]. The modern sample 0.PE3, also called Angola, was removed from the dataset due to its long branch.

For tip dating, all modern genomes were set to an age of 0. The dates of the ancient samples presented in this study plus the two complete genomes from Rasmussen et al. (2015) [7] were recalibrated with Calib 7.1 (http://calib.qub.ac.uk/calib/) to the IntCal13 calibration curve. The ancient samples were given the median calibrated probability as their age, and the 2 sigma interval was used as the boundaries for a uniform prior sampling (Table 2). The dates published for previous historical genomes were transformed to cal BP assuming 1950 as age 0 and given the mean as the age with the interval as the boundaries of a prior uniform distribution: RISE509 4729 (4625-4836) [15], RISE505 3635 (3575-3694) [15], Black Death 603 (602-604) [66]; Observance 229 (228-230) [69], Bolgar 569 (550-588) [68] and Justinian 1453 (1382-1524) [67].

The molecular clock was tested and rejected using MEGA6 (See Molecular Clock Test in QUANTIFICATION AND STATISTICAL ANALYSIS section below). Therefore, we followed previous work and used an uncorrelated relaxed clock with lognormal distribution [7, 63] with the substitution model GTR+G4, selected using ModelGenerator [57]. Tree model was set up to coalesce assuming a constant population size and a rooted ML tree was provided as a starting tree. Two independent 1,000,000,000 MCMC chains were computed sampling every 5,000 steps. The two chains were then combined using LogCombiner from BEAST 1.8.2 [15] with a 10 percent burn-in (100,000,000 steps per chain). The Effective Sample Size (ESS) of the posterior, prior, treeModel.rootHeight, tMRCA_allpestis are 4,589, 4,087, 1,054 and 7,571 respectively. The trees files for the 2 chains were combined with LogCombiner with 100,000,000 of burning and resampled every 20,000 steps giving a total number of 90,000 trees, that were used to produce a Maximum Clade Credibility tree using TreeAnnotator from BEAST 1.8.2 [15].

SNP effect analysis and virulence factors analysis

The SNP table from *MultiVCFAnalyzer* was provided to *SnpEff* [58] and the effect of the SNPs within genes present in the dataset was evaluated. Additionally the SNP table was manually assessed for possible homoplasies.

For the virulence factors, the samples were mapped as indicated above but without applying quality filtering and the percentage of coverage was calculated for each region using BEDtools [59] and plotted using the package ggplot2 [60] in R [61]. Additionally, *ureD* was manually explored for SNPs using Integrative Genomics Viewer (IGV) [62].

Indel analysis

The samples including the two complete Bronze Age genomes [7] were mapped against *Y. pseudotuberculosis* IP 32953 with bwa with non-UDG parameters (-n 0.01, -l 16), except for RK1001, GEN72, 1343UnTal85 and 6Post that were mapped with bwa with UDG

parameters (-n 0.1, -l 32),. The modern genomes from branch 0 (0.PE7, 0.PE2, 0.PE3 and 0.PE4), *Y. pestis* CO92 and *Y. pestis* KIM10 were *in-silico* cut in 100 bp fragments with 1bp tiling and mapped to *Y. pseudotuberculosis* reference using bwa with UDG parameters (-n 0.1, -l 32). The non-covered regions were extracted using the BEDtools genomecov function [59]. Missing regions larger than 1kb were comparatively explored in order to identify indels. Using the BEDtools intersect function [59], we extracted regions missing in the Neolithic genomes and present in the modern ones and also the regions missing in the modern ones but still present in the Neolithic genomes. The results were check by manual inspection in IGV [62].

QUANTIFICATION AND STATISTICAL ANALYSIS

In silico screening

A total of n = 563 samples were screened for the presence of *Y. pestis* and we determined candidates for further analysis by calculating a score as indicated above (see *In silico* screening in METHODS DETAILS).

Phylogenetic analysis

Phylogenetic analyses were performed with the NJ, MP and ML algorithms in MEGA6 [54] and PhyML 3.0 [55] (see SNP calling, heterozygocity, and phylogenetic analysis in METHODS DETAILS section above). To test the statistical support of the phylogenetic trees we bootstrapped each tree with 1,000 bootstrap replicates.

Tree topology test

To statistically test the obtained topology (see SNP calling, heterozygocity, and phylogenetic analysis in METHODS DETAILS section above), we performed statistical tests with TREE-PUZZLE(v5.3) [56] and CONSEL(v1.2) [23]. Four different topologies, described in Figure S4B, were provided together with the SNP alignment file to TREE-PUZZLE which evaluated these user defined trees with the following parameters: Neighbor joining + quartet sampling for parameter estimation uses, GTR model with 4 gamma categories. The output .sitelh file was used to perform the statistical test with CONSEL with 1,000 bootstraps.

Molecular clock test

In order to test the molecular clock hypothesis, we performed a molecular clock test (ML) in MEGA6, by providing the alignment used for phylogenetic analysis and the topology obtained with the Neighbout Joining algorithm (see Dating analysis in METHODS DETAILS above) with default parameters.

Dating analysis

Dating analysis was performed with BEAST 1.8.2 [15] as indicated above (see Dating analysis in METHODS DETAILS). We run two independent MCMC chains with 1,000,000,000 steps each.

DATA AND SOFTWARE AVAILABILITY

Raw sequencing data have been deposited at the European Nucleotide Archive under accession number ENA: PRJEB19335 (https://www.ebi.ac.uk/ena/data/view/PRJEB19335).