

# Lethal plague outbreaks in Lake Baikal hunter-gatherers 5,500 years ago

<https://doi.org/10.1038/s41586-026-10540-5>

Received: 31 October 2024

Accepted: 14 April 2026

Published online: 17 June 2026

Open access

 Check for updates

Ruairidh Macleod<sup>1,2,3,15</sup>✉, Frederik V. Seersholm<sup>1,3,15</sup>, Bianca De Sanctis<sup>1,4,5,6</sup>, Angela Lieveise<sup>7</sup>, Adrian Timpson<sup>8</sup>, Rick Schulting<sup>9</sup>, Jesper T. Stenderup<sup>1,3</sup>, Charleen Gaunitz<sup>1</sup>, Lasse Vinner<sup>1,3</sup>, Olga Ivanovna Goriunova<sup>10</sup>, Vladimir Ivanovich Bazaliiskii<sup>10</sup>, Sergei V. Vasilyev<sup>11</sup>, Erin Jessup<sup>12</sup>, Yucheng Wang<sup>1,3,4,13</sup>, Christopher Bronk Ramsey<sup>9</sup>, Mark G. Thomas<sup>9</sup>, Russell Corbett-Detig<sup>5,6</sup>, Astrid K. N. Iversen<sup>1,14</sup>, Andrzej W. Weber<sup>1,12</sup>✉, Martin Sikora<sup>1,3,16</sup>✉ & Eske Willerslev<sup>1,3,4,16</sup>✉

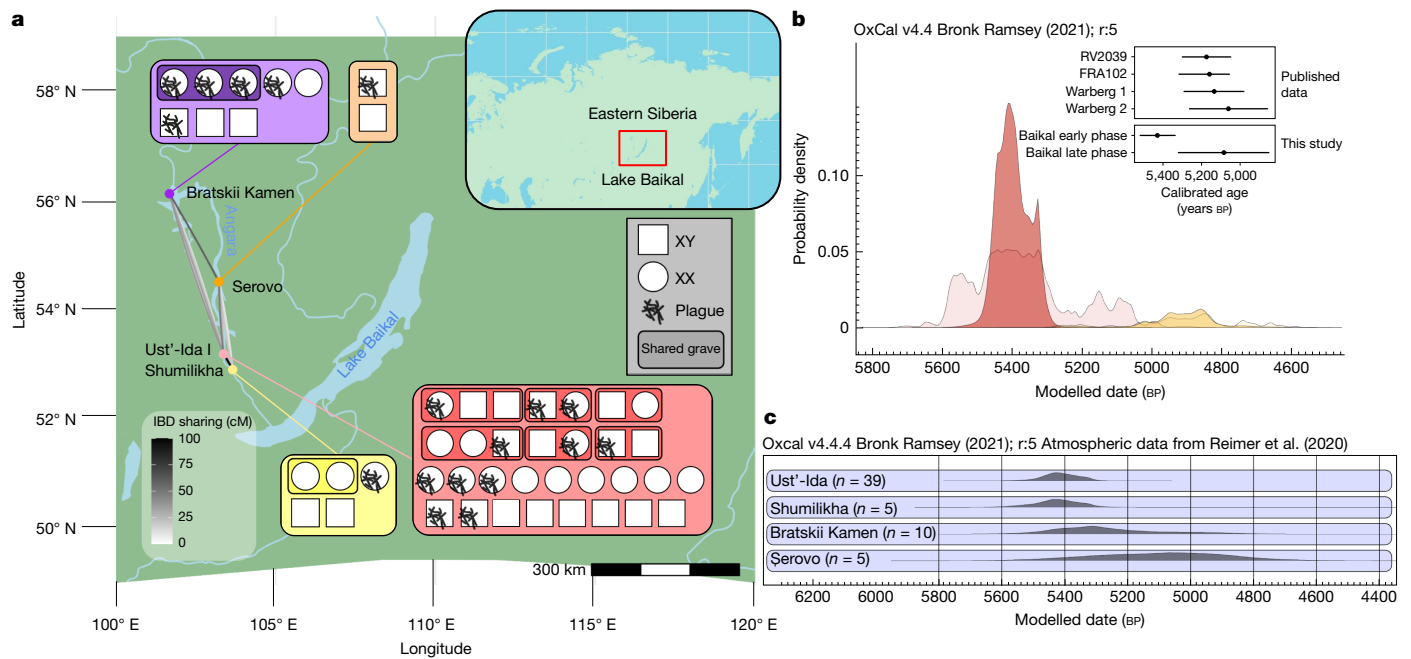
Plague is among the most devastating diseases in human history<sup>1</sup>. However, early strains of the plague-causing bacterium *Yersinia pestis* lacked virulence factors that are required for the bubonic form until around 3,800 years ago<sup>2,3</sup>. Consequently, the morbidity and mortality of early plague strains remain unclear. Here we describe early plague strains that are associated with two phases of outbreaks among mid-Holocene hunter-gatherers near Lake Baikal in southeast Siberia, beginning from about 5,500 years ago. These outbreaks occur across four hunter-gatherer cemeteries, with a 39% detection rate for plague infection. By reconstructing kinship pedigrees, we show that small familial groups were affected, consistent with human-to-human spread of disease, and that the first outbreak occurred within a single generation. The infections appear to have resulted in acute mortality, especially among children (aged 8 to 11 years). We further note functional differences, including in the *ypm* superantigen locus, which is also present in present day *Yersinia pseudotuberculosis*. The new strains diverge ancestrally to known *Y. pestis* and constrain the timing of its emergence, indicating that this happened before approximately 5,700 years ago. These findings show that plague outbreaks happened earlier than previously thought and were indeed lethal. We contend that the occurrence of outbreaks among mid-Holocene hunter-gatherer communities well outside the sphere of Late Neolithic Europe challenges the notion that higher population densities and lifestyle changes during the Neolithic agricultural transition were prerequisites for plague epidemics.

The analysis of ancient pathogen genomes has significantly expanded our understanding of the evolutionary history of human infectious diseases (for example, *Salmonella enterica*<sup>4</sup> and hepatitis B<sup>5</sup>), although this has principally been in the context of farming or pastoralist communities. *Y. pestis*, the aetiological agent of plague, is perhaps the most studied in this regard, and has had devastating consequences on human populations for millennia. Historical outbreaks of plague account for some of the most fatal events in human history<sup>1</sup>. The recovery of ancient DNA from plague victims has afforded extraordinary insights into the origins and evolution of plague at the time of these events<sup>6,7</sup>, and, remarkably, revealed infections in prehistoric individuals across Europe<sup>8</sup>. Historically and today, plague is associated with transmission via fleas from rodents, which successfully adapted to a human commensal niche in the Neolithic<sup>9</sup>. Genomic analysis of prehistoric

plague indicates that in early diverging strains, key genetic adaptations required for flea-mediated transmission of the disease and bubonic infection are absent<sup>2,3</sup>, leading to uncertainty over the transmission route and severity of these strains.

The detection of early plague cases across multiple generations of Late Neolithic farmers has been used to link outbreaks of the disease to a prolonged demographic decline between about 5,300–4,900 calibrated years before the present (cal BP)<sup>10,11</sup>, although an alternative explanation attributes the decline to agricultural crisis<sup>12,13</sup>. The former interpretation has been controversial, with others suggesting infections as more closely resembling benign foodborne enteritis<sup>14</sup>. The similarity or otherwise of these early strains to *Y. pseudotuberculosis*—the closest relative of *Y. pestis*—has been an important point of interest through such discussions, and based on existing ancient genomes, *Y. pestis* has

<sup>1</sup>Lundbeck Foundation Geogenetics Centre, Globe Institute, University of Copenhagen, Copenhagen, Denmark. <sup>2</sup>All Souls College, University of Oxford, Oxford, UK. <sup>3</sup>Centre for Ancient Environmental Genomics, Globe Institute, University of Copenhagen, Copenhagen, Denmark. <sup>4</sup>Department of Genetics, University of Cambridge, Cambridge, UK. <sup>5</sup>Department of Biomolecular Engineering, UC Santa Cruz, Santa Cruz, CA, USA. <sup>6</sup>Genomics Institute, UC Santa Cruz, Santa Cruz, CA, USA. <sup>7</sup>Department of Anthropology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada. <sup>8</sup>Research Department of Genetics, Evolution and Environment, University College London, London, UK. <sup>9</sup>School of Archaeology, University of Oxford, Oxford, UK. <sup>10</sup>Scientific Research Centre “Baikalskii Region”, Irkutsk State University, Irkutsk, Russia. <sup>11</sup>Institute of Ethnology and Anthropology, Russian Academy of Sciences, Moscow, Russia. <sup>12</sup>Department of Anthropology, University of Alberta, Edmonton, Alberta, Canada. <sup>13</sup>Group of Alpine Paleoeology and Human Adaptation (ALPHA), State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources (TPESER), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China. <sup>14</sup>Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. <sup>15</sup>These authors contributed equally: Ruairidh Macleod, Frederik V. Seersholm. <sup>16</sup>These authors jointly supervised this work: Martin Sikora, Eske Willerslev. ✉e-mail: ruairidh.macleod@all-souls.ox.ac.uk; aweber@ualberta.ca; martin.sikora@sund.ku.dk; ewillerslev@sund.ku.dk



**Fig. 1 | Overview of the spatiotemporal distribution of ancient humans and plague infections in this study.** **a**, Locations of affected cemeteries on the Angara River northwest of Lake Baikal, and IBD sharing between the sampled occupants of cemeteries (pairwise sharing lines between sites; greyscale ramp indicates total IBD sharing in segments larger than 3 centimorgans (cM) totalling more than 10 cM in pairwise relationships between sites; Supplementary Note 2). Inset represent the sampled individuals at each site (31 Ust'-Ida I, 8 Bratskii Kamen, 2 Serovo and 5 Shumilkha), with plague detections indicated, and shared graves indicated by shaded areas around individuals. Maps created using Natural Earth Data. **b**, Kernel density estimates plotted within Bayesian models of the

date ranges for the early (red) and late (dark yellow) phases of plague outbreaks at Baikal. Lighter shaded areas correspond to the summed probability distributions prior to modelling; dates used are from individuals identified with plague only. Inset, 95.4% confidence intervals for modelled date ranges at Baikal are shown compared with those from other pre-LNBA plague cases: RV2039 from Latvia<sup>15</sup>, Warberg 1 and Warberg 2 from Germany<sup>61</sup> and FRA102 from Sweden<sup>11</sup>. **c**, Kernel density estimate modelled radiocarbon date distributions for the four cemetery sites, for all radiocarbon-dated post-weaning age humans (or associated deer tooth pendants), irrespective of DNA sampling. Radiocarbon date modelling undertaken with OxCal v.4.4.4<sup>65,66</sup>.

been estimated to have diverged from *Y. pseudotuberculosis* some time in the past 50,000 years (refs. 8,11,15).

Studies of prehistoric plague genomes from Late Neolithic and Bronze Age (LNBA) strains predominantly date to between 4700–2400 cal BP (refs. 3,8,16), and are typically defined as one of two lineages, depending on the presence (LNBA+) or absence (LNBA-) of the *ymt* gene<sup>3</sup>. *ymt* encodes *Yersinia* murine toxin, which enhances bacterial survival in the flea digestive tract during the transition period between rodent and human hosts, and thereby the flea bite-transmitted bubonic form of plague in humans<sup>17</sup>. Lineages of *Y. pestis* that diverged prior to these LNBA clades have also been identified in a handful of Neolithic Swedish individuals (5200–4850 cal BP)<sup>10,11</sup> and a Latvian individual with western hunter-gatherer ancestry (5300–5050 cal BP)<sup>15</sup>. These genomes lack classic virulence genes (*YpfΦ* prophage and *ymt*), although pangenomic analysis revealed the presence of the locus encoding for *Y. pseudotuberculosis*-derived mitogen (YPM), a superantigenic toxin associated with *Y. pseudotuberculosis* (but not later *Y. pestis* strains). This raises intriguing questions about the possible severity of early strains of plague; subsequent LNBA- strains show substantial gene loss, although the virulence potential of these are unknown<sup>3</sup>. Evidence regarding the demographic impact of plague infection on prehistoric populations has so far been lacking in these studies.

Middle Holocene hunter-gatherers around Lake Baikal, southeast Siberia, have been the focus of intensive archaeological study by the Baikal Archaeology Project, yielding important datasets for framing prehistoric hunter-gatherer lifeways<sup>18,19</sup>. These groups demonstrate remarkable continuity of hunter-gatherer lifeways and subsistence, evidenced by an extensive archaeological record of mortuary sites from between about 8500–3500 cal BP (ref. 20). The genomes of sampled hunter-gatherers indicate a long-term continuum of Ancient North

Eurasian and North East Asian ancestry until c. 4500–4000 cal BP (refs. 21,22) (Extended Data Figs. 1 and 2). By this period, cases of plague from human remains corresponding to the LNBA- strain are documented sporadically among Early Bronze Age burials<sup>22,23</sup>. Zoonotic spillover events causing plague infections in this region remain a major health concern to this day<sup>24</sup>. These are principally associated with marmots, the primary zoonotic reservoir of plague in the region<sup>25,26</sup>. To explore health and community structure in prehistoric hunter-gatherer groups, we analysed ancient human and pathogen DNA from four cemetery sites in Cis-Baikal (the lake's western and northern region) across two separate outbreaks dated to 5520–5265 cal BP and 5315–4235 cal BP (95.4% confidence intervals for modelled date ranges based on individuals with detected plague cases, corrected for freshwater reservoir effects; Supplementary Note 4). The long tail for the second outbreak date range (Fig. 1b) is due to this only consisting of two direct dates, although the highest likelihood date range for this is approximately 5050–4850 cal BP.

### Outbreaks of basal plague strains

We generated shotgun-sequenced ancient DNA from 46 Late Neolithic individuals and examined this data for presence of pathogens (Methods). This revealed a conspicuously high occurrence of *Y. pestis* among these individuals, more so than any other pathogen. *Y. pestis* was detected in 18 individuals, indicating 2 distinct phases of outbreaks of plague infection—separated by between 4 and 6 centuries—in 4 cemeteries (Fig. 1a,b). These occur across two phases at Shumilkha, Ust'-Ida I, Bratskii Kamen and Serovo (see Fig. 1c), with cases from Bratskii Kamen in both the first and second phases. These sites are all located on banks of the River Angara, a major watercourse draining from Lake Baikal,

with a rich fishery<sup>27</sup>. Stable carbon and nitrogen isotopic data from individuals at Ust'-Ida I evidence consumption of both local fish and terrestrial game<sup>27</sup>. Burials at Ust'-Ida I and Shumilikha correspond to the Isakovo mortuary tradition (characterized by bodies that are typically oriented parallel to the river, and the presence of grave goods such as mitre-shaped clay vessels, lithic arrowheads and bone or antler points), whereas those at Bratskii Kamen and Serovo correspond to the Serovo mortuary tradition (with bodies frequently oriented perpendicular to the river; bifaces and egg-shaped pots as grave goods are the main features; see also Supplementary Note 1). At Ust'-Ida I, we also detect reads aligned to the zoonotic pathogen *Brucella*, the cause of brucellosis, in one individual (#26.04; Supplementary Note 3). The two plague outbreaks are grouped by the predominant burial practices at each cemetery: Isakovo-style graves in the first outbreak and Serovo-style graves in the second phase (Supplementary Note 1), which are contemporaneous at Lake Baikal between around 6000–5000 cal BP (ref. 27). This period is defined locally as the Late Neolithic, following Siberian archaeological terminology, where the Neolithic is defined on technological criteria such as the introduction of the bow and arrow, clay vessels and stone grinding techniques (domestic plants and animals other than dogs are absent), although these communities remain as hunter-gatherers until the encroachment of pastoralism in the Late Bronze Age. Grave sites comprise the vast majority of the Cis-Baikal archaeological record, and designations such as the Late Neolithic are later categorizations, applied to distinct sets of burial characteristics and grave goods that broadly correspond to different periods. All four cemeteries were also used during the Early Neolithic (7650–6660 cal BP) and Early Bronze Age (4970–3470 cal BP)<sup>27</sup>, although only their Late Neolithic components are considered here.

Pairwise sharing of identity-by-descent (IBD) segments between individuals at these cemeteries indicates recent shared ancestry. Although they are up to 340 km apart, the Angara river would readily have facilitated travel. Very low rates of inbreeding were detected and a high effective population size based on runs of homozygosity was inferred using hapROH (maximum likelihood estimate: 18,219 individuals, 95% confidence interval 9,445–42,062). This is consistent with the scenario of highly mobile, exogamous hunter-gatherer groups.

Within the hunter-gatherer individuals analysed here, the highest number of detected plague infections was at Ust'-Ida I, which is also the largest Isakovo mortuary site in Cis-Baikal. Here, we found a 35% detection rate (11 out of 31 individuals sequenced), including burials #14 and #56.01, for which human genome data were previously reported<sup>21</sup>. Across other sites, we identify one high-coverage plague genome at Shumilikha, four lower coverage genomes from Bratskii Kamen, and one medium coverage genome from Serovo. Overall, we observe a 39% detection rate across Late Neolithic individuals at these cemeteries (from dental cementum). In comparison, quantitative PCR screening of known Mediaeval plague victims at Smithfield, London, UK<sup>28</sup> returned a detection rate of 5.7% from bone and 37% from dental pulp tissue (overall 20%), indicating a high rate of false negative plague detection using ancient DNA. To prevent misrepresentation of data, all ancient individuals with screening data from the affected sites are reported here (human autosomal genome coverage ranges from 0.001× to 1.9×, average 0.65×). Direct radiocarbon dates were obtained from nearly all the individuals within the Late Neolithic components of these cemeteries (a total of 58, including those previously reported from Ust'-Ida I<sup>29</sup>; Supplementary Data 7).

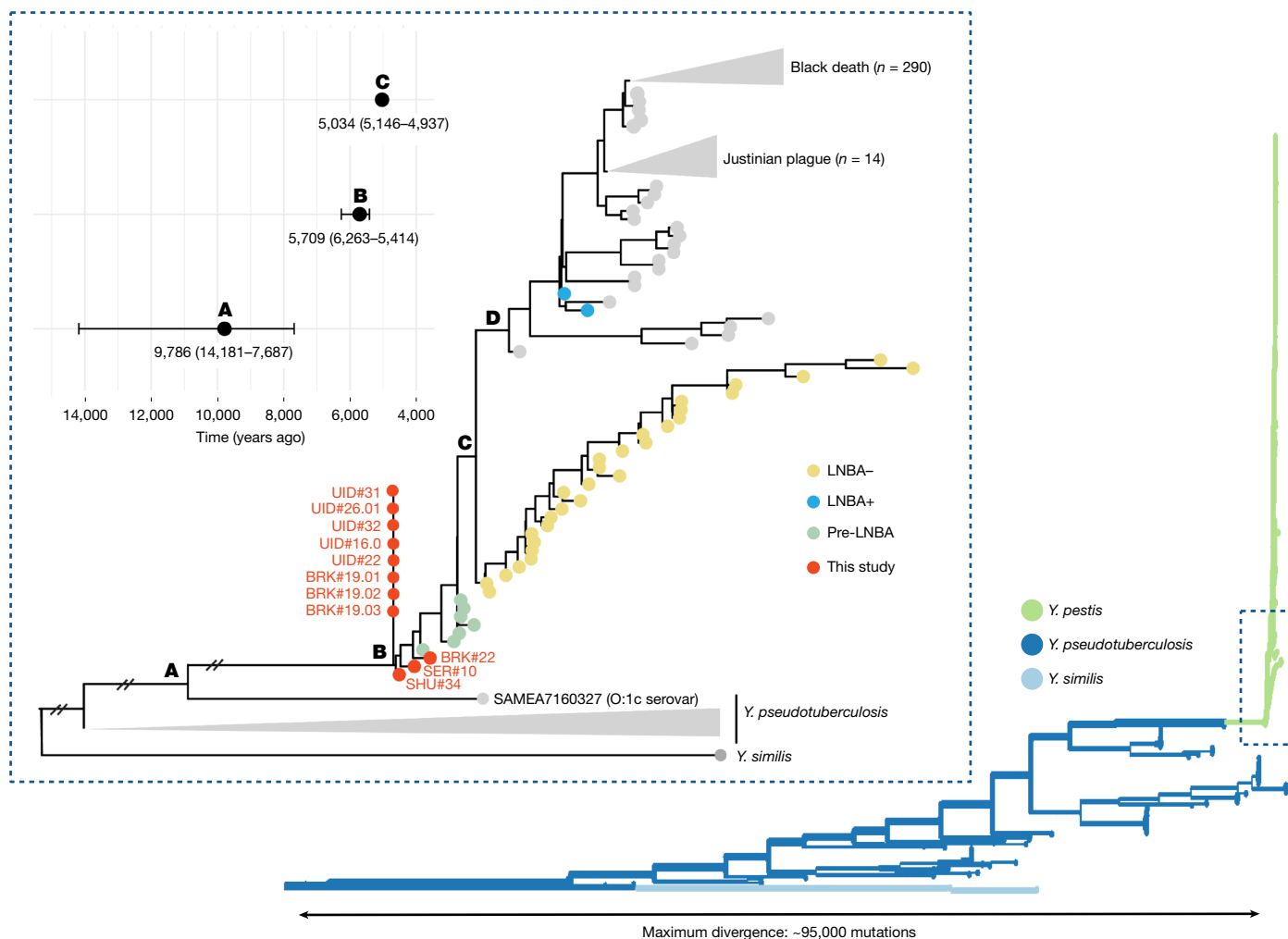
*Y. pestis* genomes identified between the two Baikal phases of outbreaks were found to diverge ancestrally to the current known clade of ancient and modern plague strains (Fig. 2). We confidently assign these to *Y. pestis* from their phylogenetic position, and also the presence of virulence genes and plasmids characteristic of *Y. pestis* (Extended Data Fig. 3 and Supplementary Note 3). This phylogeny was built using genomes obtained from Shumilikha Burial #34 (6.4× coverage) from the first phase, and from Bratskii Kamen Burial #22 (1.6×) and Serovo

Burial #10 (1.0×) from the second phase. Eight lower coverage genomes were phylogenetically placed using USHER<sup>30</sup>. The USHER algorithm finds the most parsimonious placement on the tree, selecting the node with the greatest number of descendants if multiple are equally parsimonious, and ignores missing genotypes. Placement of all low-coverage genomes at the same basal node is partly due to data missingness, although consistent with the position of the three higher coverage Baikal genomes. Bayesian inference of node dates was undertaken following an approach to account for the effects of recombination within bacterial phylogenies<sup>31,32</sup> (Methods and Supplementary Note 3). The emergence of *Y. pestis* as a clonal species of *Y. pseudotuberculosis* occurs some time between the divergence of the lineage that gives rise to *Y. pestis* (labelled node A in Fig. 2) and the most recent common ancestor of available *Y. pestis* genomes (node B in Fig. 2). The upper bound provided by the former is likely to be substantially affected by the paucity *Y. pseudotuberculosis* genomes sequences, and could well be more recent if phylogenetically closer serovars were identified. Nonetheless, this lower bound (with a mean date of 5,709 years ago) revises a previous divergence estimate of 4,810–5,122 years ago<sup>33</sup>, as would be expected by including *Y. pestis* genomes older than this range (other estimates of this have ranged from 6,000 to 50,000 years ago<sup>8</sup> and 7,400 years ago<sup>15</sup>). The phylogeny supports the conclusion that *Y. pestis* first evolved from a variant of the O:1 *Y. pseudotuberculosis* strain (represented by a genome from serotype O:1c, European Nucleotide Archive (ENA) accession: SAMEA7160327), consistent with previous findings<sup>34</sup> reporting the inactivation of the O-antigen gene cluster as a step towards the evolution of *Y. pestis*. Between the two phases, we observe small genetic differences between strains in distinct private mutations in the first and second phase strains (with strict filters for genotype calling; Methods and Supplementary Note 3); this is also clear from the position of nodes in Fig. 2. Although mutation rates in *Y. pestis* are known to be highly variable within different lineages<sup>33</sup>, this result is consistent with a scenario of related strains resulting from separate zoonotic spillover events from a local animal reservoir.

### Baikal hunter-gatherer plague mortality

To contextualize these plague outbreaks, we considered biological kinship patterns, burial treatment and age at death within the affected hunter-gatherer cemeteries. At the site with the highest positive detection of plague (and largest sample), Ust'-Ida I, radiocarbon dates for the Late Neolithic Isakovo component are exceptionally tightly clustered for a relatively large cemetery<sup>29</sup> (Extended Data Fig. 4). Modelled date ranges for all the early phase plague victims indicate a very narrow temporal span, on the order of a few decades (Supplementary Note 4), supporting the scenario that these burials were contemporaneous. This is further corroborated by the high similarity among plague genomes consistent with plague infections occurring in a single outbreak, or over a very brief time span. By reconstructing the most likely familial pedigrees, we find that the relationships and ages of family members are consistent with a mortality event over a time span of less than a single generation (Fig. 3). None of the age at death–relationship pairings indicate, for example, children that reached a similar age to their parents, or siblings and half-siblings with very different ages (the greatest sibling age gap is nine years, separated by a middle sibling). Where multiple generations are present, their inferred age-at-death ranges are generally consistent with those expected if all relatives had died at the same time (for example, a 12–15 year old has a 35–50-year-old father).

The Isakovo mortuary group at Ust'-Ida I is unusual in several other ways among Cis-Baikal hunter-gatherer cemeteries. In addition to the tightly clustered radiocarbon dates, childhood mortality is disproportionately high (also observed at Bratskii Kamen, see Fig. 4), and there is a high incidence of multiple-interment graves (more than half at the site) with no evidence of subsequent grave opening and addition of new burials. This suggests co-occurrence of deaths within shared



**Fig. 2 | Phylogenetic relationships and inferred internal node dates between hunter-gatherer plague samples from this study and previously published data.** Right, the overall topology of the *Y. pseudotuberculosis* species complex is shown from a mutation-annotated tree based on 448 genomes (branch lengths indicate distance by mutations). Inset, a simplified version of this phylogeny, with prehistoric plague strains shown in particular (some branch lengths truncated). The three Baikal samples with higher coverage were incorporated

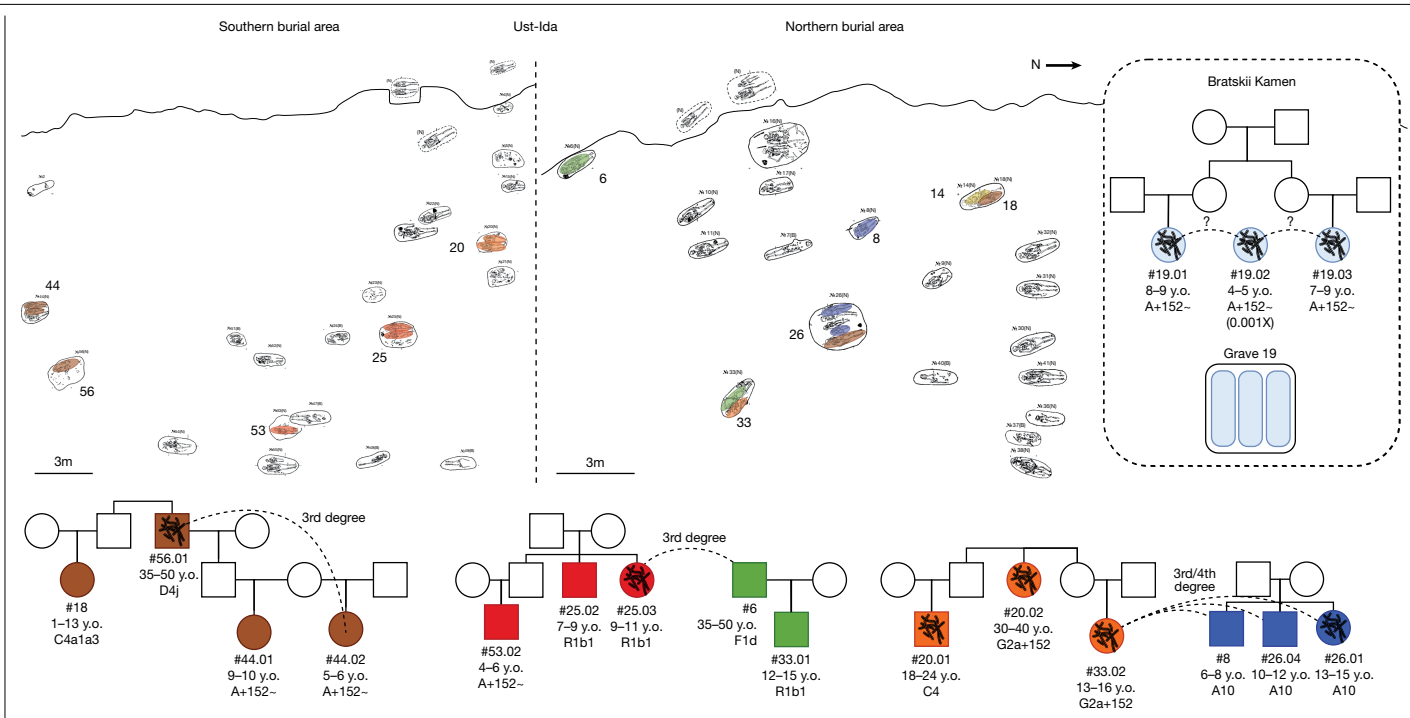
directly into the construction of a RAxML phylogeny, whereas the eight lower coverage samples were phylogenetically placed afterwards, all of which shared their most parsimonious placement at the most basal *Y. pestis* node. Their placement at this node does not constitute a branching position, thus their inclusion adjacent to this node. Top left, internal node dates estimated using BactDating. BRK, Bratskii Kamen; SER, Serovo; SHU, Shumilkha; UID, Ust'-Ida I.

graves, consistent with a catastrophic mortality event. An avuncular relationship was also detected using KIN<sup>35</sup> between the Ust'-Ida I and Shumilkha cemeteries, but this was not substantiated by the expected IBD-sharing pattern (Supplementary Note 2). Nonetheless, the high degree of IBD sharing between individuals (Fig. 1) across a distance of only 37 km along the Angara river, suggests that the concurrent plague outbreaks might be linked to the groups being in close contact at this time point.

In terms of plague detection within grave groups, we find no statistically significant pattern of plague co-occurrence among relatives (Supplementary Note 3), although affected individuals appear to be associated in a number of cases. The burial at Bratskii Kamen features a shared grave of 3 young girls, aged between 4 and 9 years (Fig. 3, left), with similar radiocarbon dates (Supplementary Note 4). Two of them (#19.01 and #19.03) were inferred as third-degree related (most probably cousins); the third had insufficient DNA preservation to confidently infer relatedness, but all three shared a mitochondrial haplotype with three rare private mutations and so were likely to be close maternal relatives. Genome data for *Y. pestis* were identified in all three, suggesting an outbreak of plague infection in a family, with synchronous deaths of the three children. Similarly, at Ust'-Ida I, a nephew and aunt (#20.01

and #20.02) are buried in a shared grave, with *Y. pestis* identified from both (Fig. 3, orange pedigree). The teenage niece of the aunt, however, is buried in a different shared grave with an unrelated teenage male (possibly suggesting non-biological kinship); his father in turn (green pedigree) is buried in an entirely separate grave.

Additionally, some pairs of siblings who are buried together in shared graves show only one individual detected as positive for plague, as is the case with the siblings in grave #25 (Fig. 3, red pedigree). In another example, for a sister (#26.01) and brother (#26.04), the sister is inferred as positive, whereas the brother is not (although *Y. pestis* reads are detected at just below the threshold for confident identification; Supplementary Data). These observations are consistent with a high false negative detection rate in the palaeogenomic analysis of plague<sup>28</sup>. The brother was also infected with probably non-lethal brucellosis (Supplementary Note 3). In several cases close family members are found in different graves within the cemetery, for example, Burial #8, the third sibling of the pair in Grave 26. A pattern is visible, where two closely related family members are buried together, and one or more others are buried further away. This may be consistent with a more drawn-out sequence of deaths instead of a single mortality event if shared graves indicate concurrent deaths, reflecting a scenario of



**Fig. 3 | Familial pedigree groups identified from ancient genomes and the site plan for the Ust'-Ida I cemetery.** Individuals detected for plague are marked with bacilli silhouettes. Pedigrees are drawn from 30 sampled individuals at Ust'-Ida I and 8 from Bratskii Kamen; only close familial relationships are shown;

delayed person-to-person disease transmission. No causes of death were apparent other than genetically detected plague infection (although other microbes detected might reflect bacterial coinfections at the time of death; Supplementary Note 3). Notably, survivors must have existed to bury the deceased, with the typical Isakovo mortuary treatment and grave goods as well as acknowledgement of biological kinship suggesting a more prolonged sequence of mortality events.

### Epidemiological implications

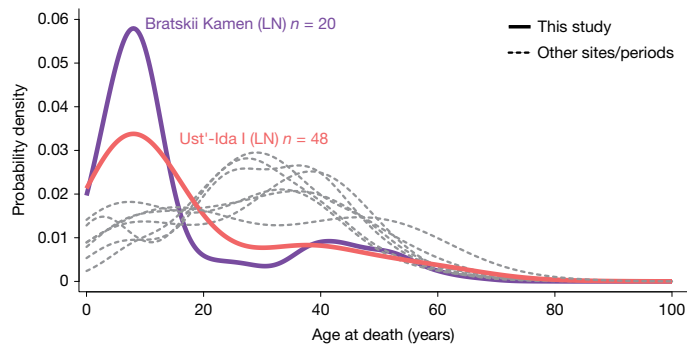
At Baikal, the principal contemporary zoonotic reservoir of plague is the marmot (*Marmota sibirica*), and marmot hunting for meat and fur has historically resulted in perennial plague infections especially in young men, who are exposed during skinning and butchery<sup>36</sup>. Since the nineteenth century, marmots were the most targeted game species by Indigenous hunters in this region, originally by trapping<sup>37</sup>, and there are extensive historical accounts of 'tarbagan plague' from consumption of infected marmots around Lake Baikal<sup>38</sup>. Prehistoric hunter-gatherer marmot procurement is clearly evidenced by the presence of numerous marmot teeth as grave goods in Early Neolithic Kitoi graves<sup>19,39</sup>, although these have not been found in Late Neolithic graves. Consumption of raw or undercooked marmot organs results in the septicaemic form of infection following the faecal-oral transmission route, whereas close contact with marmots infected by present-day *Y. pestis* strains causes bubonic or pneumonic plague (or often both), with the latter often occurring secondarily to septicaemic infection<sup>40</sup> or inhalation of infectious blood droplets during, for example, skinning<sup>41</sup>. The incidence of detected infections among co-buried kin described above would be consistent with the transmission of plague among humans, particularly via pneumonic transmission in the scenario of concurrent deaths.

A striking aspect of the osteological age-at-death data at Ust'-Ida I and Bratskii Kamen—the two cemeteries with multiple instances of plague detected—is that their demographic profiles are highly skewed towards childhood mortality. These both show a peak in mortality

although many other third- or fourth-degree relationships are detected at Ust'-Ida I (see Supplementary Note 2). Pedigrees are reconstructed based on deaths occurring contemporaneously. The two samples from Serovo (not shown) were found to be fourth-degree relatives. y.o., years old.

at the age range of 7.5–11 years—that is, in children before puberty (Supplementary Note 1). In an analysis of mortality profiles across mid-Holocene Cis-Baikal hunter-gatherer cemeteries, these two cemeteries are clearly outliers in terms of the proportion of childhood deaths (Fig. 4). This result was found to be highly statistically significant given a null model of mortality profiles (Supplementary Note 5). Conversely, the 20–25-year age range shows the lowest mortality at Ust'-Ida I, and deaths between 20–35 years of age are completely absent at Bratskii Kamen (Supplementary Fig. 4). Parents are also conspicuously absent from the pedigree groups; although there are many sibling and cousin relationships, there is only one instance of a parent-offspring relationship. The sex ratio in these individuals appears unaffected however (22 XY and 24 XX).

In the context of widespread infection with a plague strain of unknown virulence, this differential mortality between children and adults might be interpreted in a number of ways, given the available bioarchaeological data and current understanding of human immunity. First, adults could largely consist of those who had already been exposed to and recovered from the plague as children, and consequently acquired protective immunity, preventing reinfection or death. This would imply that outbreaks were regularly recurring, which our findings are not able to attest to, and would also imply that older individuals would be more likely to have acquired immunity, yet mortality actually increases slightly after the age range of 20–35 years (after the primary peak around 10 years of age). Alternatively, variation in mortality due to behavioural differences between age groups (for example, division of group tasks or roles by age, resulting in higher childhood exposure to marmots) cannot be ruled out, although there is little analogous precedent for this with regard to marmots specifically, and this is not supported by the lack of heightened childhood mortality in any other Baikal hunter-gatherer cemeteries (Fig. 4). Finally, it is possible that children could be at a greater risk of death owing to inherent differences in immune responses between adults and prepubescent children. Children are known to be more susceptible to infection from Gram-negative bacteria<sup>42</sup>, as evidenced by the



**Fig. 4 | Mortality profiles at Bratskii Kamen and Ust'-Ida I compared with other Baikal hunter-gatherer cemetery populations.** Kernel density plot of modelled age-at-death probabilities, based on a null model of a continuous probability of death at any age (see Supplementary Note 5). All relevant assemblages of human skeletal remains from the Cis-Baikal region studied by A.L. with more than 20 individuals are shown. Sample sizes for sites depicted: Bratskii Kamen (Late Neolithic (LN)),  $n = 20$ ; Khuzhir-Nuge XIV,  $n = 81$ ; Lokomotiv,  $n = 101$ ; Shamanka II,  $n = 156$ ; Shumilikha,  $n = 36$  (Early Bronze Age); Ust'-Ida I,  $n = 48$ ; Verkholensk,  $n = 27$ . Individuals are only included from the predominant period of mortuary use at each site. Outlier individuals from later or earlier mortuary traditions are excluded.

epidemiological profile of *Yersinia enterocolitica* and *Y. pseudotuberculosis* infections today<sup>43</sup>.

### Functional variants in basal *Y. pestis*

The evolution of *Y. pestis* lineages is shaped considerably by processes of gene loss<sup>44</sup>, a pattern typical of pathogenic bacteria in the transitional process to obligate parasitism, which has also been identified across the LNBA–plague strains<sup>3</sup>. From analysis of the coverage of the classic plague virulence genes, we find that virulence genes absent in published LNBA– and pre-LNBA strains from Rinnukalns (RV 2039) and Falbygden<sup>11</sup> are also absent at Baikal (*ymt* and *YpfΦ* prophage; Fig. 5), prohibiting the manifestation of bubonic plague (Extended Data Fig. 3). However, since this virulence gene analysis relies on traditional single-reference mapping, it is restricted to genetic content that is present in the modern reference. To characterize possible ancestral *Y. pseudotuberculosis* variation in the Cis-Baikal strains which might contribute to our interpretation of their pathogenicity, we mapped sequenced reads to a pan-genome variation graph representing genetic diversity across 82 complete assemblies of the *Y. pseudotuberculosis* species complex (56 *Y. pestis*, 24 *Y. pseudotuberculosis* and 1 *Yersinia similis*, based on ref. 11). We found that the two plague strains from Lake Baikal carried similar levels of ancestral *Yersinia* diversity that were only found in *Y. pseudotuberculosis* and *Y. similis* as other pre-LNBA strains (Fig. 5e). For example, we detected the presence of *ypm*, the gene encoding the YPM superantigen known from modern-day *Y. pseudotuberculosis* strains<sup>45</sup>, and recently observed in pre-LNBA and LNBA–plague strains<sup>11</sup>. Three alleles of this gene exist in modern *Y. pseudotuberculosis*: *ypmA*, *ypmB* and *ypmC*, with *ypmA* being regarded as the most virulent form of the gene<sup>46</sup>.

YPM binds to the invariant region of human leukocyte antigen (HLA) class II molecules and interacts with the variable domain of the  $\beta$  chain of the T cell receptor. By bridging HLA class II and a T cell receptor, YPM promotes T cell activation and the release of a range of proinflammatory cytokines, further amplifying the immune response<sup>47,48</sup>.

These YPM-associated immune responses have been suggested to be the cause of various inflammatory complications, including encephalopathy, Far East scarlet-like fever (FESLF; also known as Izumi fever; especially associated with *ypmA*) and a Kawasaki-like syndrome<sup>47,49–51</sup>. Today, FESLF occurs primarily in children aged less than 14 years and

Kawasaki disease occurs mostly in children aged 5 years or less. However, Kawasaki disease following *Y. pseudotuberculosis* infection may also affect older children<sup>52</sup>. These YPM-related inflammatory complications are likely to also have primarily affected prepubescent children in the past, further exacerbating the early *Y. pestis* induced morbidity and mortality in the young.

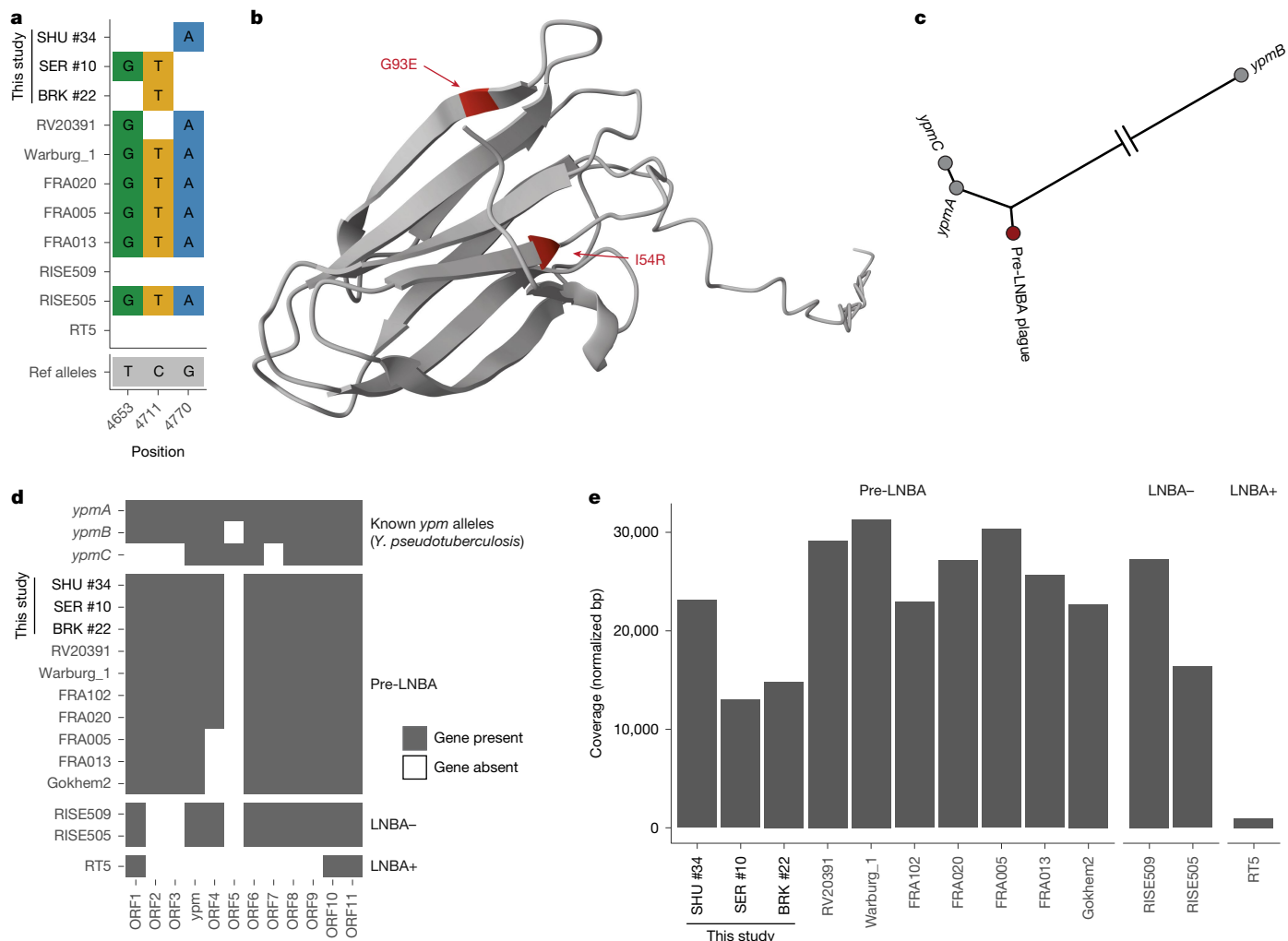
Notably, we find that the *ypm* gene from the two plague strains from Cis-Baikal is closest in sequence similarity to *ypmA*, differing at only 3 base positions: 4653 (T>G (isoleucine>arginine)), 4711 (C>T (synonymous threonine)) and 4770 (G>A (glycine>glutamate); Fig. 5a). These three single nucleotide polymorphisms (SNPs) appear to be fixed in all plague strains in which the gene is present (pre-LNBA and LNBA–strains; Fig. 5a). Since two of these three variants are non-synonymous mutations (I54R and G93E are located in distinct beta-sheets of the YPM structure), they could potentially influence protein secondary, tertiary or quaternary structure, protein–protein interactions, and recognition by the immune system (Fig. 5b). Furthermore, by reconstructing the most likely phylogeny of the three known *ypm* variants together with our data, we found that the *ypmB* version of the gene is highly divergent from *ypmA*, *ypmC* and the ancient plague *ypm*. Under the assumption that the root of the tree is located between *ypmB* and the remaining diversity, the *ypm* from ancient plague appears to diverge ancestrally to both *ypmA* and *ypmC* (Fig. 5c).

In addition, we identified ten open reading frames (ORFs) around the *ypm* locus that are present in the ancestral form of the plague but are absent in later forms. The ORF region is similar to an unstable region of the *Y. pseudotuberculosis* genome with notable low GC content. We found that these ORFs in the Baikal *Y. pestis* genomes were similar to those surrounding the *ypmB* variant as reported in other pre-LNBA strains<sup>11</sup> (Fig. 5d). This pattern with a *ypm* gene similar to *ypmA* combined with a *ypm* locus similar to *ypmB* has—to our knowledge—not been observed previously. A possible explanation could be that this diversity in the Lake Baikal pre-LNBA strains, the most basal of the sampled plague strains, reflects ongoing local adaptation to marmots and other rodent hosts to a greater extent than humans because the regional animal host reservoir presumably at this time far exceeded that of humans. This, so far unique, combination might affect, for example, gene methylation and *ypm* transcription levels.

These genetic features of the Baikal pre-LNBA strains might, together with age-dependent differences in the immune system, partly help explain why prepubescent children predominate among the plague victims, although assessing their actual impact requires functional studies.

### Discussion

Our findings demonstrate that the earliest known outbreaks of plague occurred in prehistoric hunter-gatherers centuries before infections are observed in Neolithic farmers. These outbreaks were probably the result of zoonotic spillover from wild marmot populations at Lake Baikal. These results support a central or northeast Asian origin for plague, whereas previously the earliest samples had only been reported in northern Europe<sup>11,15</sup>. This is in line with estimates based on the analysis of modern *Y. pestis* diversity<sup>53</sup>. Our phylogenetic analysis reveals that these virulent plague strains are temporally relatively close to the most recent common ancestor of *Y. pestis* and *Y. pseudotuberculosis*, possibly indicating rapid diversification with the transfer to rodent hosts from one or several of the other animal hosts of *Y. pseudotuberculosis*. Additionally, this raises questions around the differentiation of taxa within the *Y. pseudotuberculosis* species complex (which includes *Y. pestis* and *Y. similis*) that ancient genome data alone may not be adequate to answer (given that conventional distinctions can be based on pathogenic potential and host range as well). Furthermore, the inferred high lethality of these outbreaks is directly evidenced by mortality profiles and coinciding radiocarbon dates in affected burial sites, indicating that



**Fig. 5 | Variation around the *ypm* locus. a**, SNPs within the *ypm* gene (missing genotypes shown as blank). Ref, reference. **b**, Positions of SNPs (red arrows) on the protein structure of YPMa (Protein Data Bank: 1PM4). This figure was made using data from the AlphaFold database, accession A0A0UIQV71. **c**, Unrooted neighbour-joining tree from the gene sequences of *ypm* variants, including the

pre-LNBA plague form described here. **d**, Presence of open reading frames around the *ypm* locus. **e**, Comparison of ancestral gene content between prehistoric plague strains from reference graph alignments, using normalized breadth of coverage.

children and adolescents were especially vulnerable; these are insights that were previously lacking for prehistoric plague infections. Until now, the earliest detected strains of plague were of uncertain pathogenicity; their virulence has been the focus of considerable debate, based on genetic data alone<sup>3,10,15,16,54</sup>. Here we have integrated multiple lines of evidence from mortuary sites affected by plague (including plague genomes, biological kinship patterns, mortality profiles and modelled radiocarbon date ranges) to characterize what we argue are the lethal consequences of infection during this period.

The context of these outbreaks is important for the interpretation of health and epidemiology in the past. That these outbreaks occur in relatively small, mobile prehistoric hunter-gatherer groups emphasizes that increased population density, animal domestication and lifestyle changes resulting from the Neolithic transition are not necessary conditions for significant zoonotic outbreaks. This further revises interpretations of plague as a unique factor contributing to demographic decline during the Late Neolithic in Europe, as previously suggested<sup>10,11</sup>, especially given the apparent severity of outbreaks identified here. The mortality profile at Baikal also contrasts sharply with the expectation of proponents of the Neolithic epidemiological transition theory that the greatest burden of zoonotic disease in prehistoric hunter-gatherers would fall on producers (20 to 40 year olds)<sup>55</sup>. Our findings further

reveal insights into the social dimension of these communities during outbreaks, evidencing care for the dead (from the co-interment of close relatives and apparently contemporaneous victims), and concurrently detected infections among kin which evince interpersonal contact in life. This evidence of person-to-person transmission contrasts with previous expectations for basal plague strains<sup>15</sup>.

Of note, children seem to have borne the brunt of lethality from plague infections at Cis-Baikal. Different mortality rates among age groups have been observed in historical records of plague outbreaks. Similar to our findings, Parish records from the bubonic plague outbreak in London (UK) in 1603 showed a considerably higher child mortality rate (around 5× higher)<sup>56,57</sup>. One distinct difference between the Cis-Baikal outbreaks and mediaeval bubonic plague epidemic is the probable transmission route. Although airborne and faecal-oral transmission might have occurred in both, flea bite transmission associated with bubonic plague is unlikely in the Cis-Baikal outbreaks (given the absence of *ymt*). Spread of infectious droplets or aerosols through coughing is documented as the primary transmission mode of pneumonic plague<sup>58</sup>, matching our findings for human-to-human spread inferred from the biological kinship and archaeological data. Notably, our results are consistent with previous interpretations that early *Y. pestis* strains could have presented as fatal respiratory pathogens<sup>54</sup>.

Our results indicate that the earliest observed zoonotic spillover was not a one-off event but re-occurred several centuries later, highlighting the prominence of zoonotic infections in prehistoric communities across many different cultural and environmental settings. Additionally, a low-coverage identification of brucellosis suggests evidence of animal-to-human zoonotic transmission in these groups (infection is acquired by direct contact with infected animals<sup>59</sup>, Supplementary Note 3). Recurrent outbreaks of ancestrally diverged plague strains in Cis-Baikal groups between 5500–5000 cal BP further suggests a long history of wild rodents as a perennial reservoir for plague spillover. The subsequent plague strains genetically closest to those at Cis-Baikal are from around 5,000 km west in northern Europe. Given this distance, and that there is little evidence for external contact with non-hunter-gatherer groups at this time, this supports the hypothesis that a substantial, continent-spanning rodent reservoir of *Y. pestis* might account for frequent isolated spillover events in the subsequent millennia, instead of continuous human-to-human transmission. Moreover, the potential association of prehistoric spillover with human procurement of marmots at Cis-Baikal emphasizes the probable key role of rodent species other in the composition of *Y. pestis* reservoirs. Some 352 reservoir species have been identified from present-day surveillance, a number of which are ecologically long established (for example, ground squirrels and gerbils)<sup>60</sup>. The scenario of a persistent prehistoric reservoir aligns both with previous findings of rapid repeated infections of diverged plague strains within the same familial lineage from Neolithic Sweden<sup>11</sup> as well as at Cis-Baikal. We note that recent findings show that early plague cases in domesticates<sup>61,62</sup> may equally be congruous with reverse zoonosis, although given the long history of rodent species reported as reservoirs for plague<sup>60</sup>, we feel that this currently remains the most parsimonious source for these outbreaks.

Together, our findings underscore the universality of zoonotic infection, given the markedly different lifeways of prehistoric hunter-gatherers from European Neolithic farmers. These insights are as relevant for the challenges faced by the world today as they were 5,500 years ago, with 75% of new human pathogens emerging from animal transmission<sup>63</sup>. Insights into the evolutionary history of pathogens across periods of substantial demographic and technological change (before the impact of the Neolithic transition in this case) can provide data to contextualize the major challenges humanity is currently facing, such as the climate change-driven disruption of ecological niches around the world<sup>64</sup>.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-026-10540-5>.

- Barbieri, R. et al. *Yersinia pestis*: the natural history of plague. *Clin. Microbiol. Rev.* <https://doi.org/10.1128/cmr.00044-19> (2020).
- Spyrou, M. A. et al. Analysis of 3800-year-old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic plague. *Nat. Commun.* **9**, 2234 (2018).
- Andrades Valtueña, A. et al. Stone Age *Yersinia pestis* genomes shed light on the early evolution, diversity, and ecology of plague. *Proc. Natl Acad. Sci. USA* **119**, e2116722119 (2022).
- Key, F. M. et al. Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization process. *Nat. Ecol. Evol.* **4**, 324–333 (2020).
- Mühlemann, B. et al. Ancient hepatitis B viruses from the Bronze Age to the Medieval period. *Nature* **557**, 418–423 (2018).
- Harbeck, M. et al. *Yersinia pestis* DNA from skeletal remains from the 6th century AD reveals insights into Justinianic Plague. *PLoS Pathog.* **9**, e1003349 (2013).
- Spyrou, M. A. et al. The source of the Black Death in fourteenth-century central Eurasia. *Nature* **606**, 718–724 (2022).
- Rasmussen, S. et al. Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell* **163**, 571–582 (2015).
- Munshi-South, J., Garcia, J. A., Orton, D. & Phifer-Rixey, M. The evolutionary history of wild and domestic brown rats (*Rattus norvegicus*). *Science* **385**, 1292–1297 (2024).
- Rascovan, N. et al. Emergence and spread of basal lineages of *Yersinia pestis* during the Neolithic Decline. *Cell* **176**, 295–305.e10 (2019).
- Seersholm, F. V. et al. Repeated plague infections across six generations of Neolithic farmers. *Nature* **632**, 114–121 (2024).
- Shennan, S. et al. Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nat. Commun.* **4**, 2486 (2013).
- Colledge, S., Conolly, J., Crema, E. & Shennan, S. Neolithic population crash in northwest Europe associated with agricultural crisis. *Quat. Res.* **92**, 686–707 (2019).
- Hama, H. O. et al. *Yersinia pestis* infection is not synonymous with deadly plague in Neolithic Scandinavia. *Am. J. Biol. Anthropol.* **189**, e70200 (2025).
- Susat, J. et al. A 5,000-year-old hunter-gatherer already plagued by *Yersinia pestis*. *Cell Rep* **35**, 109278 (2021).
- Andrades Valtueña, A. et al. The stone age plague and its persistence in Eurasia. *Curr. Biol.* **27**, 3683–3691.e8 (2017).
- Bland, D. M., Miarinjara, A., Bosio, C. F., Calarco, J. & Hinnebusch, B. J. Acquisition of *Yersinia murine* toxin enabled *Yersinia pestis* to expand the range of mammalian hosts that sustain flea-borne plague. *PLoS Pathog.* **17**, e1009995 (2021).
- Weber, A. W., Katzenberg, M. A. & Schurr, T. A. *Prehistoric Hunter-Gatherers of the Baikal Region, Siberia* (University of Pennsylvania Press, 2011).
- Weber, A. W., Bazaliiskii, V. I. & Jessup, E. (eds.) *Shamanka II: An Early Neolithic Cemetery on the Southwest Shore of Lake Baikal, Siberia: Archaeological and Osteological Materials* (Deutsches Archäologisches Institut, 2024).
- Weber, A. W. Middle Holocene hunter-gatherers of Cis-Baikal, Eastern Siberia: combined impacts of the boreal forest, bow-and-arrow, and fishing. *Archaeol. Res. Asia* **24**, 100222 (2020).
- Damgaard, P. D. B. et al. The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, eaar7711 (2018).
- Yu, H. et al. Paleolithic to Bronze Age Siberians reveal connections with First Americans and across Eurasia. *Cell* **181**, 1232–1245.e20 (2020).
- Kılınc, G. M. et al. Human population dynamics and *Yersinia pestis* in ancient northeast Asia. *Sci. Adv.* **7**, eabc4587 (2021).
- Xu, L. et al. Climate-driven marmot-plague dynamics in Mongolia and China. *Sci. Rep.* **13**, 11906 (2023).
- He, Z. et al. Distribution and characteristics of human plague cases and *Yersinia pestis* isolates from 4 Marmota plague foci, China, 1950–2019. *Emerg. Infect. Dis.* **27**, 2544–2553 (2021).
- Toole, R., Swamy, K., Moa, A. & Quigley, A. An overview of the first human case of bubonic plague detected in Bugat settlement, Gobi-Altai province in Mongolia. *Glob. Biosecurity* **6**, e246 (2024).
- Weber, A. W., Ramsey, C. B., Schulting, R. J., Bazaliiskii, V. I. & Goriunova, O. I. Middle Holocene hunter-gatherers of Cis-Baikal, Eastern Siberia: chronology and dietary trends. *Archaeol. Res. Asia* **25**, 100234 (2021).
- Schuenemann, V. J. et al. Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of *Yersinia pestis* from victims of the Black Death. *Proc. Natl Acad. Sci. USA* **108**, E746–52 (2011).
- Bronk Ramsey, C., Schulting, R. J., Bazaliiskii, V. I., Goriunova, O. I. & Weber, A. W. Spatio-temporal patterns of cemetery use among Middle Holocene hunter-gatherers of Cis-Baikal, Eastern Siberia. *Archaeol. Res. Asia* **25**, 100253 (2021).
- Turakhia, Y. et al. Ultrafast sample placement on existing tRees (USHER) enables real-time phylogenetics for the SARS-CoV-2 pandemic. *Nat. Genet.* **53**, 809–816 (2021).
- Croucher, N. J. et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15 (2015).
- Didelot, X., Croucher, N. J., Bentley, S. D., Harris, S. R. & Wilson, D. J. Bayesian inference of ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Res.* **46**, e134 (2018).
- Eaton, K. et al. Plagued by a cryptic clock: insight and issues from the global phylogeny of *Yersinia pestis*. *Commun. Biol.* **6**, 23 (2023).
- Skurnik, M., Peippo, A. & Ervelä, E. Characterization of the O-antigen gene clusters of *Yersinia pseudotuberculosis* and the cryptic O-antigen gene cluster of *Yersinia pestis* shows that the plague bacillus is most closely related to and has evolved from *Y. pseudotuberculosis* serotype O:1b. *Mol. Microbiol.* **37**, 316–330 (2000).
- Popli, D., Peyrègne, S. & Peter, B. M. KIN: a method to infer relatedness from low-coverage ancient DNA. *Genome Biol.* **24**, 10 (2023).
- Galdan, B., Baatar, U., Molotov, B. & Dashdavaa, O. Plague in Mongolia. *Vector Borne Zoonotic Dis.* **10**, 69–75 (2010).
- Saveljev, A., Soloviev, V., Scopin, A., Shar, S. & Otgonbaatar, M. Contemporary significance of hunting and game animals use in traditional folk medicine in north-west Mongolia and adjacent Tuva. *Balk. J. Wildl. Res.* **1**, 76–81 (2014).
- Lynteris, C. The native knowledge hypothesis. in *Ethnographic Plague: configuring disease on the Chinese-Russian frontier* 43–68 (Palgrave Macmillan, 2016).
- Losey, R. J., Ivanov, A. V., Palesskiy, S. V. & Bazaliiskii, V. I. Geochemical analyses of marmot teeth to evaluate the potential for overlapping foraging ranges in two Siberian human cemetery populations. *J. Ethnobiol.* **36**, 493–511 (2016).
- Kehrmann, J. et al. Two fatal cases of plague after consumption of raw marmot organs. *Emerg. Microbes Infect.* **9**, 1878–1880 (2020).
- Wong, D. et al. Primary pneumonic plague contracted from a mountain lion carcass. *Clin. Infect. Dis.* **49**, e33–e38 (2009).
- Al-Hasan, M. N., Huskins, W. C., Lahr, B. D., Eckel-Passow, J. E. & Baddour, L. M. Epidemiology and outcome of Gram-negative bloodstream infection in children: a population-based study. *Epidemiol. Infect.* **139**, 791–796 (2011).
- European Centre for Disease Prevention and Control. *Yersiniosis: Annual Epidemiological Report for 2018* (ECDC, 2019).
- Chain, P. S. G. et al. Insights into the evolution of *Yersinia pestis* through whole-genome comparison with *Yersinia pseudotuberculosis*. *Proc. Natl Acad. Sci. USA* **101**, 13826–13831 (2004).
- Carnoy, C. et al. The superantigen gene *ypm* is located in an unstable chromosomal locus of *Yersinia pseudotuberculosis*. *J. Bacteriol.* **184**, 4489–4499 (2002).

46. Carnoy, C., Mullet, C., Müller-Alouf, H., Leteurtre, E. & Simonet, M. Superantigen YPMa exacerbates the virulence of *Yersinia pseudotuberculosis* in mice. *Infect. Immun.* **68**, 2553–2559 (2000).
47. Goubard, A. et al. Superantigenic *Yersinia pseudotuberculosis* induces the expression of granzymes and perforin by CD4<sup>+</sup> T cells. *Infect. Immun.* **83**, 2053–2064 (2015).
48. Chung, L. K. & Bliska, J. B. *Yersinia* versus host immunity: how a pathogen evades or triggers a protective response. *Curr. Opin. Microbiol.* **29**, 56–62 (2016).
49. Sato, K., Ouchi, K. & Taki, M. *Yersinia pseudotuberculosis* infection in children, resembling Izumi fever and Kawasaki syndrome. *Pediatr. Infect. Dis.* **2**, 123–126 (1983).
50. Tahara, M., Baba, K., Waki, K. & Arakaki, Y. Analysis of Kawasaki disease showing elevated antibody titres of *Yersinia pseudotuberculosis*. *Acta Paediatr.* **95**, 1661–1664 (2006).
51. Vincent, P., Salo, E., Skurnik, M., Fukushima, H. & Simonet, M. Similarities of Kawasaki disease and *Yersinia pseudotuberculosis* infection epidemiology. *Pediatr. Infect. Dis. J.* **26**, 629–631 (2007).
52. Ono, Y. et al. Kawasaki disease caused by *Yersinia pseudotuberculosis* infection. *QJM* **113**, 679–680 (2020).
53. Morelli, G. et al. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat. Genet.* **42**, 1140–1143 (2010).
54. Zimble, D. L., Schroeder, J. A., Eddy, J. L. & Latham, W. W. Early emergence of *Yersinia pestis* as a severe respiratory pathogen. *Nat. Commun.* **6**, 7487 (2015).
55. Armelagos, G. J., Goodman, A. H. & Jacobs, K. H. The origins of agriculture: Population growth during a period of declining health. *Popul. Environ.* **13**, 9–22 (1991).
56. Hollingsworth, M. F. & Hollingsworth, T. H. Plague mortality rates by age and sex in the parish of St. Botolph's without Bishopsgate, London, 1603. *Popul. Stud.* **25**, 131–146 (1971).
57. Evans, C. M. & Evans, A. E. Plague — a disease of children and servants? A study of the parish records of St Peter upon Cornhill, London from 1580 to 1605. *Contin. Chang.* **34**, 183–208 (2019).
58. Wu, L. T. *A Treatise on Pneumonic Plague* (Publications of the League of Nations, 1926).
59. Franco, M. P., Mulder, M., Gilman, R. H. & Smits, H. L. Human brucellosis. *Lancet Infect. Dis.* **7**, 775–786 (2007).
60. Mahmoudi, A. et al. Plague reservoir species throughout the world. *Integr. Zool.* **16**, 820–833 (2021).
61. Susat, J. et al. Neolithic *Yersinia pestis* infections in humans and a dog. *Commun. Biol.* **7**, 1013 (2024).
62. Light-Maka, I. et al. Bronze Age *Yersinia pestis* genome from sheep sheds light on hosts and evolution of a prehistoric plague lineage. *Cell* **188**, 5748–5762.e18 (2025).
63. Jones, K. E. et al. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
64. Carlson, C. J. Climate change increases cross-species viral transmission risk. *Nature* **607**, 555–562 (2022).
65. Bronk Ramsey, C. OxCal v4.4. <https://c14.arch.ox.ac.uk/oxcal.html> (2021).
66. Reimer, P. J. et al. The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

## Methods

### Laboratory work

Ancient DNA was extracted from the dental cementum of molar or premolar teeth from archaeological skeletal remains studied by the Baikal Archaeology Project. Sampling for ancient DNA (aDNA) was undertaken in dedicated clean laboratory facilities at the Lundbeck Foundation Centre for GeoGenetics (Copenhagen) and at the Institute of Archaeology, University College London (London). Cementum was isolated specifically from the roots of teeth<sup>67</sup> using a sterilized handheld rotary saw, and pulverized prior to demineralization and enzymatic digestion. Sampled aliquots were approximately 50–100 mg of material. Extraction, purification and library preparation of aDNA for shotgun sequencing followed the approach described in Allentoft et al.<sup>68</sup>, using a double-stranded library protocol following Margaryan et al.<sup>69</sup> in the first instance, and the ‘Santa Cruz Reaction’ single-stranded library protocol<sup>70</sup> for samples with low template DNA content. Concentrations for resulting libraries were obtained using an Agilent FragmentAnalyzer and pooled at equimolar concentration for sequencing on Illumina NovaSeq 6000 S4 flowcells (100 bp paired-end reads) at the GeoGenetics Sequencing Core (Copenhagen). All samples were screened without partial Uracil-DNA Glycosylase (UDG) treatment, and in some cases subsequent additional libraries were built with UDG treatment (following<sup>71</sup>).

For libraries where screening sequencing indicated the presence of *Y. pestis* DNA, in-solution capture enrichment was undertaken. Hybridization capture was performed using the Arbor Sciences myBaits kit following Wagner et al.<sup>72</sup>, using the manufacturer’s High Sensitivity protocol, but only with a single round of enrichment. Pooled libraries from the capture reactions were then re-amplified for 16 cycles and sequenced on the same platform as above.

### Preliminary bioinformatics

Following base-calling of Illumina data using CASAVA (v.1.8.2)<sup>73</sup>, adapter sequences and polyN tails were trimmed from demultiplexed fastq files using AdapterRemoval (v.2.0). Reads were aligned to the human reference genome GRCh38 using bwa aln (v.0.7.18)<sup>74</sup> (reference genome hg19 was also used for hapRoH analysis, see below). Aligned reads were converted to BAM files, merged across libraries at sample level, sorted, filtered and indexed using Samtools (v.1.21)<sup>75</sup>, then duplicates identified using MarkDuplicates from Picard (v2.18.7), with the following options in place: ‘OPTICAL\_DUPLICATE\_PIXEL\_DISTANCE = 12000 REMOVE\_DUPLICATES = false TAGGING\_POLICY = All VALIDATION\_STRINGENCY = LENIENT’. Duplicate reads were then filtered out using Samtools alongside reads with a mapping quality of <30. Summary statistics for sequencing depth and coverage were generated using BEDtools (v2.23.0)<sup>76</sup> and pysam (<https://github.com/pysam-developers/pysam>). Estimation of human DNA contamination and damage patterns were performed at a library level, using contamMix<sup>77</sup>, ANGSD (v.0.940)<sup>78</sup>, and mapDamage2.0<sup>79</sup>.

### Human DNA Analysis

Chromosomal sex was inferred based on the ratio of Y and X chromosome aligned reads, following existing confidence intervals<sup>80</sup>. Chromosomal aneuploidies were not detected. Mitochondrial haplogroups were assigned using haplogrep (v.2.4.0)<sup>81</sup> following annotation of variants using mutserve (v.1.3.0)<sup>82</sup>. Y chromosomal haplogroups were assigned following the approach in ref. 11.

For exploratory analysis of ancestry through principal component analysis (PCA), pseudohaploid genotypes were called by randomly selecting a variant from a pileup generated with Samtools. Samples were then projected into the variation space obtained from using smartpca<sup>83</sup> to undertake PCA on 2,086,279 SNPs (filtered for transversions only and with minor allele frequency >0.1%) from a reference panel of ancient Eurasian populations<sup>68</sup>. The latter was lifted over from hg19

using hgLiftOver (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>), and the effects of liftover evaluated (Extended Data Fig. 2). All reported samples were included in the PCA by projection.

Diploid genotypes were called using bcftools (v.1.21), and for the analysis of IBD segment sharing, missing diploid genotypes were imputed using GLIMPSE<sup>84</sup> (for samples with a minimum autosomal genome coverage of 0.1×) following the approach in ref. 68, and IBD segments called using IBDseq (v.r1206)<sup>85</sup>, followed by genetic clustering by IBD<sup>68</sup>.

Runs of homozygosity were detected from homozygous-by-descent segments obtained from IBDseq, and from pseudohaploid data subset to the 1240k SNP positions using hapRoH (v.1)<sup>86</sup>. Biological kinship was inferred using KIN<sup>35</sup>, initially running KINgaroo on filtered BAM files targeting the 2,086,279 SNPs described above, and then validated based on IBD sharing inferred from IBDseq. Pedigrees were then reconstructed taking into account the resulting log-likelihood estimates for kinship scenarios, uniparental haplotypes, sex and age-at-death (Supplementary Note 2).

### Screening for pathogen taxa

Shotgun sequencing data generated from dental cementum was screened for the presence of known human pathogens using the pathopipe workflow (<https://github.com/martinsikora/pathopipe/>) detailed in Sikora et al.<sup>87</sup>. Reads were classified using a fast k-mer approach, KrakenUniq<sup>88</sup> (v.0.5.8), based on a custom database of human pathogens and environmental microbes. For each genus identified in each sample, pairwise alignments using bowtie2<sup>89</sup> (v.2.5.4) were made for all reads classified to that genus against all available species reference genomes for the same genus. Assignments are then made for the presence of pathogen taxa on the basis of the following detection thresholds: unique read count >30, k-mer rank = 1, corrected coverage ratio >0.5, average nucleotide identity >0.97, average number of soft clipped bases <8, based on those applied by Seersholm et al.<sup>11</sup>.

### *Y. pestis* DNA analysis

For samples where plague infection was identified through the pathogen screening pipeline, we carried out traditional single-reference mapping with Bowtie2<sup>89</sup> against the plague reference genome (CO92; GCA\_000009065.1), with the parameters ‘-D 20 -R 3 -N 1 -L 20 -i S,1,0.50-end-to-end-no-unal’. Next, duplicate reads and low mapping quality reads (MQ < 30) were removed with samtools<sup>75</sup>, followed by calculations of the average depth of coverage in each sample using BEDtools genomcov<sup>76</sup>. We characterized plague cases based on their coverage as either: tentative detections (<0.01×), lower-coverage plague cases (0.01–1×) or higher-coverage plague genomes (>1×).

For the three higher coverage genomes we called genotypes in a sample-wise manner using HaplotypeCaller from GATK<sup>90</sup>, followed by a subsequent step of joint haplotype calling using GenotypeGVCFs on the merged dataset. Using VariantFiltration (GATK) we removed low-confidence calls of either: low genotype quality (<50), an allele balance of less than 0.9, a read depth of less than 3 or a read depth higher than 1,000. Next, we converted the dataset to multifasta format using bcftools consensus. In doing so, we applied a mask across regions containing the highest proportions of zero mapping quality reads, typically localized in repetitive regions, following the approach in Seersholm et al.<sup>11</sup>. To maintain the coordinates of the reference genome and for consistency with the other aligned samples, sequencing data for published reference genomes for *Y. pseudotuberculosis* and *Y. similis* were downloaded from ENA and the reads were realigned to the *Y. pestis* reference genome GCA\_000009065. Genotypes were then called and filtered as described above, yielding a multiple sequence alignment for the full *Yersinia* chromosome with 448 sequences on the coordinates of the reference genome. A phylogenetic tree was inferred from the full alignment file including all reference sequences and the three high-coverage samples from this study using RAXML-NG<sup>91</sup> with the GTR + G substitution model and using the *Y. similis* reference

genome (SAMEA5779183) as an outgroup (see Extended Data Fig. 3). We converted the multiple sequence alignment to a haploid VCF using faToVcf<sup>30</sup> with the *Y. pestis* reference genome NC\_003143.1.fa as a reference, then built a mutation-annotated tree object from this VCF and the RAxML phylogeny using UShER<sup>30</sup>. This new UShER phylogeny maintains the original topology but directly assigns substitutions in the VCF to branches on the tree using the Fitch-Sankoff algorithm<sup>92,93</sup>, so that edge lengths are in units of real substitutions. We used matUtils<sup>94</sup> to extract a .json file from the mutation-annotated tree protobuf, available interactively online at [https://bit.ly/Ypestis\\_MAT](https://bit.ly/Ypestis_MAT).

For the 8 lower coverage samples, we called a SNP-only vcf using bcftools, filtering for a minimum mapping quality of 30 and bases, a minimum base quality of 30, and a maximum depth of 1,000. We only kept sites which were variable in the reference panel or in which more than one lower coverage sample had a variant called, and further removed all variant sites in the low mapping quality mask described above, yielding a filtered vcf for the lower coverage samples. This low-coverage vcf was used to phylogenetically place the low-coverage samples into the mutation-annotated tree using UShER. All 8 lower coverage samples shared a single, maximally parsimonious placement at the root node of the *Y. pestis* clade.

Finally, we ran Gubbins (v.3.4.3)<sup>31</sup> on a multifasta of all LNBA- and pre-LNBA genomes, as well as the O:1c serovar genome (SAMEA7160327) and the SAMN03121000 genome as outgroup (see Supplementary Fig. 14). To date the phylogeny, while taking recombination into account, we ran BactDating<sup>32</sup> on the output tree from Gubbins based on non-recombinant variation. We used 100,000 iterations and a relaxed gamma model as suggested in ref. 33. Convergence was confirmed through the trace file as shown in the Supplement. We report median age estimates and 95% confidence intervals for nodes of interest in Fig. 2.

### Variation graph analysis

To characterize the full diversity of ancient plague we built a pan-genome variation graph of all known diversity within the *Y. pseudotuberculosis* species complex (*Y. pestis*, *Y. pseudotuberculosis* and *Y. similis*). We used Pangenome Graph Builder (pggb)<sup>95</sup> on all available *Y. pseudotuberculosis* complex assemblies from NCBI with assembly level characterized as either 'chromosome' or 'complete'. To ensure correct construction of the graph around the plasmids, we built separate graphs for the chromosome and the plasmids and merged these afterwards using vg tools<sup>96</sup>. Next, we indexed the graph and carried out Giraffe<sup>97</sup> short read mapping to the variation graph of the data from this study and all publicly available ancient shotgun data. Lastly, we identified graph nodes present in the Lake Baikal plague strains but absent in all modern plague assemblies and classified these based on their presence/absence pattern in *Y. pseudotuberculosis* and *Y. similis*. Each node was classified as either ancestral (present in both *Y. pseudotuberculosis* and *Y. similis*) or either *Y. pseudotuberculosis*-derived or *Y. similis*-derived.

### Age-at-death estimation

Age-at-death estimation was based on a variety of established anthropological methods. For non-adult individuals (generally <20 years), it was assessed through dental formation and eruption, epiphyseal and long bone diaphyseal measurements, and epiphyseal union, as summarized in refs. 49–51,98,99. Adult age estimation focused on skeletal morphological changes, namely those of the pubic symphysis<sup>100,101</sup> and iliac auricular surface<sup>102–104</sup>, but also palatine and ectocranial suture closure<sup>105–107</sup>. For all individuals, as many methods as possible were considered based on the state of skeletal and/or dental preservation.

### Radiocarbon dating

Radiocarbon dating was undertaken at the Oxford Radiocarbon Accelerator Unit following an established protocol at that facility<sup>108</sup>. New determinations from Bratskii Kamen, Serovo and Shumilkha are

presented here for the first time, alongside previously published radiocarbon dates<sup>27,109</sup> (Extended Data Fig. 4 and Supplementary Table 1). All human dates are corrected for the freshwater reservoir effect (FRE), using the regression equation for southwest Baikal/Angara<sup>110,111</sup> (Supplementary Note 4). A small number of dates on red deer tooth pendants are preferred over human dates where available, as they avoid the FRE. Bayesian modelling of the radiocarbon dates was undertaken in OxCal 4.4<sup>65</sup>, using non-informative, single-phase models with uniform boundaries. To visualize summed multiple dates, kernel density estimation (KDE) models and plots within Bayesian models were employed<sup>112</sup>.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Raw sequencing data generated for the analysis of ancient genomes here is available on European Nucleotide Archive under accession PRJEB111316. Aligned sequences and the mutation annotated tree described here can be accessed at [https://github.com/ramacleod/Prehistoric\\_plague\\_MAT](https://github.com/ramacleod/Prehistoric_plague_MAT).

- Hansen, H. B. et al. Comparing ancient DNA preservation in petrous bone and tooth cementum. *PLoS ONE* **12**, e0170940 (2017).
- Allentoft, M. E. et al. Population genomics of post-glacial western Eurasia. *Nature* **625**, 301–311 (2024).
- Margaryan, A. et al. Population genomics of the Viking world. *Nature* **585**, 390–396 (2020).
- Kapp, J. D., Green, R. E. & Shapiro, B. A fast and efficient single-stranded genomic library preparation method optimized for ancient DNA. *J. Hered.* **112**, 241–249 (2021).
- Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Philos. Trans. R. Soc. B* **370**, 20130624 (2015).
- Wagner, D. M. et al. *Yersinia pestis* and the plague of Justinian 541–543 AD: a genomic analysis. *Lancet Infect. Dis.* **14**, 319–326 (2014).
- Hosseini, P., Tremblay, A., Matthews, B. F. & Alkharouf, N. W. An efficient annotation and gene-expression derivation tool for Illumina Solexa datasets. *BMC Res. Notes* **3**, 183 (2010).
- Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
- Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- Fu, Q. et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).
- Korneliusson, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014).
- Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
- Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).
- Weissensteiner, H. et al. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–63 (2016).
- Weissensteiner, H. et al. Contamination detection in sequencing studies using the mitochondrial phylogeny. *Genome Res.* **31**, 309–316 (2021).
- Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
- Rubinacci, S., Ribeiro, D. M., Hofmeister, R. J. & Delaneau, O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat. Genet.* **53**, 120–126 (2021).
- Browning, B. L. & Browning, S. R. Detecting identity by descent and estimating genotype error rates in sequence data. *Am. J. Hum. Genet.* **93**, 840–851 (2013).
- Ringbauer, H., Novembre, J. & Steinrücken, M. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. *Nat. Commun.* **12**, 5425 (2021).
- Sikora, M. et al. The spatiotemporal distribution of human pathogens in ancient Eurasia. *Nature* **643**, 1011–1019 (2025).
- Breitwieser, F. P., Baker, D. N. & Salzberg, S. L. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. *Genome Biol.* **19**, 198 (2018).
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
- Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10.33 (2013).
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).

92. Fitch, W. M. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Biol.* **20**, 406–416 (1971).
93. Sankoff, D. Minimal mutation trees of sequences. *SIAM J. Appl. Math.* **28**, 35–42 (1975).
94. McBroome, J. et al. A daily-updated database and tools for comprehensive SARS-CoV-2 mutation-annotated trees. *Mol. Biol. Evol.* **38**, 5819–5824 (2021).
95. Garrison, E. et al. Building pangenome graphs. *Nat. Methods* **21**, 2008–2012 (2024).
96. Garrison, E. et al. Variation graph toolkit improves read mapping by representing genetic variation in the reference. *Nat. Biotechnol.* **36**, 875–879 (2018).
97. Sirén, J. et al. Pangenomics enables genotyping of known structural variants in 5202 diverse genomes. *Science* **374**, abg8871 (2021).
98. Buikstra, J. E. & Ubelaker, D. H. *Standards for Data Collection from Human Skeletal Remains* (Arkansas Archaeological Survey, 1994).
99. Schaefer, M., Black, S. & Scheuer, L. *Juvenile Osteology, A Laboratory and Field Manual* (Academic Press, 2009).
100. Brooks, S. & Suchey, J. M. Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Hum. Evol.* **5**, 227–238 (1990).
101. Katz, D. & Suchey, J. M. Age determination of the male os pubis. *Am. J. Phys. Anthropol.* **69**, 427–435 (1986).
102. Buckberry, J. L. & Chamberlain, A. T. Age estimation from the auricular surface of the ilium: a revised method. *Am. J. Phys. Anthropol.* **119**, 231–239 (2002).
103. Lovejoy, C. O., Meindl, R. S., Pryzbeck, T. R. & Mensforth, R. P. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. *Am. J. Phys. Anthropol.* **68**, 15–28 (1985).
104. Meind, R. S. & Lovejoy, C. O. in *Age Markers in the Human Skeleton* (ed. Iscan, M. Y.) 137–168 (Charles C. Thomas, 1989).
105. Mann, R. W., Symes, S. A. & Bass, W. M. Maxillary suture obliteration: aging the human skeleton based on intact or fragmentary maxillae. *J. Forensic Sci.* **32**, 148–157 (1987).
106. Krogman, W. & Iscan, M. Y. *The Human Skeleton in Forensic Medicine* (Charles C Thomas, 1986).
107. Meind, R. S. & Lovejoy, C. O. Ectocranial suture closure: a revised method for the determination of skeletal age at death based on the lateral-anterior sutures. *Am. J. Phys. Anthropol.* **68**, 57–66 (1985).
108. Brock, F., Higham, T., Ditchfield, P. & Ramsey, C. B. Current pretreatment methods for AMS radiocarbon dating at the Oxford Radiocarbon Accelerator Unit (Orau). *Radiocarbon* **52**, 103–112 (2010).
109. Weber, A. W. et al. Chronology of middle Holocene hunter-gatherers in the Cis-Baikal region of Siberia: corrections based on examination of the freshwater reservoir effect. *Quaternary* **419**, 74–98 (2016).
110. Schulting, R. J., Ramsey, C. B., Bazaliiskii, V. I., Goriunova, O. I. & Weber, A. Freshwater reservoir offsets investigated through paired human–faunal <sup>14</sup>C dating and stable carbon and nitrogen isotope analysis at Lake Baikal, Siberia. *Radiocarbon* **56**, 991–1008 (2014).
111. Schulting, R. J. et al. Freshwater reservoir effects in Cis-Baikal: an overview. *Archaeol. Res. Asia* **29**, 100324 (2022).
112. Ramsey, C. B. Methods for summarizing radiocarbon datasets. *Radiocarbon* **59**, 1809–1833 (2017).

**Acknowledgements** We thank E. Tilby for assistance in drafting figures; M. Madrona, M. Hjorth, A. Poersksen and L. Kjærsgaard Hansen for assistance with laboratory work; A. Hiob for administration of the Baikal Archaeology Project; P. Selmer Olsen for technical assistance;

A. Razeto and L. Olsen for administration of the Lundbeck Foundation GeoGenetics Centre; S. Shennan, J. Parkhill and R. Durbin for insightful discussions; and N. N. Mamonova for providing additional data for age-at-death estimations. R.M., F.V.S., J.T.S., C.G., L.V., Y.W. M.S. and E.W. disclose support for this research from the Lundbeck Foundation (grant numbers R491-2024-1351, R302-2018-1799, R302-2018-2155 and R155-2013-16338), the Novo Nordisk Foundation (NNF18SA0035006, NNF24SA0092560 and NNF25SA0103965), the Wellcome Trust (UNS69906), the Carlsberg Foundation (CF18-0024), the Danish National Research Foundation (DNR94 and DNR174) and the University of Copenhagen (KU2016 programme). E.W. further discloses support for this research from Ferring Pharmaceuticals A/S and from Illumina. F.V.S. and M.S. disclose funding from Riksbankens Jubileumsfond (M 21-0018). F.V.S. discloses support from the Lundbeck Foundation (R491-2024-1351 and R322-2019-2610). R.M., A.L., R.S., O.I.G., V.I.B., C.B.R. and A.W.W. disclose support from the Social Sciences and Humanities Research Council of Canada (Major Collaborative Research Initiatives 410-2000-1000, 412-2005-1004, and 412-2011-1001, and Partnership Grant 895-2018-1004), and from the University of Alberta. O.I.G., V.I.B., and A.W.W. disclose support from a Russian Federation research grant (075-15-2019-866 “Baikal Siberia in the Stone Age: At the crossroads of the worlds”). B.D.S. discloses support from the President’s Postdoctoral Fellowship Programme (University of California). Y.W. discloses support from the Excellent Research Group Program for Tibetan Plateau Earth System (42588201). A.K.N.I. discloses funding from the OAK foundation (OFIL-20-095). R.M., A.T. and M.G.T. disclose support from the European Research Council Horizon 2020 programme (95183/COREX). M.G.T. discloses support from the European Research Council Horizon 2020 programme (865515/SUSTAIN). R.M. discloses further support from a research fellowship awarded by All Souls’ College, Oxford. S.V.V. and R.C.-D. declare no relevant funding.

**Author contributions** Ancient DNA laboratory work was undertaken by R.M., F.V.S., J.T.S., C.G. and L.V. Computational analysis was undertaken by R.M., F.V.S., B.D.S. and M.S., and supervised by M.S., as well as R.C.-D. for phylogenetic analysis specifically. Archaeological research contributing to this study was undertaken by A.W.W., A.L., R.S., E.J., O.I.G. and V.I.B. and directed by A.W.W. Osteological analysis specifically was undertaken by A.L. Analysis and modelling of the radiocarbon dates was undertaken by R.S. and C.B.R. Samples were curated by E.J., S.V.V., O.I.G. and V.I.B. Specific areas of expertise for interpretation and further analysis were provided by A.K.N.I. (immunology and pathology); A.T. and M.G.T. (modelling mortality profiles); Y.W. (ecology); R.C.-D. (computational phylogenetics). The initial draft writing was led by R.M., together with F.V.S. and A.K.N.I., with subsequent contributions from B.D.S., A.L., E.J., R.S., M.G.T., A.W.W., M.S. and E.W. All authors reviewed, commented on and approved the final version of the manuscript for submission. This project was undertaken in the context of the PhD research of R.M., which was conceived by A.W.W. and E.W. Research design for plague analysis was conceived by M.S. and E.W., with input from R.M. and F.V.S. E.W. initiated, led and was primarily responsible for supervising this research.

**Competing interests** The authors declare no competing interests.

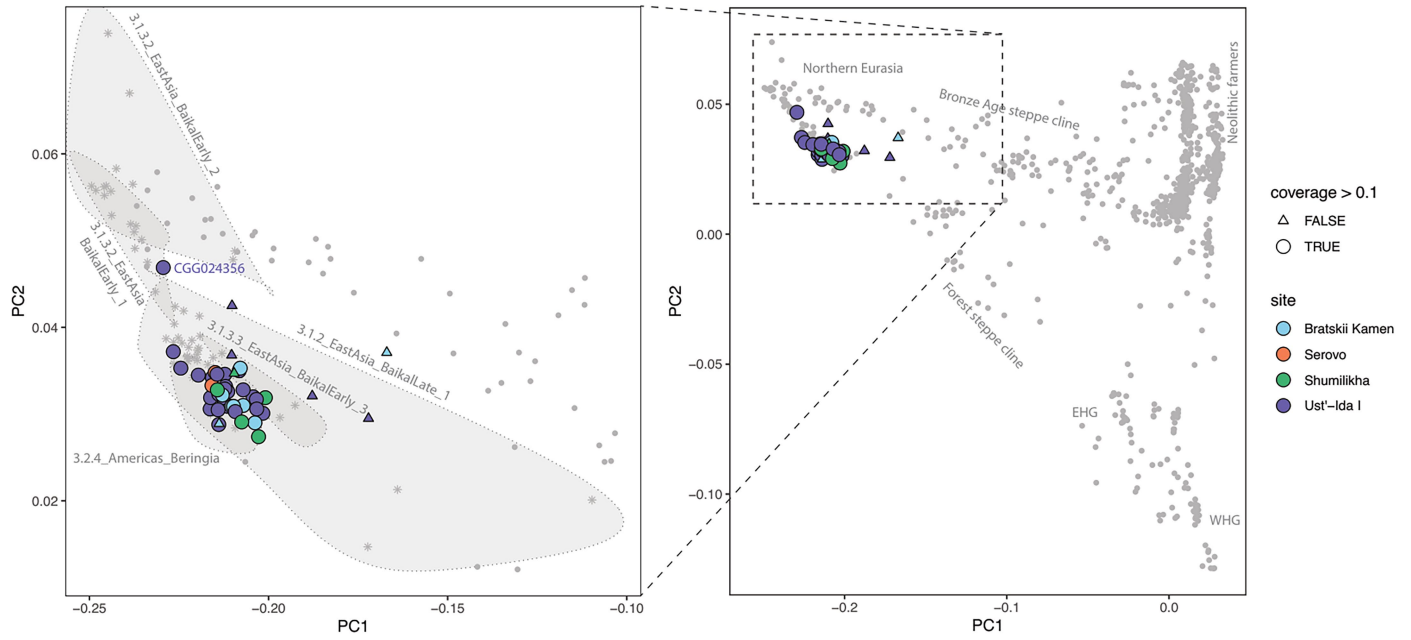
#### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-026-10540-5>.

**Correspondence and requests for materials** should be addressed to Ruairidh Macleod or Martin Sikora.

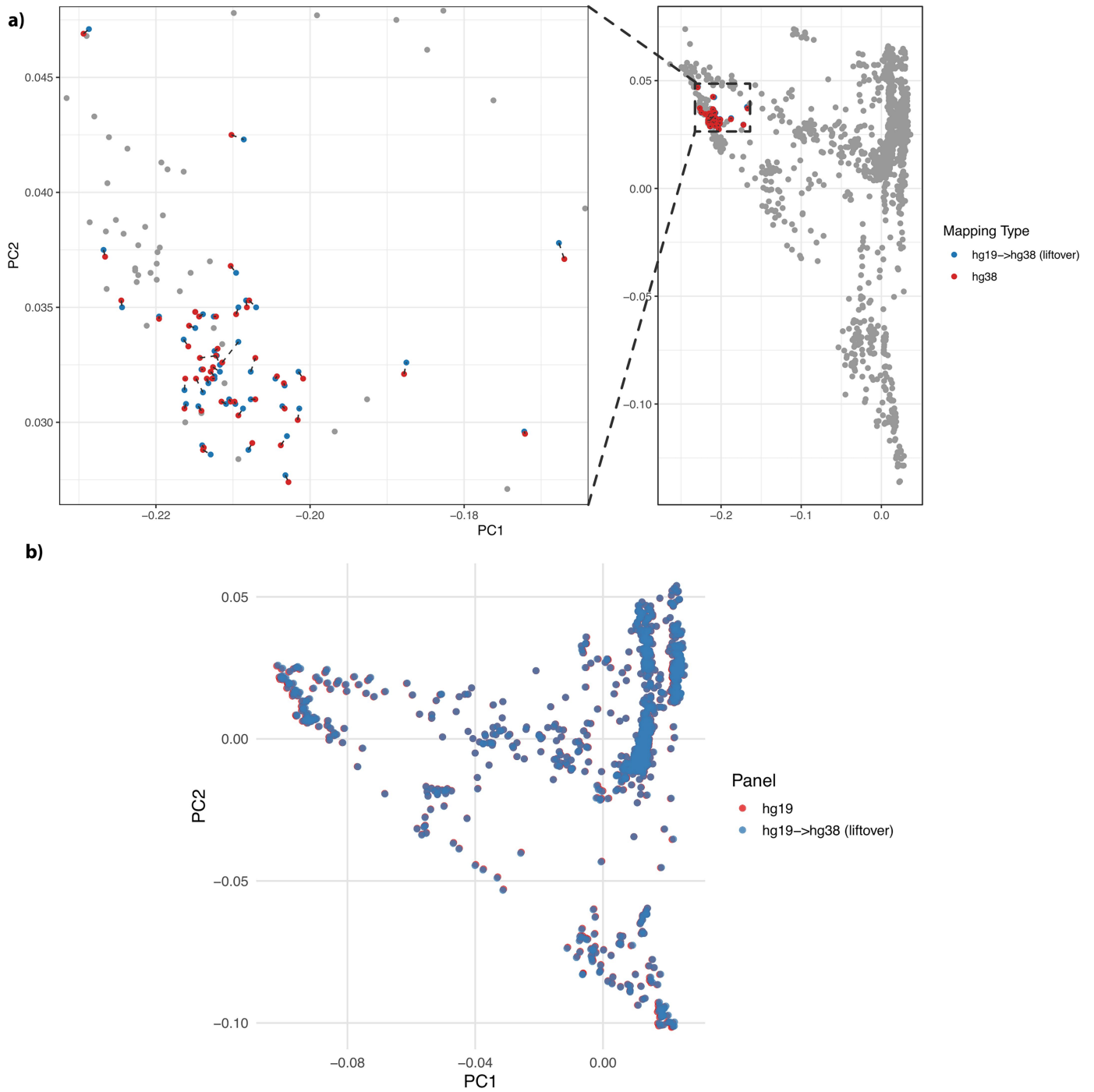
**Peer review information** Nature thanks Kevin Daly, Vanessa Villalba-Mouco and the other, anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



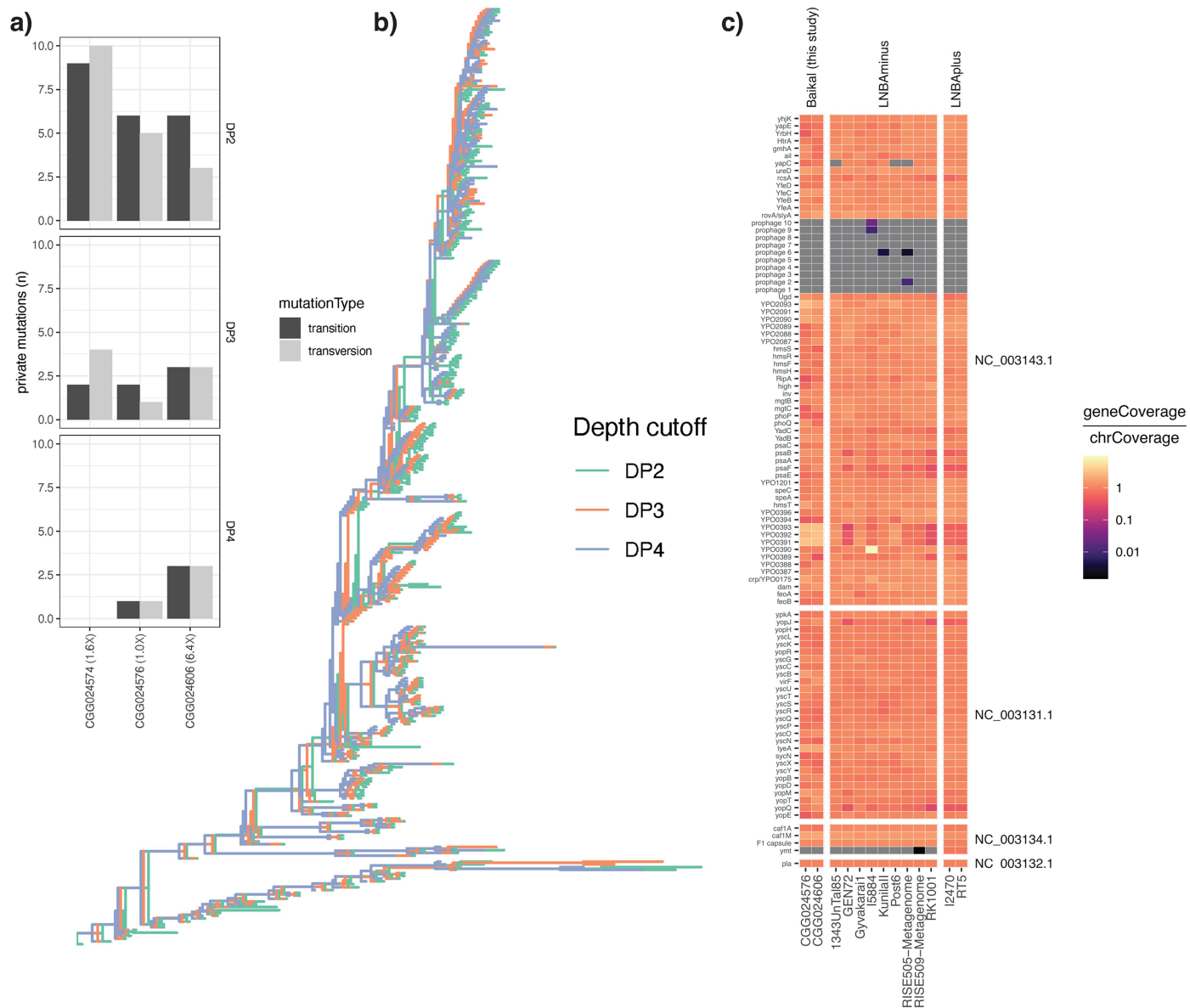
**Extended Data Fig. 1 | Principal Components Analysis (PCA) of ancient human genomes from Baikal hunter-gatherers studied here.** Ancient genomes for 46 ancient individuals (31 Ust'-Ida I; 8 Bratskii Kamen; 2 Serovo; and 5 Shumilkha) are projected onto a background PCA using as a reference the dataset from Allentoft et al.<sup>68</sup>. Reference data is coloured grey, while test

samples are colored by sample site. Low coverage samples are represented by triangles. On the left hand insert plot, the five populations with the highest similarity to our test samples have been highlighted. Further genetic ancestry analyses, including F-statistic tests, are described in Supplementary Note 2.



**Extended Data Fig. 2 | Comparing the effects of mapping to human genome builds hg19 and hg38.** a) PCA analysis comparing test samples mapped directly to hg38 (red color) and mapped to hg19 and lifted over to hg38 (blue color), using smartPCA. b) The effect of hg19 to hg38 liftover using smartPCA, shown

by procrustes analysis comparing the smart PCA space for our original panel mapped to hg37 and that same panel lifted over to hg38. Further details and results of comparisons (including comparisons of euclidean distances and F statistic tests) are provided in Supplementary Note 6.

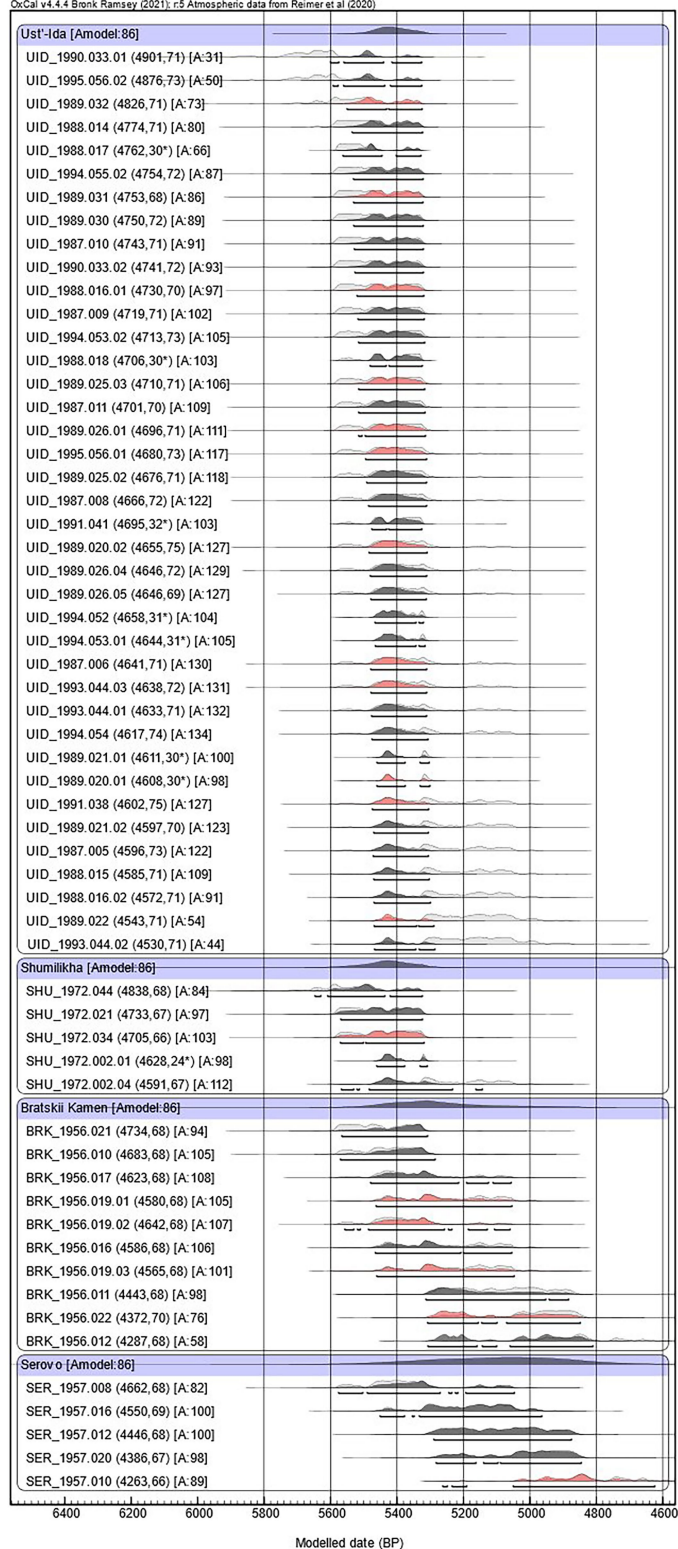


**Extended Data Fig. 3 | Genome analysis of *Yersinia pestis* sequences.**

a) Number of private mutations for each higher coverage plague genome using different depth cut offs of variant calling (DP2, DP3 and DP4 corresponding to minimum depth thresholds of 2x, 3x and 4x, respectively) stratified in transitions and transversions. b) Comparison of phylogenetic trees inferred with the same depth threshold values. Trees inferred using different depth thresholds for calling genotypes are overlaid on top of each other. c) Heatmap

of coverage for classic virulence-associated genes from *Y. pestis*. Colour gradient corresponds to the average depth of coverage of the specified gene, divided by the average depth of coverage of the entire chromosome. The two highest coverage Baikal samples, Shumilikha #34 (CGG024606) and Serovo #10 (CGG024576) are shown at the bottom alongside previously LNBA- and LNBA+ strains.

OxCal v4.4.4 Bronk Ramsey (2021); r5 Atmospheric data from Reimer et al (2020)



**Extended Data Fig. 4 | Radiocarbon dating results from remains studied here.** Site-based KDE models of all radiocarbon-dated individuals (or associated red deer tooth pendants, marked by asterisks) from the four Late Neolithic cemeteries on the Angara River with plague victims (indicated in red). Radiocarbon dates obtained from human remains are corrected for the freshwater reservoir effect (see Supplementary Note 4).

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	Not applicable
Data analysis	<p>CASAVA (v.1.8.2)            AdapterRemoval (v.2.0)            bwa aln (v.0.7.18)            Samtools (v.1.21)            Picard (v2.18.7)            BEDtools (v2.23.0)            pysam (<a href="https://github.com/pysam-developers/pysam">https://github.com/pysam-developers/pysam</a>)            ANGSD (v.0.940)            mapDamage2.0            haplogrep (v.2.4.0)            mutserve (v.1.3.0)            Smartpca (v. 18140)            hgLiftOver (<a href="https://genome.ucsc.edu/cgi-bin/hgLiftOver">https://genome.ucsc.edu/cgi-bin/hgLiftOver</a>)            bcftools (v.1.21)            GLIMPSE (v. 1.1.1)            IBDseq (v.r1206)            hapRoH (v.1)            KIN            Pathopipe (<a href="https://github.com/martinsikora/pathopipe/">https://github.com/martinsikora/pathopipe/</a>)</p>

KrakenUniq (v.0.5.8)  
 bowtie2 (v.2.5.4)  
 RAxML-NG  
 UShER  
 Pangenome Graph Builder (pgraph)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All genetic sequencing data generated in the course of this study will be available on publication, under ENA accession no. PRJEBXXXXXX.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

### Reporting on sex and gender

*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*

### Reporting on race, ethnicity, or other socially relevant groupings

*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status). Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.) Please provide details about how you controlled for confounding variables in your analyses.*

### Population characteristics

*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*

### Recruitment

*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*

### Ethics oversight

*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

### Study description

We analyse ancient genomic data from prehistoric human archaeological remains, to study the pathogens affecting them, in particular *Yersinia pestis*, the cause of plague; as well as biological kinship and ancestry of humans.

### Research sample

We studied the archaeological remains of individuals corresponding the Serovo and Isakovo periods of occupation of the Hunter-gatherer cemetery sites of Ust-Ida I, Shumilikha, Serovo and Bratskii Kamen, in the cis-Baikal region.

Sampling strategy	All samples from which ancient DNA data were generated are reported and included in the studied sample.
Data collection	Data collection for ancient DNA sequence data was made through the sequencing of DNA extracted from the dental cementum of molar or premolar teeth from archaeological skeletal remains; undertaken in dedicated clean lab facilities at the Lundbeck Foundation Centre for GeoGenetics (Copenhagen) and at the Institute of Archaeology, University College London (London).
Timing and spatial scale	We describe samples from four sites on the Angara river, in the cis-Baikal region (northwest of Lake Baikal, Russia). These sites consist of an earlier phase at Shumilikha (5580–5320 cal. BP) and Ust'-Ida I (5600–5320 cal. BP), and later phase at the sites of Bratskii Kamen (5475–5052 cal. BP) and Serovo (5290–4870 Cal. BP).
Data exclusions	No ancient DNA data were excluded where available for the sites and phases described.
Reproducibility	No experimental findings are reported.
Randomization	N/A.
Blinding	N/A.

Did the study involve field work?  Yes  No

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

### Methods

- | n/a                                 | Involvement in the study  |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms              |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern             |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                                   |

- | n/a                                 | Involvement in the study                        |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Palaeontology and Archaeology

Specimen provenance	Specimens were collected by researchers linked to the Baikal Archaeology Project, and curated by Andrzej Weber and Erin Jessup in this context.
Specimen deposition	Samples used for bioarchaeological analysis (here, teeth) are curated at present at the Department of Anthropology, University of Alberta, Canada; other archaeological remains are deposited at Scientific Research Centre "Baikalskii Region", Irkutsk State University, Russia; and Institute of Ethnography and Anthropology, Russian Academy of Sciences, Russia.
Dating methods	Radiocarbon dates were generated at the University of Oxford Radiocarbon Accelerator Unit facility, and calibrated using the OxCal and the IntCal curve.
<input checked="" type="checkbox"/>	Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.
Ethics oversight	The Universities of Cambridge and Copenhagen provided ethical approval for this study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Plants

### Seed stocks

*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*

### Novel plant genotypes

*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*

### Authentication

*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.*